Optical Coating Technology

Optical Coating Technology

Philip W. Baumeister



SPIE PRESS
A Publication of SPIE—The International Society for Optical Engineering
Bellingham, Washington USA

Library of Congress Cataloging-in-Publication Data

Baumeister, Philip.
Optical coating technology / by Philip Baumeister
p. cm. – (SPIE Press monograph ; PM137)
Includes bibliographical references and index.
ISBN 0-8194-5313-7 (hardcover)
1. Optical coatings. I. Title. II. Series.

TS517.2.B38 2004
681'.4—dc22 2003069358

Published by

SPIE—The International Society for Optical Engineering
P.O. Box 10
Bellingham, Washington 98227-0010 USA
Phone: +1 360 676 3290
Fax: +1 360 647 1445
Email: spie@spie.org
Web: http://spie.org

Copyright © 2004 The Society of Photo-Optical Instrumentation Engineers

All rights reserved. No part of this publication may be reproduced or distributed in any form or by any means without written permission of the publisher.

The content of this book reflects the work and thought of the author.
Every effort has been made to publish reliable and accurate information herein, but the publisher is not responsible for the validity of the information or for any outcomes resulting from reliance thereon.

Printed in the United States of America.

CONTENTS

Foreword / xix
Preface / xxi

Chapter 1 How coatings are used and integrated into optical systems

1.1 Fabrication of multilayer interference devices / 1-1
- 1.1.1 How do multilayers function? / 1-3
- 1.1.2 Nonoptical attributes of surface coatings / 1-4

1.2 Nomenclature and general properties / 1-5
- 1.2.1 Nomenclature / 1-5
- 1.2.2 Reflectance and transmittance / 1-6
- 1.2.3 Wavelength and frequency / 1-10
- 1.2.4 Angle shift / 1-11

1.3 Antireflection coatings / 1-11
- 1.3.1 Choice of an antireflection coating / 1-12
- 1.3.2 Antireflection coatings for the IR / 1-16
- 1.3.3 Antireflection coatings for fiber-optic communication devices / 1-18
- 1.3.4 Antireflection coatings at nonnormal incidence / 1-18
- 1.3.5 Out of bandwidth reflectance / 1-19
- 1.3.6 Specification of an antireflection coating / 1-19
- 1.3.7 Alternatives to vacuum-deposited antireflection coatings / 1-20

1.4 Spectral filtering and narrowband rejection / 1-21
- 1.4.1 Comparison of interference and absorption filters / 1-21
- 1.4.2 Selective absorbers / 1-22
- 1.4.3 Reflection filter / 1-22
- 1.4.4 Narrowband attenuator / 1-23

1.5 Filters with broad spectral bandwidth / 1-25
- 1.5.1 Introduction / 1-25
- 1.5.2 Edge filters / 1-26
- 1.5.3 Dichroic coating / 1-27
- 1.5.4 Coatings for the IR / 1-27
- 1.5.5 Effect of humidity / 1-27
- 1.5.6 Specification of an edge filter / 1-28
- 1.5.7 Edge filter for color control / 1-29
- 1.5.8 Thermal control coatings / 1-29
- 1.5.9 Regenerative coatings / 1-31
- 1.5.10 Coatings for solar cells / 1-31
- 1.5.11 Thermal control for satellites / 1-32
- 1.5.12 Short-wave pass – a case study / 1-33

1.6 Bandpasses
- 1.6.1 Applications / 1-34
- 1.6.2 Single-cavity bandpass filters / 1-34

1.6.3 Classes of bandpass filters / 1-36
1.6.4 Attributes of conventional bandpass filters / 1-38
1.6.5 Blocking of a bandpass / 1-42
1.6.6 Bandpass in convergent flux / 1-45
1.6.7 Attributes of a bandpass for wavelength multiplexing or wavelength demultiplexing / 1-48
1.7 Reflectors – used principally at normal incidence / 1-49
1.7.1 Introduction / 1-49
1.7.2 Criteria for the selection of reflectors / 1-50
1.7.3 Overview of reflectors / 1-52
1.7.4 Metallic reflectors / 1-53
1.7.5 Overcoated metals / 1-56
1.7.6 All-dielectric reflectors / 1-57
1.7.7 Reflectors containing metal layers / 1-63
1.7.8 Reflector for an optical waveguide / 1-64
1.7.9 A checklist for the specification of reflector / 1-65
1.8 Beamdividers, dichroics and polarizers / 1-65
1.8.1 Introduction / 1-65
1.8.2 Beamdividers / 1-71
1.8.3 Linear polarizers / 1-74
1.8.4 Dichroic reflectors / 1-78
1.8.5 Miscellaneous topics / 1-86
1.8.6 Coatings: costs and specifications / 1-89
1.9 Miscellaneous topics / 1-92
1.9.1 Clear aperture, jig marks and bevels / 1-92
1.9.2 Cementing of coatings / 1-94
1.9.3 Neutral density / 1-96
1.9.4 Absorbers – both selective and broadband / 1-96
1.9.5 Coatings for glass on buildings / 1-97
1.9.6 Interference photocathode / 1-99
1.9.7 Electrically conducting coatings / 1-100
1.9.8 Lateral variation of R or T / 1-102
1.9.9 Unusual coatings and systems / 1-104
1.9.10 Coatings for fiber optic communication systems / 1-105
1.10 Appendices
1.10.1 The creation of an environment for an optical coating when the impinging light is noncollimated / 1-106
1.10.2 Coating designs / 1-109

Chapter 2 Fundamentals

2.0 Overview of chapter / 2-1
2.1 Overview of intuitive approach / 2-1
2.1.1 Waves propagating in a nonabsorbing medium / 2-1
2.1.2 Properties of an absorbing medium / 2-3

2.1.3 Dispersion of the medium in which the wave propagates / 2-5
2.2 Reflection and transmission at an interface / 2-6
2.2.1 Overview and establishment of terminology / 2-6
2.2.2 Boundary conditions at an interface / 2-7
2.2.3 Fresnel coefficients at an interface / 2-9
2.2.4 Radiant reflectance and transmittance / 2-14
2.2.5 Analogy between admittance and an optical coating / 2-20
2.2.6 Addition of waves at an interface / 2-23
2.3 Phase shift upon reflection and node of the standing wave / 2-27
2.3.1 Definition of phase shift / 2-27
2.3.2 Standing waves normal to an interface / 2-28
2.3.3 Lateral standing waves / 2-30
2.3.4 Differential phase shift and reflectance vs angle of incidence / 2-30
2.3.5 Constraints on the reflection and transmission coefficients / 2-32
2.3.6 Conventions relating to the phase shift upon reflection / 2-33
2.4 Properties of a multilayer / 2-34
2.4.1 Introduction / 2-34
2.4.2 Radiant reflectance and transmittance of a single layer / 2-34
2.4.3 Use of the E^+E^- matrix / 2-47
2.4.4 Characteristic matrix of a homogeneous layer / 2-52
2.4.5 Recursion methods of computing the reflectance / 2-56
2.4.6 Comparison of computational methods / 2-58
2.5 Design concepts used at nonnormal incidence / 2-59
2.5.1 Introduction / 2-59
2.5.2 Optical thickness and effective thickness / 2-59
2.5.3 Effective index / 2-61
2.6 Aids to computation / 2-64
2.6.1 Introduction / 2-64
2.6.2 Matrix relationships / 2-64
2.6.3 Special layers / 2-69
2.6.4 Partitioning of a multilayer / 2-71
2.6.5 Admittance transformation layers / 2-75
2.6.6 Optimization / 2-75
2.6.7 Equivalent layers / 2-76
2.7 Properties of a stack with equal optical thickness layers / 2-83
2.7.1 Introduction / 2-83
2.7.2 Conventions for specifying thicknesses of layers / 2-83
2.7.3 Admittance of a stack with layers of equal optical thickness / 2-85
2.7.4 *R/T* polynomial / 2-86
2.8 Graphical aids to multilayer design / 2-90
2.8.1 Introduction / 2-90
2.8.2 Graphical presentation / 2-91
2.8.3 Vector addition of amplitudes / 2-91
2.9 Standing waves, net flux ratio and absorption / 2-95
2.9.1 Introduction / 2-95

2.9.2 Net flux ratio – definition and properties / 2-95
2.9.3 Methods of computing the net flux ratio / 2-97
2.9.4 Maximum net flux ratio and its attainment / 2-102
2.9.5 Computation of absorption using net flux ratio / 2-105
2.9.6 Loss in weakly absorbing layers / 2-107
2.10 Appendices – Propagation of electromagnetic waves / 2-112
2.10.1 Assumptions about waves and the medium in which they propagate / 2-112
2.10.2 Propagation of EM waves in isotropic homogeneous media / 2-113
2.10.3 Propagation in an inhomogeneous and anisotropic medium / 2-116
2.11 Appendices / 2-118
2.11.1 Dispersion of the medium / 2-118
2.11.2 Equivalent layers in terms of dimensionless parameters / 2-119
2.11.3 Synthesis of an equivalent layer / 2-121
2.11.4 Spectral bandwidth of a single-cavity bandpass / 2-122
2.11.5 Reflection coefficients in terms of sheet resistance / 2-125

Chapter 3 Thin films, the building blocks of multilayers

3.0 Purpose of this chapter / 3-1
3.1 Methods of depositing a thin film – a conceptual view / 3-2
3.1.1 Transport and attachment / 3-2
3.2 Deposition methods – hardware and procedures / 3-3
3.2.1 Physical vapor deposition / 3-3
3.2.2 Chemical processes / 3-5
3.2.3 Miscellaneous processes / 3-11
3.3 Overview of physical vapor deposition and film formation / 3-12
3.3.1 Conditions for thin film deposition / 3-12
3.3.2 Steps in film formation / 3-12
3.4 Process parameters influencing optical properties / 3-18
3.4.1 Method of deposition / 3-18
3.4.2 Vapor impingement angle / 3-19
3.4.3 Post deposition environment of a film / 3-20
3.4.4 Temperature shift of coatings / 3-23
3.5 Criteria for thin film material selection / 3-25
3.5.1 Introduction / 3-25
3.5.2 Producibility / 3-25
3.5.3 Refractive index, absorption and inhomogeneity / 3-25
3.5.4 Scatter and other losses / 3-28
3.5.5 Mechanical stress / 3-30
3.5.6 Thickness limits / 3-30
3.5.7 Adhesor layers / 3-30
3.5.8 Sacrificial layers / 3-32
3.5.9 Other criteria / 3-33

3.6 Survey of coating materials / 3-36
3.6.1 Introduction / 3-36
3.6.2 Dielectric coating materials / 3-36
3.6.3 Mixtures – mostly all-dielectric / 3-37
3.6.4 Absorbing materials / 3-38
3.6.5 Commonly used coating materials / 3-39
3.7 (Appendix) List of useful coating materials / 3-40

Chapter 4 Reflection reducing coatings

4.0 Introduction / 4-1
4.0.1 Development of the antireflection coating / 4-1
4.1 Antireflection coating design by computer optimization / 4-1
4.1.1 Can optimization produce an adequate antireflection coating? / 4-1
4.1.2 Design procedure / 4-1
4.1.3 Examples / 4-2
4.1.4 Antireflection coating that functions at nonnormal incidence / 4-5
4.1.5 Optimization or analytical design – which should be used? / 4-5
4.2 Design methods and evaluation of antireflection coatings / 4-6
4.2.1 Design methods / 4-6
4.2.2 Evaluation of an antireflection coating / 4-7
4.2.3 Electrical impedance mismatch and optical coating / 4-9
4.2.4 Symmetry of the reflectance vs wave number curve / 4-10
4.3 Multiple quarterwave and other narrowband designs / 4-11
4.3.1 Introduction / 4-11
4.3.2 Single layers / 4-11
4.3.3 Multiple layers of quarterwave optical thickness / 4-12
4.3.4 Two layers of unequal optical thickness / 4-13
4.4 All-dielectric antireflection coatings deposited upon metallic layers / 4-17
4.4.1 Introduction / 4-17
4.4.2 Design procedure / 4-17
4.4.3 Antireflection coating deposited upon a transition metal / 4-18
4.4.4 Antireflection coating deposited upon a semitransparent metal / 4-18
4.5 Coatings with broader spectral bandwidth – maximally flat designs / 4-21
4.5.1 Introduction / 4-21
4.5.2 Quasimaximally flat designs / 4-22
4.6 Coating with zero reflectance at two or more wavelengths / 4-25
4.6.1 Introduction / 4-25
4.6.2 Quarter-half coating / 4-25
4.6.3 Zeroes of reflectance – manifold solutions / 4-26
4.7 Chebyshev antireflection coatings / 4-30
4.7.1 Introduction: their attributes / 4-30
4.7.2 Procedure for design / 4-31
4.7.3 Comparison with maximally flat antireflection coating / 4-32

4.8 Step-up and step-down of admittance / 4-32
4.8.1 Introduction / 4-32
4.8.2 Quasimaximally flat quarter-half-quarter / 4-33
4.8.3 Quarter-half-quarter: wavelengths of zero reflectance / 4-33
4.8.4 Simulation of the bottom layer with two materials / 4-34
4.8.5 Design with an equivalent center layer / 4-35
4.9 Miscellaneous topics / 4-36
4.9.1 Simultaneous reduction of the reflectance in the visible and IR / 4-36
4.9.2 Simulated graded refractive index using equivalent layers / 4-38
4.9.3 Realization of coatings / 4-40
4.9.4 Design using linear programming / 4-41
4.9.5 Antireflection coatings for nonnormal incidence / 4-42
4.9.6 Antireflection coatings for multiple substrates / 4-46
4.9.7 Antireflection coating that matches optical cement to glass / 4-47
4.10 Appendix: Proofs, derivations and designs / 4-47
4.10.1 Two-layer unequal thickness / 4-47
4.10.2 Maximally flat / 4-47
4.10.3 Simultaneous and manifold solutions / 4-48
4.10.4 Quarter-half coating / 4-48
4.10.5 Reflectance zeros of Q Q and Q Q Q coatings / 4-48
4.10.6 Chebyshev antireflection coatings / 4-49
4.10.7 Zero reflectance of quarter-half-quarter antireflection coating / 4-50
4.10.8 Appendix – designs of antireflection coatings / 5-50
4.10.9 Additional information about antireflection coatings / 5-50

Chapter 5 Reflectors, edge filters and periodic structures

5.1 Introduction / 5-1
5.1.1 Examples – use of vector diagrams to calculate reflectance / 5-2
5.1.2 Example – output coupler for a carbon dioxide laser / 5-2
5.1.3 Nonquarterwave coatings / 5-4
5.1.4 Narrowing of the bandwidth / 5-5
5.2 Analysis of the basic period / 5-7
5.2.1 Introduction / 5-7
5.2.2 Stop band and passband – conceptual / 5-7
5.2.3 Quantitative definition of the stop band / 5-8
5.2.4 Spectral width of a stop band / 5-10
5.2.5 Higher order stop bands / 5-15
5.2.6 Splitting of the layers in a basic period / 5-18
5.3 "Single-stack" coatings / 5-19
5.3.1 Decision rules that produce an optimal quarterwave reflector / 5-19
5.3.2 Reflectance of a quarterwave stack – derived from admittance / 5-22
5.3.3 Rejection filters / 5-26
5.3.4 Blocking filters and reflectors / 5-28
5.3.5 Suppression of reflection peaks of higher order stop bands / 5-30

5.3.6 Spectral region of flattened reflectance / 5-37
5.3.7 Dual-band reflectors / 5-40
5.3.8 Performance of stacks at nonnormal incidence / 5-41
5.3.9 Minus filter / 5-44
5.3.10 Loss and absorption in a quarterwave stack / 5-45
5.4 Edge filter design / 5-58
5.4.1 Introduction / 5-58
5.4.2 Optimization / 5-58
5.4.3 Equivalent layers in edge filters / 5-59
5.5 Broadband reflectors and rejection filters / 5-68
5.5.1 Two-component stacks with overlapping stop bands / 5-69
5.5.2 Hot mirror without overlapping stop bands / 5-72
5.5.3 Broadband IR rejection filter / 5-72
5.5.4 Design techniques / 5-74
5.6 Phase shift upon reflection / 5-74
5.6.1 Introduction / 5-74
5.6.2 Quarterwave stacks and similar periodic media / 5-74
5.6.3 Broadband reflectors / 5-76
5.6.4 Phase shift due to thickness changes / 5-77
5.6.5 Phase shift upon reflection – nonnormal incidence / 5-77
5.7 Miscellaneous topics / 5-80
5.7.1 Overcoated metallic reflector / 5-80
5.7.2 Broadband reflector with relatively low reflectance / 5-83
5.7.3 IR reflector with visible transmittance / 5-84
5.7.4 TIR differential phase shift coatings / 5-85
5.7.5 Protective coatings / 5-86
5.8 Appendices / 5-86
5.8.1 Analysis of a periodic structure using matrices / 5-86
5.8.2 Amplitude reflection coefficient of a periodic structure / 5-89
5.8.3 Reflectance and phase envelopes / 5-90
5.8.4 Missing stop bands – solution for the refractive indices / 5-91
5.8.5 Coating designs / 5-93
5.8.6 Koppelmann's equation / 5-93

Chapter 6 Beamdividers and polarizers

6.1 Introduction / 6-1
6.1.1 Examples of beamdividers / 6-1
6.1.2 Terminology associated with polarizer design and usage / 6-5
6.2 Nonimmersed linear polarizers containing quarterwave layers / 6-7
6.2.1 Introduction / 6-7
6.2.2 Single layers on a germanium substrate / 6-9
6.2.3 Quarterwave stack linear polarizers – odd layer count / 6-13
6.2.4 Stack with low refractive index outer layer / 6-14
6.2.5 Applications / 6-15

6.3 Immersed linear polarizers / 6-15
6.3.1 Introduction / 6-15
6.3.2 Selection of the refractive indices / 6-16
6.3.3 Extension of the angular bandwidth / 6-21
6.3.4 Broadening of the spectral bandwidth / 6-25
6.3.5 Broadening of both the spectral and angular bandwidth / 6-25
6.3.6 Coating that suppresses the *s* reflectance / 6-26
6.4 Nonpolarizing nonimmersed designs / 6-29
6.4.1 Designs containing layers of nonquarterwave optical thickness / 6-29
6.4.2 Quarterwave layer thickness designs – introduction / 6-29
6.4.3 Analytical solution for a coating containing two refractive indices / 6-30
6.4.4 Designs containing three refractive indices / 6-32
6.5 Nonpolarizing immersed coatings / 6-34
6.5.1 Introduction / 6-34
6.5.2 Nonpolarizing – independent of angle / 6-34
6.5.3 Nonpolarizing at a single angle / 6-35
6.5.4 Tolerances of the refractive index and layer thickness / 6-39
6.5.5 Comparison of beamdividers: all-dielectric and containing silver / 6-39
6.6 Miscellaneous topics / 6-41
6.6.1 Joining prisms / 6-41
6.6.2 Linear retarder with 180° differential phase shift upon reflection / 6-42
6.6.3 References to beamdividers and polarizers in other chapters / 6-42
6.7 Appendices containing derivations / 6-42
6.7.1 Condition for the existence of a nonpolarizing angle / 6-42
6.7.2 Nonpolarizing, nonimmersed system – solution for the indices / 6-43
6.7.3 Three-component immersed system – solution for the indices / 6-43
6.7.4 Solution for the reciprocal index squared / 6-44
6.7.5 Two-component immersed system – solution for the indices / 6-4
6.8 Appendices containing multilayer designs / 6-44
6.8.1 Design: polarizer with 52 layers / 6-44
6.8.2 Design: polarizer with 67 layers / 6-44
6.8.3 Design: linear polarizer / 6-45
6.8.4 Design: polarizer with 98 layers / 6-45
6.8.5 Design: polarizer with 64 layers / 6-45
6.8.6 Design: 50%-50% beamdivider / 6-45
6.8.7 Design: 50%-50% beamdivider / 6-45
6.8.8 Design: all-dielectric beamdivider with minimal polarization splitting / 6-46
6.8.9 Design: beamdivider with a single silver layer / 6-46

Chapter 7 All-dielectric bandpass filters

7.1 Introduction / 7-1
7.1.1 Usage of all-dielectric filters / 7-1

7.1.2 Filter classification / 7-2
7.1.3 Design methods / 7-2
7.1.4 Mutual-passband filter / 7-3
7.1.5 Periodic structures without halfwave layers / 7-4
7.2 Control of the spectral bandwidth / 7-6
7.2.1 Spectral bandwidth control – alteration of the thickness of the spacer / 7-7
7.2.2 Reflectance of the reflectors / 7-7
7.2.3 Phase dispersion / 7-8
7.3 Periodic-structure bandpass design method / 7-11
7.3.1 Periodic-structure method of bandpass design –overview / 7-12
7.3.2 Periodic-structure method of bandpass design – procedure / 7-13
7.4 Filter design using two components / 7-14
7.4.1 Selection of the periodic structure / 7-14
7.4.2 Use of a single-section AR / 7-15
7.4.3 Use of a three-section AR / 7-16
7.4.4 Comparison of bandpasses / 7-17
7.4.5 Use of a two-section AR / 7-17
7.4.6 Phase-dispersion narrowing of the passband / 7-18
7.4.7 Bandpass with 7% spectral bandwidth / 7-18
7.5 Periodic structures containing three materials / 7-20
7.5.1 Design procedure / 7-20
7.5.2 Design example – 7% spectral bandwidth / 7-21
7.5.3 Design example – a bandpass with inferior offband rejection / 7-23
7.5.4 Use of higher order spacers to narrow the spectral bandwidth / 7-24
7.5.5 Improvement of the average transmittance of the passband / 7-24
7.5.6 A bandpass viewed as a multiple-cavity filter / 7-25
7.5.7 "Cavity" – possible meanings of this word / 7-26
7.5.8 Microwave and periodic-structure design methods – a comparison / 7-26
7.6 Microwave design method / 7-28
7.6.1 Attributes of a bandpass / 7-28
7.6.2 A prototype multiple-cavity filter / 7-29
7.6.3 Choice of the number of cavities – introduction / 7-37
7.6.4 Control of the spectral bandwidth of a multiple-cavity bandpass / 7-39
7.6.5 Design procedure / 7-44
7.7 Examples of conventional bandpass design / 7-48
7.7.1 Offband rejection / 7-48
7.7.2 Two-cavity designs / 7-49
7.7.3 Discarding of halfwaves / 7-51
7.7.4 Three-cavity designs / 7-52
7.7.5 Four-cavity bandpasses / 7-54
7.7.6 Five-cavity bandpasses – specifications / 7-57
7.7.7 A six-cavity bandpass – mixed-cavity design / 7-58
7.7.8 Seven-cavity bandpass – homologous cavity / 7-59

7.7.9 Bandpasses containing more than seven cavities / 7-59
7.8 Bandpasses for optical fiber communication / 7-59
7.8.1 Choice of the number of cavities / 7-60
7.8.2 Three-cavity bandpass / 7-60
7.8.3 Four-cavity bandpasses / 7-61
7.8.4 Design of five-cavity bandpasses / 7-63
7.8.5 A comparison of the results / 7-66
7.8.6 Use of a composite spacer layer / 7-66
7.8.7 Temperature shift of the passband center wavelength / 7-67
7.9 Additional topics / 7-68
7.9.1 Phase conjugate bandpass filter / 7-68
7.9.2 Miscellaneous design methods / 7-69
7.9.3 Cavity shifting / 7-69
7.9.4 Effects of absorption upon passband transmittance / 7-72
7.9.5 Production of a bandpass – effects of layer thickness errors / 7-75
7.9.6 Performance of a bandpass at nonnormal incidence / 7-76
7.9.7 Single-cavity bandpass filter used as a linear polarizer / 7-78
7.9.8 Selection of the number of cavities / 7-79
7.9.9 Bandwidth and spectral slope of a maximally flat prototype / 7-82
7.9.10 Design of a bandpass with a large number of cavities / 7-84
7.9.11 Pulse propagation through a WDM bandpass filter / 7-89
7.10 Miscellaneous topics and appendices / 7-91
7.10.1 Design of a prototype bandpass with Chebyshev transmittance in its passband / 7-91
7.10.2 Shape factor of a prototype WDM bandpass / 7-95
7.10.3 Designs of bandpass filters and other coatings / 7-96
7.10.4 Standing wave ratios of the reflectors of a prototype bandpass with a Chebyshev passband transmittance – the ANSI C source code / 7-100
7.10.5 Historical notes / 7-102

Chapter 8 Coatings that contain absorbing layers

8.1 Introduction / 8-1
8.1.1 Uses of absorbing films / 8-1
8.1.2 Generic absorbing materials / 8-1
8.1.3 Absorption controlled by the standing wave electric field / 8-1
8.2 Bandpass filters – general properties / 8-3
8.2.1 Overview of bandpass filters / 8-3
8.2.2 Trade-off between maximum transmittance T_{max} and offband rejection / 8-4
8.2.3 Simple filters containing one or two metal layers / 8-4
8.2.4 One-M bandpass filter / 8-4
8.2.5 Properties of the Two-M bandpass filter / 8-11
8.3 Design procedures for metal-dielectric bandpass filters / 8-17
8.3.1 Recommended metals of metal-dielectric bandpass filters / 8-17

8.3.2 Trade-off between spectral bandwidth and peak transmittance / 8-18
8.3.3 Overview of multiple-cavity bandpass filter design / 8-19
8.3.4 Guidelines for the design of multiple-cavity bandpass filter / 8-20
8.3.5 Bandpass filter design procedures / 8-21
8.4 Bandpass filter design examples / 8-21
8.4.1 Introduction / 8-21
8.4.2 Design procedure – the number of cavities is even / 8-22
8.4.3 Design procedure – the number of cavities is odd / 8-24
8.4.4 General discussion / 8-26
8.4.5 Design of an asymmetrical three-M filter / 8-26
8.5 Dark mirror absorber / 8-30
8.5.1 Characterization of the metal / 8-30
8.5.2 Optimization as a design method / 8-31
8.5.3 Analytical design procedure / 8-31
8.6 Reflectors / 8-32
8.6.1 Reflector with different reflectance from air and substrate sides / 8-32
8.6.2 Reflection filter / 8-33
8.7 Beamdivider containing silver / 8-33
8.8 Neutral density coatings / 8-34
8.8.1 Effects of inserting the coating into a beam of high irradiance / 8-34
8.8.2 Precise control of absorbance / 8-35
8.8.3 Coating with relatively low reflectance / 8-35
8.8.4 Constant absorbance vs wavelength / 8-35
8.9 Miscellaneous topics / 8-36
8.9.1 Amplitude reflection coefficient of an overcoated matching stack / 8-36
8.9.2 Optical constants of titanium / 8-37
8.9.3 Historical notes / 8-37

Chapter 9 Coating deposition

9.1 Introduction / 9-1
9.1.1 Overview of the vacuum evaporation process / 9-1
9.2 Cleaning of a substrate prior to coating it / 9-2
9.2.1 Removal of particulates on the substrate by mechanical means / 9-2
9.2.2 Cleaning of the substrate in liquids / 9-2
9.2.3 Drying of the substrate / 9-3
9.2.4 Chemical etching of the substrate surface / 9-3
9.3 Tooling, initial pumpdown, ion bombardment and heating of substrates / 9-3
9.3.1 Tooling – introduction / 9-4
9.3.2 Pumpdown of the vacuum coating tank / 9-5
9.3.3 Exposure of the substrate to a gas discharge / 9-5
9.3.4 Heating of the chamber and substrates / 9-6
9.4 Thin film deposition / 9-9
9.4.1 Introduction / 9-9

9.4.2 Thermal evaporation / 9-13
9.4.3 Electron beam heating of evaporant / 9-14
9.4.4 Ion beam sputtering as method of depositing coatings / 9-17
9.4.5 Introduction of gas during the deposition of a coating / 9-18
9.4.6 Rate control of evaporation sources / 9-21
9.4.7 Pinholes in coatings / 9-21
9.4.8 Procedures and hardware for deposition of layers / 9-22
9.5 Collection of the evaporant upon the substrates / 9-23
9.5.1 Introduction – fundamental equations / 9-23
9.5.2 Overview of methods of controlling layer thickness / 9-26
9.5.3 Molecular intensity distribution / 9-27
9.5.4 Layer thickness distribution on a flat nonrotating plate / 9-29
9.5.5 Single rotation of the tooling containing the substrates / 9-30
9.5.6 Planetary rotation of the plano tooling / 9-35
9.5.7 The use of masking to achieve thickness uniformity, single rotation / 9-38
9.5.8 A comparison of single rotation and planetary rotation / 9-41
9.5.9 Use of masking to produce a nonuniform thickness distribution / 9-44
9.6 The control of layer thickness during deposition / 9-45
9.6.1 Non-optical methods / 9-45
9.6.2 Overview of "optical" monitoring / 9-46
9.6.3 The hardware of an optical monitor / 9-49
9.6.4 Indirect monitoring in reflection / 9-53
9.6.5 Composite direct monitoring / 9-64
9.6.6 Composite, direct monitoring of a bandpass filter / 9-65
9.6.7 The optical monitor as a process control instrument / 9-71
9.7 Mechanical stress in optical coatings / 9-71
9.7.1 Overview / 9-71
9.7.2 Measurement of the mechanical stress / 9-72
9.7.3 Control of the mechanical stress / 9-74
9.8 Appendices / 9-75
9.8.1 Planetary motion – equations / 9-75
9.8.2 Design of an edge filter / 9-76
9.8.3 Thickness uniformity of a WDM bandpass – introduction / 9-77
9.8.4 Composite optical monitoring – termination at a maximum or minimum / 9-78

Chapter 10 Miscellaneous topics

10.0 Overview of chapter / 10-1
10.1 Graphical aids to multilayer design / 10-1
10.1.1 Use of graphics / 10-1
10.1.2 Graphically presented parameters / 10-2
10.1.3 Example – an absorbing stack / 10-5
10.1.4 Choice of method / 10-6

10.2 Optimization / 10-7
10.2.1 Introduction / 10-7
10.2.2 Mode of operation – correction / 10-7
10.2.3 Mode of operation – minimization / 10-8
10.2.4 Design procedure / 10-9
10.2.5 Miscellaneous topics / 10-12
10.2.6 Alternative starting designs / 10-14
10.2.7 Further exploration of the subject / 10-15
10.2.8 Appendix – design of edge filters / 10-15
10.3 Overall transmittance of an array of coated objects / 10-16
10.3.1 Introduction / 10-16
10.3.2 Effects of illumination with non-monochromatic light / 10-16
10.3.3 Equations for reflectance and transmittance – incoherent illumination / 10-20
10.3.4 Example of an "element" / 10-22
10.3.5 Application to the transmittance of a single slab / 10-23
10.3.6 Examples / 10-24
10.3.7 Transmittance in convergent flux / 10-26
10.4 Performance of coatings – their optical characteristics / 10-29
10.4.1 Introduction / 10-29
10.4.2 Specular and diffuse measurements / 10-30
10.4.3 Spectrophotometry – a brief overview / 10-31
10.4.4 Absorption / 10-35
10.4.5 Loss / 10-36
10.4.6 Other optical tests / 10-36
10.5 Performance of coatings and their non-optical characteristics / 10-36
10.5.1 Adhesion / 10-37
10.5.2 Moderate abrasion / 10-37
10.5.3 Severe abrasion / 10-37
10.5.4 Humidity / 10-37
10.5.5 Salt fog / 10-37
10.5.6 Temperature shock / 10-37
10.5.7 Specialized tests – fungus and sand erosion / 10-38
10.5.8 Sequence of the tests / 10-38
10.6 Phase relations in multilayers / 10-38
10.6.1 Introduction / 10-38
10.6.2 Transmissive phase retardance / 10-39
10.6.3 Reflective phase shift / 10-40
10.6.4 Constraints upon the phase shifts / 10-45
10.7 The influence of a coating upon a transmitted or reflected wavefront / 10-46
10.7.1 Introduction / 10-46
10.7.2 Transmitted wavefront / 10-46
10.7.3 Phase of the reflected beam / 10-51

Chapter 11 References to the literature

Chapter 12 Notation and definitions of terminology

12.1 Symbols – general comments / 12-1
12.2 Symbols / 12-1
12.3 Glossary of terminology / 12-9

Chapter 13 Index

FOREWORD

[*Publisher's Note:* Philip Baumeister passed away during the production of this book. In consideration of his wishes, we are publishing the manuscript essentially in the form it existed the time of his death, with the addition of this foreword.]

This book represents the fruits of Professor Philip Baumeister's experience teaching a class entitled Optical Thin Film Technology over a period of twenty-five years. This class was offered as a short course by the UCLA Extension Program and was taught on the UCLA Campus in Los Angeles, California. The class ran over the period of one week and consisted of ten lectures.

This course was an extremely successful offering by UCLA and earned the distinction of being one of the longest running short courses in the UCLA Engineering Extension catalog. The goal of this class was to provide the students with both a depth of understanding and the breadth of knowledge to tackle important topics in the field of thin film technology. The extensive notes presented here captures the essence of these lectures and summarizes the various topics covered in this class. Professor Baumeister constantly updated these notes in order to reflect current topics and to cover key issues faced by those designing, producing and/or using optical thin film technology.

An element of the success of this short course was based on the limited training options available to scientists and engineers working in the field of optical thin film technology. Optical thin films are an enabling technology whose applications cross many boundaries and are continually finding new applications. The optical coating industry tends to be diffuse and only few companies possess the resources necessary to provide in-house training. Courses such as this provide a valuable aid to those in the field.

Professor Baumeister's design expertise and breadth of thin film knowledge was well established. His interests in thin films technology began in graduate school in 1955 at the University of California, Berkeley. He worked with Professor John Strong on the forefront of optical coating technology. Professor Baumeister was one of the pioneers in applying computer analysis and computer optimization techniques to the design and analysis of optical thin films.

While these notes were made available to the attendees of the short course, others became interested in obtaining additional copies. Prodded by the requests for these notes from those individuals not able to take the class, Professor Baumeister had entertained plans of self-publishing this volume. However, these plans were never realized. The struggle to find the right venue to publish these notes was complicated by Professor Baumeister's desire to continually update the notes and to keep these notes timely. These notes were last edited in the summer and fall of 2003.

The information in this book represents a resource that has no comparison. While the presentation is terse, the data presented can be extremely valuable to the practicing technologist. Information in this book is also widely backed up with citations to patents and the published literature. Many questions can be answered using this book, such as, what are the conventions for a given analysis formalism? Or, what other design approaches have been tried for this application?

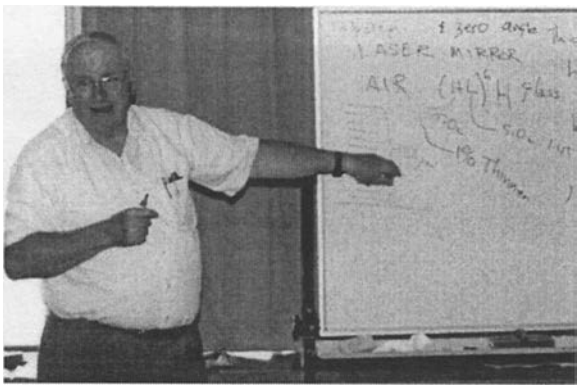
The book is largely organized around topics associated with functions of optical thin film performance. Chapter 1 provides an overview of the many applications and functionalities that optical thin films can provide when implemented in optical systems. Chapter 2 reviews the fundamentals of thin film analysis and presents various methods to calculate optical multilayer performance. Chapter 3 surveys thin film deposition methods, processes, and materials, while Chapter 4 provides a thorough understanding of the design of ARs for a wide variety of applications. Chapter 5 provides design approaches for multilayer reflectors, edge filters, and other periodic structures. Chapter 6 discusses designs and topics associated with beam dividers for both polarizing and non-polarizing applications. Chapter 7 covers the design and analysis of sophisticated dielectric bandpass filters, and Chapter 8 provides information associated with designing thin film performance using absorbing multilayers. Chapter 9 reviews concepts and the practical aspects of implementations associated with methods of producing thin films, and Chapter 10 collects and presents topics ranging from optimization of a coating design and phase relations in multilayer films to measurement and characterization of thin films. This is followed by the extensive list of references in Chapter 11, the summary of definitions and notations in Chapter 12 and the index in Chapter 13.

I use this book extensively in my daily routine as a practicing thin film engineer and hope that others will find it a valuable resource.

Ric Shimshock
Sunnyvale, California
December 2003



Philip Baumeister during his tenure at the University of Rochester (1959-1978).



Phil on a recent consulting trip to Japan.

PREFACE

This book is intended to instruct a technologist in the design and fabrication of thin film coatings that function by means of optical interference. Two years or more of college level instruction in the physical sciences is a desirable prerequisite for the reader. The book is an outgrowth of a course that was taught at the UCLA Extension from 1978 to 1999.

What this book is not: (1) This book is *not* a survey of the literature. The exigencies of time have made it impossible to reference all previous publications on all subjects. The author apologizes to those deserving authors whose works are not mentioned. (2) This book is *not* a catalogue from which a reader may select coatings and procure them. In Chap. 1, reflectance and transmittance curves of coatings from various manufacturers merely show what that manufacturer produced *at one time*. That product may, or may not still be in production and if it is, its performance could easily be far superior to that shown in this book. In other instances the manufacturer that produced certain coatings no longer exists. Commercial vendors do not assume any product liability by virtue of the presentation of data from their sales literature.

What is the author's right to speak on this subject? He began his study of optical coatings at the University of California (Berkeley) Physics Department in 1955. Later he instructed graduate students in coating fabrication and design at the University of Rochester. He served as Chief Scientist at OCLI (Santa Rosa, California) for a half decade, where he supervised a group that developed new processes and automated coating chambers. He was employed as a production engineer in the coating department of Coherent, Inc. and as a Senior Project Engineer with Deposition Sciences, Inc., in Santa Rosa, California. He is now a consultant, dwelling in the bucolic rural solitude of Sebastopol, California.

The preparation of this lucubration spans several decades and its chief benefit to the author has been the mental discipline demanded to organize thoroughly and write tersely. The desideratum could be Nobel Laureate Chandrasekhar's statement: "I don't want to be trivial." (Quoted in *The Christian Science Monitor*, 2 December 1983).

The stylist conventions of the American Institute of Physics are used, as are SI units where possible, although a few non-SI "Torr" pressure units appear. Radiometric terminology is precise; intensity is flux/solid angle and irradiance is flux/unit area. A few unfamiliar terms such as *dereflect* and *molecular intensity* are defined before their use.

Conversations with Oded Arnon, Verne Costich, Jay Eastman, Douglas Harrison, Erik Krisl and Doug Smith have been useful in clarifying some of the concepts in this book. The author is grateful to Konstantin V. Popov, Barbara Russell and Lita Holleman for their work in making the mathematical text and written text more accurate. Some of the clever drawings of Chap. 3 bear the stamp of George Russell. In the editorial work involved in getting the book into its final form, the contributions of Ric Shimshock and my three daughters, Dr. Nancy Charlotte Baumeister, PhD, Ms. Lynn Karin Baumeister and Ms. Carol Lisa Baumeister, are gratefully acknowledged.

My twelve days of instruction in the class of Joseph Frederick Ware, C.S.B., laid the foundation for this book.

Chapter 1

How coatings are used and integrated into optical systems

This chapter is intended for both users and manufacturers of optical coatings. Although the former need not be familiar with multilayer design, it is nevertheless useful to be aware of how coatings are selected and to gain an overview of their capabilities.

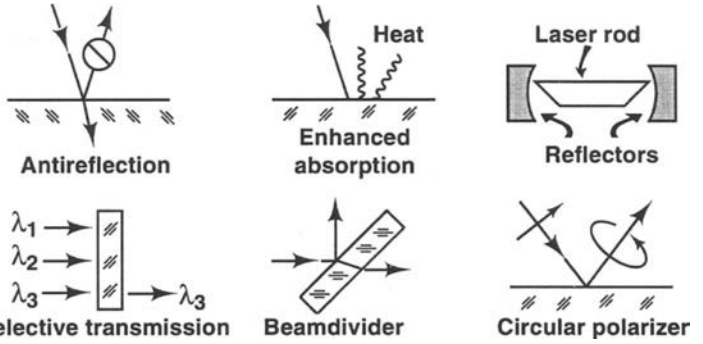


Fig. 1-1 Functions of multilayers: reduction of reflection at an interface, absorption, reflectors in a laser cavity, spectral filtering, beamdivider and circular polarizer.

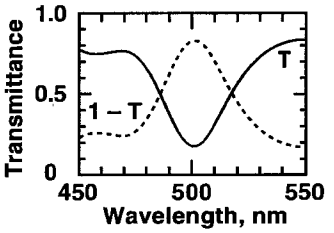


Fig. 1-2 Transmittance of a multilayer plastic sheet. After Mearl [88].

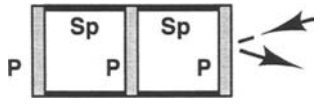


Fig. 1-3 Cross section of a resonant reflector, consisting of three circular plates with plane parallel surfaces (*shaded* P) and ring spacers (*solid* black Sp).

Figure 1-1 delineates functions of multilayer coatings, such as the reduction of reflectance at a surface, enhancement of its absorption or selective transmittance of some wavelengths in preference to others. Coatings also alter the state of polarization of a beam or divide it into several paths. Multilayer reflectors are widely used in laser cavities.

1.1 Fabrication of multilayer interference devices

Most of these devices are constructed by laying down relatively thin layers on a surface. At the time of this writing in the year 2000, physical vapor deposition, as described in Chap. 9, is the most widely used method of deposition. Chap. 3 discusses other deposition methods. Other optical interference devices deserve mention:

(a) Thin sheets of plastic containing 250 alternate layers of polystyrene (refractive index 1.6) and a lower index material are co-extruded [Alfrey 69, 73] to form a periodic structure exhibiting strong interference color. The spectral transmittance/reflectance of such material appears in Figs. 1-2 and 1-62.

(b) An ultraviolet laser forms a standing wave pattern in a plastic. The UV induced polymerization changes its refractive index in a periodic fashion, thus forming a narrowband reflector [Kaminow 71].

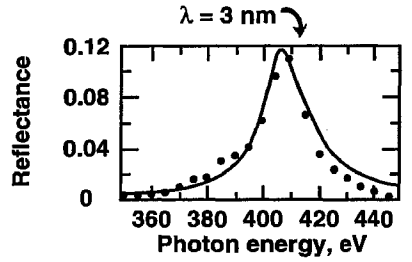


Fig. 1-4 Theoretical reflectance (*solid*) at an angle of 67° of X-ray reflector consisting of alternate layers of cobalt and carbon. Measured maximum R is 7%. Deposited by Spiller [76]. Illustration from Bartlett [84].

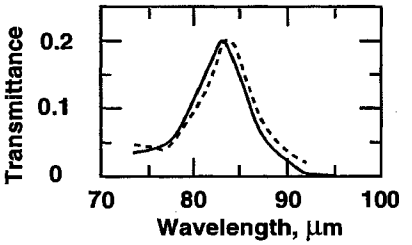


Fig. 1-5 Measured $T(\lambda)$ of single-cavity bandpasses consisting of four alternate layers of cesium iodide and germanium deposited upon a silicon substrate. After S. D. Smith [68].

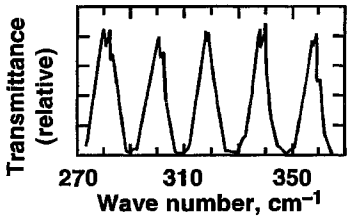


Fig. 1-6 Measured $T(\lambda)$ of an IR bandpass fabricated by laminating silicon slabs between organic materials. Wave number of 310 cm^{-1} corresponds to a wavelength of 322 μm. After Costich [90].

(c) A potassium chlorate crystal is grown so that its refractive index varies periodically, due to circular dichroism [Strong 61]. Reflection peaks are produced in the blue part of the spectrum.

(d) Hummingbird feathers and insect scales are often brilliantly colored due to a periodic variation in refractive index [Greenewalt 60; Wright 64; Steinbrecht 85].

(e) Resonant reflectors have narrowband reflectance peaks as high as 98% [Wiggins 71; Laser Optics 69]. As illustrated in Fig. 1-3, they are fabricated from slabs of bulk material rather than evaporated films. In their use inside laser cavities, their damage thresholds are considerably larger than those of evaporated structures.

(f) Cholesteric liquid crystals have a helical structure that modulates periodically the refractive index. Circularly polarized light is strongly reflected [Meier 75].

(g) The Lippmann photographic process records color by means of standing waves in a photographic emulsion. Wood [34] shows a cross section of a plate that manifests 250 undulations in its refractive index.

Interference devices have been fabricated over a wide range of wavelengths from the ultraviolet UV to the infrared IR. As shown in Fig. 1-4, a multilayer reflector with a reflectance as high as 7% at λ = 3 nm has been produced using alternate layers of carbon and cobalt [Spiller 76; Haelbich 76] or a tungsten-cobalt alloy [Spiller 78].

A bandpass filter for use at wavelengths as long as 75 μm was produced by evaporating alternate layers of cesium iodide and germanium, as depicted in Fig. 1-5 [Smith 68; Seeley 73]. Alternatively, the bandpass (see Fig. 1-6) was fabricated by laminating silicon slabs between organic materials [Costich 90]. A spectral filter at an even longer wavelength consists of alternate slabs of titania and polystyrene in a microwave

waveguide [Seeley 62]. Both Costich [90] and Iwasaki [70] constructed IR filters composed of layers of extruded polyethylene.

1.1.1 How do multilayers function?

Optical interference causes a coating to reflect more strongly at a specified wavelength and not at others. This is illustrated by the single layer in Fig. 1-7. A light wave (the incident beam *a*) impinges upon the layer and a fraction of that wave reflects at the air-film interface — this is (*c*).

The light and dark bands represent maxima and minima of the amplitude of that wave. A portion of the wave also penetrates the layer and reflects at the film-substrate interface — this is (*b*). That reflected wave then emerges into air at (*d*) and, in this example, is either in phase or out of phase with wave (*c*).

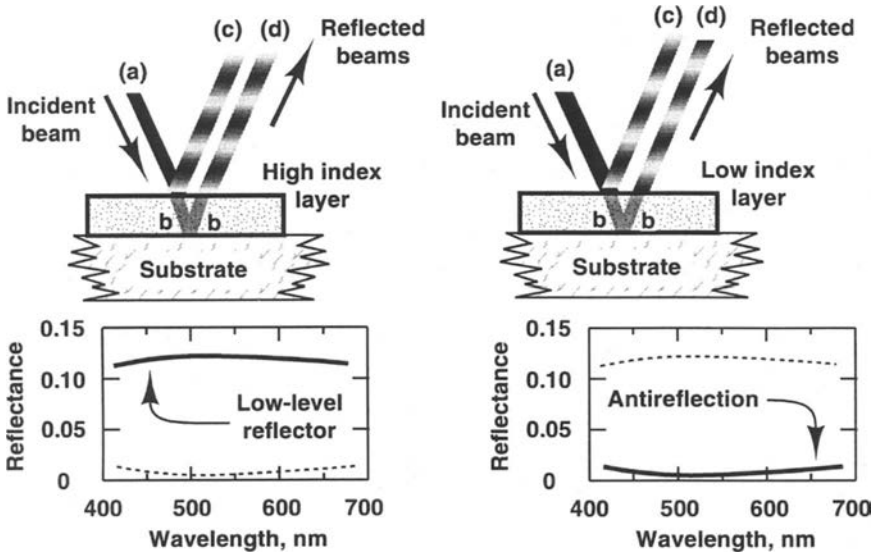


Fig. 1-7 Showing a light wave (*a*) incident upon a single layer of index 1.85 (*left panel*) or 1.38 (*right panel*) deposited upon a substrate of index 1.65. Reflected beam consists of wave (*c*) reflected at top air-film interface and (*d*) reflecting at (*b*) at film-substrate interface. Variable shading of reflected beams represents the wavelength λ. The distance from "light-to-light" or "dark-to-dark" is λ.

For the left panel of Fig. 1-7, the layer is of *higher* refractive index than the substrate. The (*c*) and (*d*) waves are *in phase* and a maximum reflectance results. This is manifested as a 12% peak reflectance in the graph. This is called "constructive interference" in books on physical optics.

The right panel depicts a layer of lower refractive index. The wave reflected from the air-film interface and the wave reflected from the film-substrate interface are 180° out of phase. Physical optics texts term this "destructive interference." The result is that the reflectance is lowered from 6% (for the uncoated substrate) to nearly zero.

The conclusions in the last two paragraphs depend upon the thicknesses of the layers relative to the wavelength. The optical thicknesses of the layers in Fig. 1-7 are quarterwaves at 530 nm, which is one of the reasons a maximum (or minimum) reflectance occurs at that wavelength. The reflectance curves could be shifted to a

shorter wavelength by reducing layer thicknesses. A thickening of a layer shifts the reflectance maxima to longer wavelengths.

An example illustrates the importance of the ordering of the layers and their relative thicknesses. The left panel of Fig. 1-8 shows a reflector with 35% peak reflectance. It consists of alternate layers of alumina, zirconia and magnesium fluoride — each with an optical thickness of a quarterwave. Two layers are interchanged and the thickness of the center layer is increased to a halfwave — the result is reduction of the reflectance over a broad bandwidth.

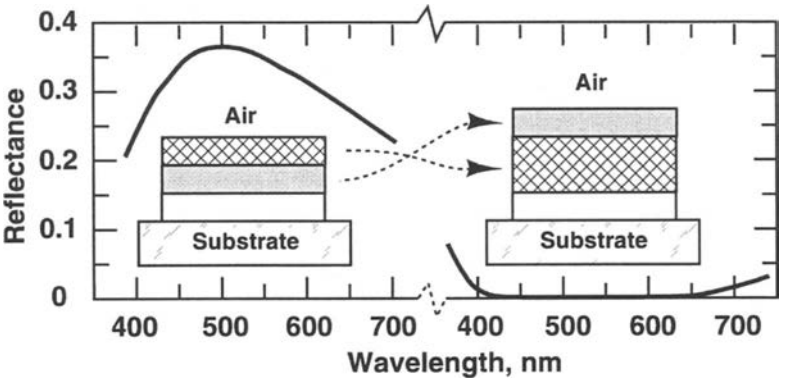


Fig. 1-8 (*Left*) $R(\lambda)$ of a reflector consisting of quarterwave optical thickness layers of magnesium fluoride (*dotted*), zirconia (*cross-hatched*) and alumina (*nonshaded*). (*Right*) Antireflection coating produced by interchanging layers and thickening the center layer to halfwave.

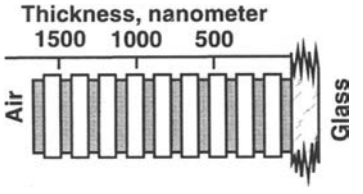


Fig. 1-9 Cross section, showing the metric thicknesses of a multilayer reflector (at 633 nm) consisting of 19 alternating layers of titania (*darker shading*) and silica (*no shading*).

Figure 1-9 depicts a multilayer consisting of 19 alternate layers of titania and silica deposited on a glass substrate. The high index layers of titania are approximately 70 nm thick and the silica layers are 110 nm thick. These thicknesses cause the stack to reflect strongly at 633 nm — for the He-Ne gas laser. An increase in the thickness of each layer would cause the stack to "resonate" at longer wavelengths, as, for example, at 1060 nm. Each titania layer has a thickness of about 300 atoms and the total thickness of the stack is close to 1.6 μm. For comparison, approximately 31 of these stacks would be piled on top of each other to equal the thickness of the paper upon which this book is printed.

1.1.2 Nonoptical attributes of surface coatings

Coatings are often deposited to alter physical properties of a surface. A plasma-deposited organic polymer can protect a surface from the environment — as, for example, when a hygroscopic material such as potassium chloride is exposed to high humidity [Wydeeven 76; Szwarc 76].

There are many uses of electrically conducting coatings, as, for example, on windows of longitudinal Pockels modulators or liquid crystal displays (see §1.9.7). The

mechanical and electrical properties of a coating are often as important as its optical properties. For example, electrically conducting coatings must be relatively transparent. The antireflection coating mentioned in the last paragraph should protect the potassium chloride from water vapor. Coatings can also increase the resistance of a surface to mechanical abrasion, as illustrated by the relatively thick layer of fused silica that is often deposited on a plastic lens.

1.2 Nomenclature and general properties

1.2.1 Nomenclature

The term *multilayer* is used generically to describe coatings and is unambiguous, except in electrical circuit board technology. *Multilayer* is especially appropriate for IR coatings that often contain as many as 100 layers. The term *dielectric* is widely used, as, for example, *all-dielectric, dielectric multilayer* or *dielectric reflector.* It means that such coatings are fabricated from materials that are essentially nonabsorbing at optical frequencies, i.e., nonmetals.

Table 1-1 Properties of materials used for substrates

Material	Refractive index			Transparency region*		Cost§	Comments
	UV	Vis	IR	Lower	Upper		
BK-7 glass	1.53	1.52	1.51	0.33 μm	2.5 μm	1	Can be hot pressed into cylindrical slabs or lenses.
Fused silica	1.50	1.47	1.45	0.18 μm	3.5 μm	10	Price varies, depending upon bubble content and purity.¶
Sapphire	1.83	1.77	1.68	0.17 μm	6 μm	5	Mechanically hard — difficult to polish.
Germanium	—	—	4.00	1.8 μm	16 μm	25	Least expensive material for FLIR optics.
Silicon	—	—	3.45	1.2 μm	8 μm	3	Used for CO_2 laser total reflectors — high thermal conductivity.
CaF_2	1.50	1.43	1.37	0.15 μm	9 μm	17	Relatively soft. Fractures easily.
Magnesium fluoride	1.42	1.39	1.37	0.12 μm	8 μm	60	Mechanically hard — difficult to polish.
Zinc sulfide	—	—	2.20	0.4 μm	12 μm	25	Less expensive than zinc selenide but more absorbing in the FLIR band.
ZnSe	—	2.65	2.40	0.6 μm	14 μm	25	Used for FLIR optics.
GaAs	—	—	3.27	1.0 μm	>30 μm	30	Used for CO_2 laser lenses.
Indium antinomide	—	—	4.2	6.0 μm	—	170	Transparent to 30 μm at temperature of 77 K.

Comments

*Typically of a slab a few millimeters in thickness.

§Typical cost (in U.S. dollars) of the raw material approximately 3.5 mm thick by 25 mm square. Cost depends upon optical quality. For example, fused silica with low bubble content and low stria is more expensive.

¶Composition of fused silica may be altered to optimize its transmittance in either the UV or the IR.

The carrier upon which the coating is deposited is the *substrate.* Table 1-1 lists substrate materials. The materials data in Ballard [72], Wolfe [78] and Palik [85, 91] are more comprehensive. Many criteria are used in the selection of a substrate, as, for example, its resistance to laser damage, cost, thermal conductivity, mechanical properties, the ease with which it can be optically polished or its transparency in a particular range of wavelengths. The latter property may be relatively unimportant if the coating contains an opaque metal layer or many dielectric films that reflect substantially all incident flux. Fused silica is widely used in laser systems because it is robust and can be polished to a high degree of smoothness.

1.2.2 Reflectance and transmittance

This section discusses the radiant reflectance R and radiant transmittance T and how they are measured.

1.2.2.1 Reflectance and transmittance of a coated part

Figure 1-10 shows an optical coating immediately after it has been removed from the chamber in which it was deposited. Although the coating may be cemented eventually to another sheet of optical glass, at this moment there is air (index n_0) on one side and the substrate (index n_s) on the other side. The face of the coating adjacent to air is its *obverse* side and next to the substrate is its *reverse* side. These definitions are independent of the direction that light hits the coating.

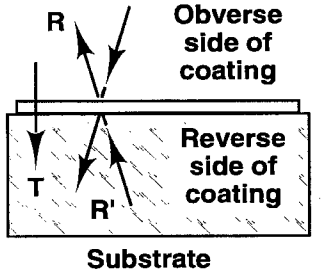


Fig. 1-10 Reflectance R on *obverse* side and R' on *reverse* side of multilayer (*shaded*) deposited on one side of substrate. T is transmittance of the multilayer.

Figure 1-11 shows a beam of radiant flux impinging upon a coating. The *incident* side of the coating is the one the flux hits first. The flux penetrates the coating and exits from its *emergent side.* It then propagates into the *emergent medium.* Whether the *obverse* or *reverse* side of the coating faces the emergent medium depends upon its orientation with respect to the incident beam.

Reflectance R and transmittance T are measured on the obverse side of the coating. R' and T' are the corresponding quantities measured on its reverse side. For many transmittance curves of coatings shown in this book, T is the transmittance from the incident medium into the bulk substrate. In other words, these $T(\lambda)$ curves correspond to those that would be measured if the substrate has no absorption and a "perfect" antireflection coating is deposited on its emergent interface.

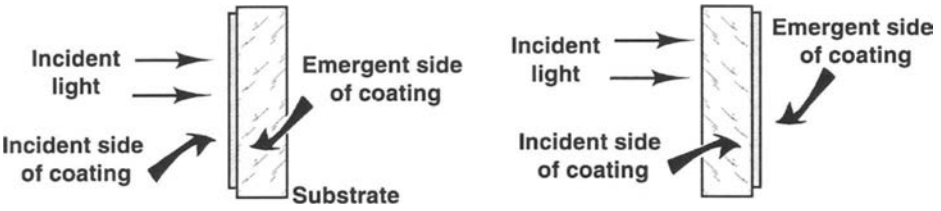


Fig. 1-11 *Incident* or *emergent* side of a multilayer (*shaded*) can be either the substrate or air, depending upon its orientation to the incident light.

It can be argued from thermodynamic principles [Fragstein 67] that

$$T = T' . \tag{1-1}$$

In other words, transmittance is invariant with respect to direction of propagation. As mentioned in Chap. 10, if monochromatic flux impinges on a filter array and the total flux over a broad band of wavelengths is collected by a detector, this reciprocity may not be true due to fluorescence in the bulk filter material. A consequence of Eq. (1-1) is that

$$R = R' \tag{1-2}$$

for any nonabsorbing coating. If such a coating produces a reflectance of 1.2% from its obverse (i.e., air) side, then its reflectance is also 1.2% at the reverse (i.e., substrate) side.

1.2.2.2 Overall transmittance and reflectance

The R and T cited above refer to the reflectance and transmittance of the multilayer alone, independent of the emergent interface of the substrate. There are additional transmission losses at that interface, as depicted in Fig. 1-12. When the flux flows through the bulk substrate and penetrates the *emergent interface*, *overall transmittance* T_{oa} is measured. Thus $T = T_{oa}$ only when there is a "perfect" antireflection coating on the emergent interface of a nonabsorbing substrate. Overall reflectance R_{oa} occurs when there are reflections from both the multilayer and the emergent side of the substrate. Equations for T_{oa} and R_{oa} of a slab appear in §10.3.5.1.

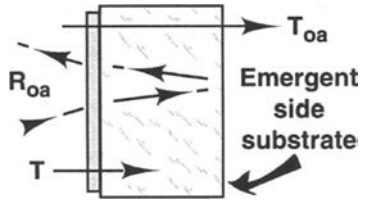


Fig. 1-12 *Overall transmittance* T_{oa} and *overall reflectance* R_{oa} are determined by both the multilayer (*shaded*) and emergent side of the substrate. T is the transmittance through the multilayer itself, from air into the substrate.

1.2.2.3 Absorption and loss

The radiant absorption A of a coating is that fraction of the flux (incident from the obverse side) converted into joule heat. A' is the absorption when the incident flux impinges upon its reverse, i.e., substrate, side. Conservation of energy requires that

$$A + R + T = 1 \tag{1-3a}$$

and

$$A' + R' + T = 1 \tag{1-3b}$$

although the reflectance and absorption need not be identical from the two sides. This is illustrated in Fig. 1-13, which shows a coating with nearly 70% reflectance from one side and a R' of below 10% from its opposite side.

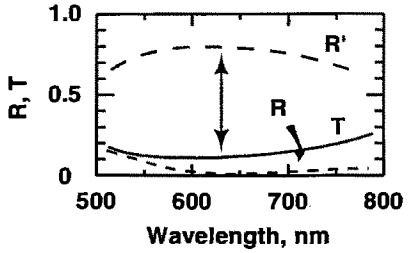


Fig. 1-13 $T(\lambda)$ (*solid*), $R(\lambda)$ from air side (*short dash*) and $R'(\lambda)$ from glass side (*long dash*) of the design shown below.

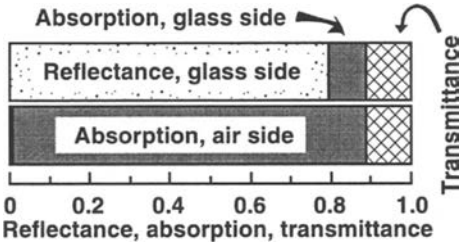


Fig. 1-14 At the arrowed wavelength in Fig. 1-13, absorption A and A' from obverse and reverse sides of the coating, respectively. R, R' and T are defined in Fig. 1-13.

For Figs. 1-13 and 1-14, the absorbing coating design is:

air **L** 0.58**H M** 0.85**L H L H** *glass*

where the indices of **H**, **L** and *glass* are 2.30, 1.52 and 1.45, respectively. The optical constant of **M** is $2.76 - j\,1.80$ and its metric thickness is 20 nm. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 633 nm.

Equation (1-1), Eq. (1-3a) and Eq. (1-3b) require that

$$R + A = A' + R' \quad . \tag{1-4a}$$

The sum of reflectance and absorption on each side of a coating must be the same. This is illustrated in Fig. 1-14, which graphs R, R', A, A' and T. The transmittance is the same in either direction.

Multilayers with such asymmetrical reflectance have been deposited upon sunglasses [Pohlack 59] and architectural glass [Apfel 72p]. When used in a Fizeau interferometer, bright fringes on a dark background appear in reflected illumination [Koch 62a] — the opposite of the fringes produced by a dielectric reflector. Another consequence is that a bandpass filter containing metal layers is usually more absorbing on one side than on the other. Such a bandpass may overheat if its more absorbing side is exposed to a high intensity light source.

The loss $\mathcal{L}$ is the fraction of the flux that is not reflected

$$\mathcal{L} \equiv 1 - R = A + T \tag{1-4b}$$

regardless of whether the loss is due to transmittance or absorption. Loss is important, for example, in the design of a coating for a laser cavity.

1.2.2.4 Opaque substrate

Equations in §2.4.4.3 compute T even when the substrate is opaque and absorbing. This T is the energy that flows past the last coating interface and into the substrate. A is the absorption in the coatings themselves.

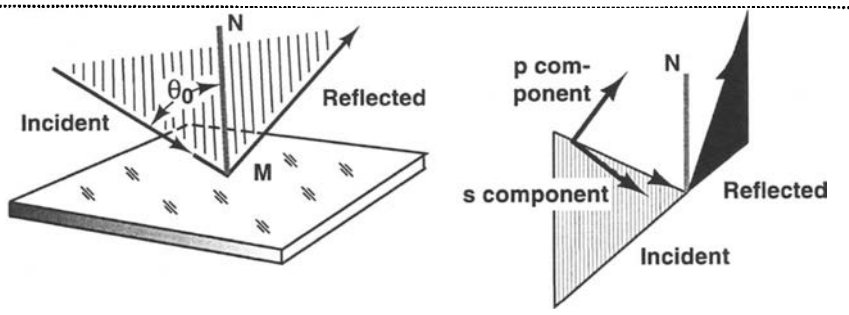


Fig. 1-15 (*Left panel*) Incident ray impinges upon multilayer M at an angle of incidence θ_0, as measured to normal N. Reflected ray R propagates at the same angle. (*Right panel*) Components of electric vector of incident light perpendicular (s) to plane of incidence and parallel (p).

1.2.2.5 *R* and *T* at nonnormal incidence

R and T are specified several ways at nonnormal incidence. The left panel of Fig. 1-15 depicts incident and reflected rays and the angle of incidence θ_0 measured to a normal. The plane of incidence contains the normal and the ray. The right panel shows that the flux can be linearly polarized with its electric field parallel to the plane of incidence (p polarization) or perpendicular (s polarization). The corresponding symbols are R_p, R_s, T_p and T_s for reflectance and transmittance of polarized flux. Unless otherwise stated, *reflectance* at nonnormal incidence refers to an arithmetic mean

$$R_a \equiv \tfrac{1}{2}(R_p + R_s) \quad , \tag{1-5a}$$

which is more accurately termed reflectance for *randomly polarized flux.* There is a similar equation for the transmittance T_a for *randomly polarized flux*

$$T_a \equiv \tfrac{1}{2}(T_p + T_s) \quad . \tag{1-5b}$$

The foregoing methods of computing R_a and T_a are meaningful as long as the source of radiation and the detector (such as a human eye) are insensitive to polarization.

A multilayer might function as a linear polarizer and transmit both the T_s and T_p. A figure of merit for such a device is the degree of polarization P_t

$$P_t \equiv \frac{T_p - T_s}{T_p + T_s} \quad , \tag{1-5c}$$

which is always bounded $-1 \le P_t \le +1$. A perfect polarizer has a P_t of either -1 or $+1$.

1.2.2.5.1 Linear polarization states

A mnemonic helps to identify the linear polarization states. Figure 1-16 shows the reflectance vs angle of incidence at an interface between air and BK-7 glass. Reflectance of p polarized flux is 4% at normal incidence and plunges to zero at 57° — the Brewster angle. Thus p is the *plunging* vector. The s polarized flux is reflected at all angles of incidence — s is the *skipping* vector.

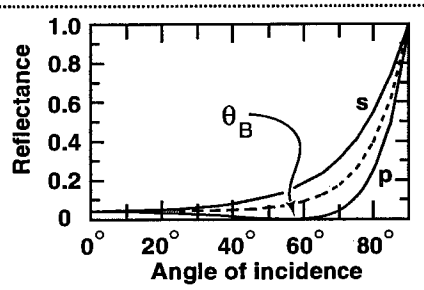


Fig. 1-16 At interface between air and BK-7 glass, reflectance vs angle of incidence θ_0 for *p* and *s* linearly polarized light. R_p is zero at Brewster angle θ_B.

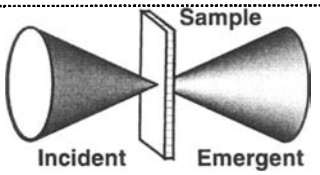


Fig. 1-17 In defining biconical transmittance, a cone of radiant flux Φ_1 is incident upon slab sample. A cone of flux Φ_2 emerges.

1.2.2.6 Absorbance and decibels

The *absorbance D* is convenient to use

$$D \equiv \log_{10}(1/T) \qquad (1\text{-}6a)$$

when transmittance ranges over many decades. The larger the absorbance (also called *optical density*), the lower the transmittance. An example is that an absorbance of 3 corresponds to a transmittance of $10^{-3} = 0.1\%$. Absorbance and absorption are distinct and should not be confused. *Absorptance* is sometimes used rather than *absorption*. There is a danger of confusion, and, for that reason, the word *absorptance* is avoided in this book.

Electrical engineers usually characterize the radiant transmittance of an optical device in decibels (dB)

$$\text{dB} \equiv 10 \log_{10}(T) \; . \qquad (1\text{-}6b)$$

Because T is fractional, the dB value is negative when the multilayer attenuates.

1.2.2.7 Biconical transmittance

Figure 1-17 depicts a conical bundle of rays converging upon a sample. The biconical transmittance T_b is

$$T_b = \frac{\text{total flux in output beam}}{\text{total flux in input beam}} \; . \qquad (1\text{-}6c)$$

§10.3.7.4 develops an equation for T_b of a coating at the focus of an aplanatic lens.

1.2.3 Wavelength and frequency

The tradition in optics is to compute or measure spectral reflectance or transmittance vs wavelength λ. The converse is true in microwave technology, where frequency ν is used almost exclusively. There are several advantages of using ν rather than λ:

(a) Quite often a coating has a transmittance curve $T(\nu)$ that manifests symmetry about a tuned frequency ν_0. §4.2.4 discusses the importance of this symmetry in antireflection coating design. This symmetry is masked when $T(\lambda)$ is plotted.

(b) A plot of $T(\lambda)$ over an extended range compresses the short wavelength spectral region and elongates features at longer wavelengths. This is illustrated in Fig. 5-28.

Notwithstanding the advantages of using frequency in optical coating design, references to yellow light as ν = 500 THz are rare, indeed. An alternative is to use inverse wavelength, i.e., wave number σ

$$\sigma \equiv 1/\lambda \; . \qquad (1\text{-}7)$$

The wave number is the number (or fraction) of wavelengths that fit into a given distance. Suppose that distance is 1.00 μm. Figure 1-18 shows at λ = 2 μm a halfwave "fits" into 1 μm and hence σ = 0.5 μm^{-1}. At a shorter wavelength of 0.4 μm, 2.5 waves "fit" into 1 μm. A non-SI unit of reciprocal centimeters is often used for σ. A conversion factor is:

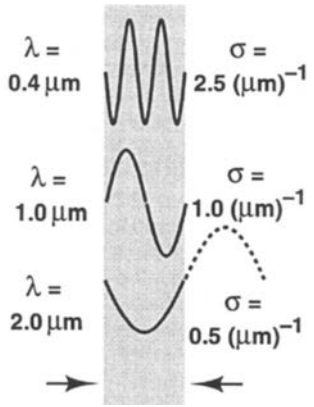


Fig. 1-18 Illustration of wave number σ as the number (or fraction thereof) of waves that "fit" into a distance (shaded) of 1 μm. λ is the corresponding wavelength.

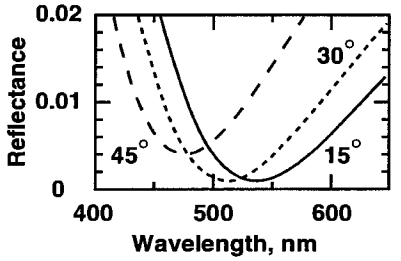


Fig. 1-19 $R(\lambda)$ of a "vee" antireflection coating (deposited on BK-7 glass) at incidence angles of 15° (*solid*), 30° (*short dash*) and 45° (*long dash*). After Matra [89].

A wavelength of 1 μm is equal to 1000 nm — both are equivalent to a wave number of 10,000 cm^{-1} or 1.0 μm^{-1}.

1.2.4 Angle shift

This term is used loosely to describe how a *feature* of a coating shifts in wavelength as angle of incidence θ_0 changes. *Feature* refers to such attributes as: (a) either a minimum or maximum in transmission or reflection; or (b) the wavelength λ_0 at which R or T attains a specific value — as, for example, $T(\lambda_0) = 40\%$.

A general property of multilayers is that features shift to shorter wavelengths (i.e., higher frequencies) with increasing angle of incidence. This author is unaware of any exceptions. §2.5.2.1 explains this phenomenon. Angle shift is illustrated in Fig. 1-19, which shows $R(\lambda)$ of a relatively narrowband antireflection coating. R_{min} shifts from 530 nm at 15° incidence angle to 480 nm (50 nm towards the blue) at $\theta_0 = 45°$.

A consequence is that if the angle of incidence is not rigidly specified, it is possible to *angle tune* a multilayer by tilting it. A bandpass is angle tuned (see §1.6.4.1) by fabricating it so that the wavelength of peak transmittance of the passband is slightly longer than required. Tilting the filter shifts its passband wavelength into specification. A multilayer polarizer may be angle tuned to optimize its transmittance at a particular wavelength, as shown in Fig. 1-173.

1.3 Antireflection coatings

A portion of the incident flux is reflected whenever it impinges upon an interface between media of different refractive indices, as, for example, an air-glass surface of a lens element. This is termed the *Fresnel loss* (see §2.2.4.1) and is graphed in Fig. 1-20 for various optical materials. *Fresnel losses* range from 4% for BK-7 glass to as large as 36% for germanium.

An appearance of a flare image in a lens is one of the undesirable effects of these reflections. Prior to the invention of the antireflection coating, many otherwise acceptable designs for photographic objectives were unusable because of objectionable flare. The use of antireflection coatings has virtually eliminated this problem. There is even greater danger when a lens is irradiated with a high power laser. The reflected beams can refocus and produce such high energy density that it destroys the glass.

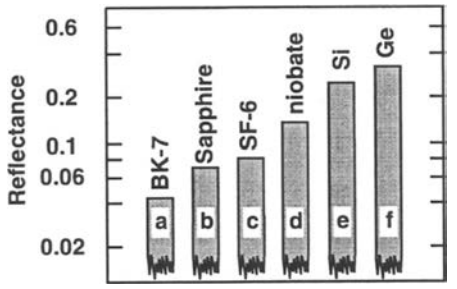


Fig. 1-20 R at interface between air and materials of refractive index n:
(*a*) BK-7 glass ($n = 1.52$);
(*b*) sapphire ($n = 1.73$);
(*c*) SF-6 glass ($n = 1.85$);
(*d*) lithium niobate ($n = 2.25$);
(*e*) silicon ($n = 3.5$);
(*f*) germanium ($n = 4.0$).

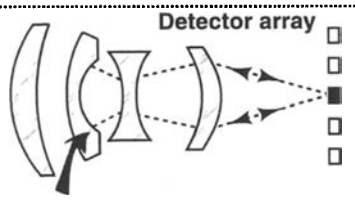


Fig. 1-21 The narcissus effect, in which a portion of the radiation from a particular IR detector (*solid black*) in a focal plane array is re-imaged upon itself because of an internal reflection (*arrowed*) from a lens surface.

Antireflection coatings are essential for IR optical components such as lenses or domes that contain germanium, silicon or other materials of high refractive index. The loss of flux at uncoated surfaces would be prohibitive otherwise.

Antireflection coatings also greatly reduce narcissus in IR imaging systems with focal plane arrays [Wolfe 78]. Figure 1-21 depicts a focal plane array of detectors. The detectors themselves radiate energy. As shown by the dashed rays, however, the radiation from one of the detectors reflects from an internal surface of a lens and retroreflects back upon itself. Thus the detector "sees its own image." This effect alters its temperature and responsivity.

1.3.1 Choice of an antireflection coating

An antireflection coating can consist of either a single layer or many layers. The spectral bandwidth over which it functions can be narrow or relatively broad. Table 1-2 lists properties of these coatings. Trade-offs are discussed in §1.3.1.4.

1.3.1.1 Single layer of magnesium fluoride

This coating is widely used to reduce the reflectance of glass or plastic. Each year the worldwide producers of consumer cameras and TV projection systems coat many million surfaces of plastic lenses with this film. Quite often a fused silica solar cell cover has a magnesium fluoride layer on its outer surface. The coating performance depends upon the refractive index of the glass.

1.3.1.1.1 Substrate refractive index

The long dash curve in Fig. 1-22 shows that the minimum reflectance on BK-7 glass is only 1.2%. Other curves illustrate that R_{min} drops to nearly zero when deposited on SF-11, which has an index as large as 1.83.

1.3.1.1.2 Wavelength of minimum R

Dashed curves in Fig. 1-22 show a minimum R at 550 nm in the green, where the human eye is most sensitive. This coating is usually applied to visual optics, such as binoculars. This layer has higher reflectance in the blue and red — hence, it appears purplish in reflection.

Table 1-2 Antireflection coatings for UV, visible and near IR

Type of coating	Figs.	Spectral region of use	Lay-er(s)	Advantages and typical performance	Disadvantages
Single MgF_2	1-22 1-23	120 nm to 1 μm	1	Low cost, performs well on high index glasses.	On BK-7, minimum R is 1.2%.
Broad-band	1-27 1-28	250 nm to 2 μm	3 to 10	R below 0.3% in spectral region 420 nm to 680 nm	Difficult to deposit on highly curved surfaces.
Narrow-band	1-25	190 nm to 2 μm	2 to 6	High laser damage threshold. $R < 0.1\%$.	Relatively narrow spectral bandwidth.

Attributes of antireflection coatings deposited on glass or fused silica.

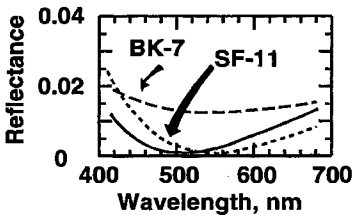


Fig. 1-22 Computed $R(\lambda)$ of a layer (optical thickness $\lambda_0/4$) of refractive index 1.38 deposited on BK-7 (*long dash*) and SF-11 glass (*short dash* and *solid*). λ_0 is 550 nm for the *dashed* curves and 510 nm for the *solid*.

The wavelength at which R_{min} occurs is altered by adjusting the layer thickness. A thinner film moves the minimum R towards the blue (*solid* curve in Fig. 1-22) — the reflected light is amber. Figure 1-23 shows an amber coating (curve A) on a BK-7 substrate and also a coating centered at longer wavelengths (curve B), which appears bluish in reflection.

Amber coatings are often applied to photographic objectives [Scharf 52; Obodovskia 67]. Higher index glasses in lens elements absorb slightly in the blue. The increased blue transmittance produced by the coating compensates for this loss and produces a better color balance, as shown in Fig. 1-24.

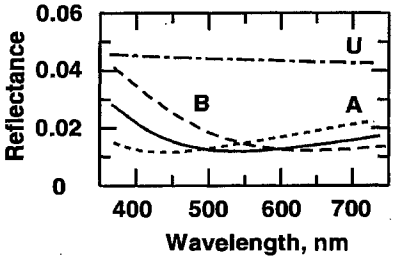


Fig. 1-23 Caption to Fig. 1-22 obtains, with the exception that the glass is BK-7. λ_0 is 450 nm (*short dash*, curve A), 550 nm (*solid*) and 650 nm (*long dash*, curve B). Curve U (*long-short*) is R of uncoated substrate.

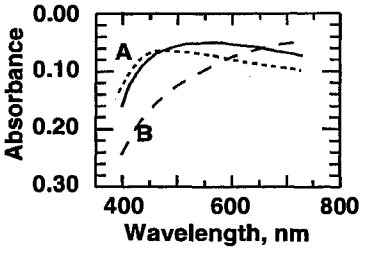


Fig. 1-24 $T(\lambda)$ (in absorbance units) of SLR camera lens, each surface of which is coated with a magnesium fluoride layer. Labeling of curves is the same as in Fig. 1-23. After Scharf [52].

1.3.1.1.3 Additional single-layer antireflection coatings

A single-layer antireflection coating of tantalum pentoxide reduces the R of the silicon surface of a photodiode or solar cell [Yeh 76; Arndt 75], thus producing higher efficiency. An antireflection coating consisting of a single-layer fluorocarbon film is

deposited on the polyester sheet used in instant camera film [Land 77]. Hard carbon (see §1.3.2.1) is also a single-layer antireflection coating for the infrared.

1.3.1.2 Narrowband antireflection coatings

A narrowband antireflection coating is called a *vee-coat* in the trade because of the shape of its $R(\lambda)$ curve, as shown in Fig. 1-19. In most cases it contains two or more layers and can be designed for any wavelength at which optical coating materials exist with low absorption and/or high laser damage threshold. Figure 1-25 shows $R(\lambda)$ of a vee-coat on fused silica for excimer laser line at 193 nm and Fig. 1-26 depicts $R(\lambda)$ of vee-coats on the same substrate for excimer laser lines at 248 nm and 308 nm. Figure 4-18 depicts $R(\lambda)$ of a vee-coat for a BK-7 substrate at 1054 nm. The coating materials (silica and titania) withstand laser pulses of high energy per unit area.

1.3.1.3 Broadband antireflection coatings

A multiple-layer antireflection coating contains anywhere from three to 20 layers. Photographic objectives coated with them are called *multicoated*. Figure 1-27 depicts $R(\lambda)$ of antireflection coatings spanning the visible spectral region. Both antireflection coatings in Fig. 1-28 function over a bandwidth approaching 2:1.

1.3.1.4 Choice and uses of antireflection coatings

A question remains: Should single layers or multicoats be deposited on photographic objectives? An answer involves the type of glass of the lens element, its curvature, cost and required level of R. In the discussion below *single layer* refers to magnesium fluoride.

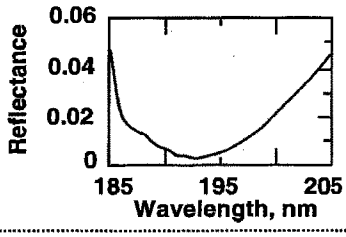


Fig. 1-25 Measured reflectance of an antireflection coating for 193 nm deposited upon a fused silica substrate. After TecOptics [99].

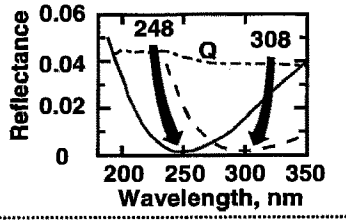


Fig. 1-26 Deposited upon fused silica substrates, $R(\lambda)$ of antireflection coatings for 248 nm and 308 nm excimer laser lines. Q (*long-short dash*) is the R of one surface of uncoated substrate. After Fischer [88].

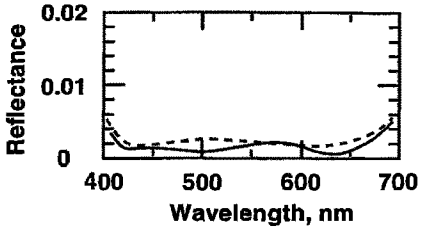


Fig. 1-27 Measured $R(\lambda)$ of antireflection coatings on BK-7. After Twardy [90] (*dashed*) and The Optical Corporation [90] (*solid*).

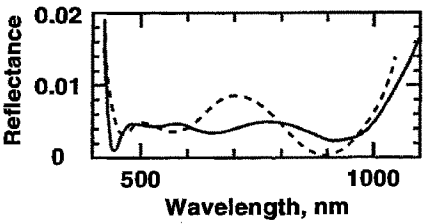


Fig. 1-28 Measured $R(\lambda)$ of antireflection coatings on BK-7. After ZCR [88] (*dashed*) and OCI [90] (*solid*).

R below 1% is usually more than adequate to reduce flare in most photographic objectives. Therefore, any R below 1% only enhances transmittance. Figure 1-22 demonstrates that single layers on higher index glass reduce R quite effectively. Furthermore, it is difficult to apply a mechanically hard antireflection coating to a plastic lens because plastic must be coated at a relatively low temperature. Deposition of a multilayer antireflection coating is more expensive than a single layer because of the complexity of equipment used to control layer thicknesses. A decision to multicoat lenses is ultimately based upon economic trade-offs. Articles by Keppler [71], Sherman [71] and Goldberg [74] review public acceptance of such lenses in the 1970s.

Exploded views of SLR camera lenses — both normal (50 mm) and wide angle (35 mm) — appear in Fig. 1-29. Single layers are applied to surfaces marked *s* and multilayers to those tagged *m*. The $R(\lambda)$ of coatings on both sides of lens element *z* in Fig. 1-29 appears in Fig. 1-30. This single layer (*upper* curve) is an amber coating (see §1.3.1.1.2). The appearance of two minima in the $R(\lambda)$ (*solid* curve) of the opposite side of the lens suggests a multicoat containing three or four layers. $R(\lambda)$ in Fig. 1-31 (pertaining to element *L* of Fig. 1-29) is characteristic of a multicoat.

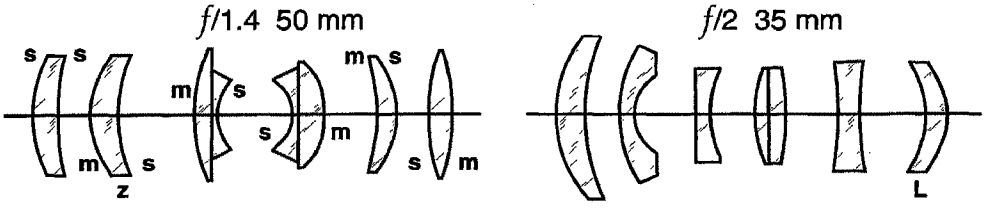


Fig. 1-29 Exploded views of the elements of SLR camera lenses. Letter *m* denotes a multilayer antireflection coating and *s* a single layer. $R(\lambda)$ of elements *z* and *L* appear i Figs. 1-30 and 1-31, respectively. After Matteucci [80].

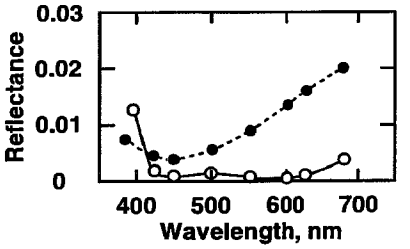


Fig. 1-30 Measured $R(\lambda)$ of each side of lens element z in Fig. 1-29. After Matteucci [80].

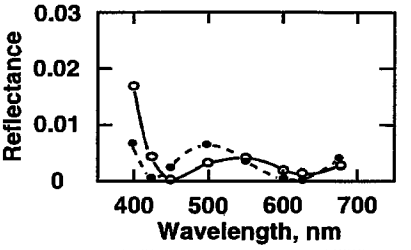


Fig. 1-31 Measured $R(\lambda)$ of each side of lens element L in Fig. 1-29. After Matteucci [80].

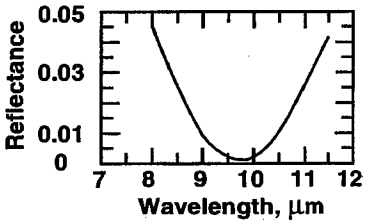


Fig. 1-32 Measured $R(\lambda)$ of a single side of germanium coated with a single layer of hard carbon. After Exotic [93].

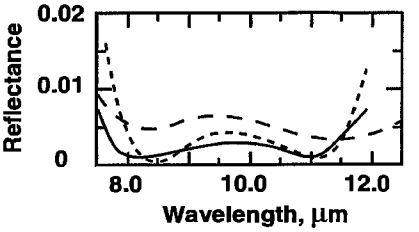


Fig. 1-33 Measured $R(\lambda)$ of antireflection coatings deposited on single surfaces of germanium (*solid*, [Corion 90]), zinc selenide (*short dash*, [Corion 90]) and zinc sulfide (*long dash*, [OCLI 81]).

It is preferable for several reasons to apply a single-layer antireflection coating to a highly curved lens surface. Layer thicknesses are usually nonuniform across the lateral extent of such a surface (see §1.9.1.3). The $R(\lambda)$ of a single layer is less sensitive to such thickness variations, as compared to a multicoat. Given the fact that the rays usually hit a highly curved lens surface at relatively high incidence angles, photopic R of a single-layer antireflection coating changes relatively less with angle of incidence, as compared to a multiple-layer antireflection coating.

Eyeglasses with multicoats are quite common in Europe and Japan. antireflection coatings are also deposited on cover plates of instruments, as, for example, altimeters in aircraft. They are also applied to glass that protects paintings in an art gallery and on the outer surface of absorbing glass that enhances the contrast of a visual display terminal.

1.3.2 Antireflection coatings for the IR

Similar to coatings for the visible spectrum (see §1.3.1), there are antireflection coatings for the IR that function over both relatively narrow and broad bandwidths.

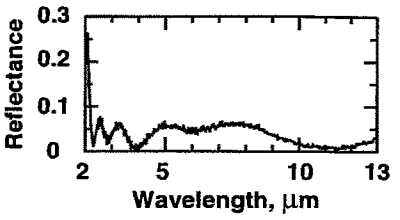


Fig. 1-34 Measured $R(\lambda)$ (single surface) of an antireflection coating on silicon. After Deposition Sciences [99].

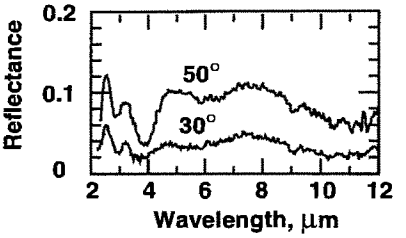


Fig. 1-35 Caption to Fig. 1-34 obtains, with the exception that incident angles are 30° and 50°.

1.3.2.1 IR antireflection coatings with relatively narrow bandwidth

An antireflection coating for the DF laser line at $\lambda = 3.8$ μm is an example. A special three-layer coating was developed that not only reduced substantially the $R(\lambda)$ of a potassium chloride substrate, but, in addition, had an extremely low absorption of 0.008% [Rudisill 74; Braunstein 78]. Baer [78a, 78b] fabricated a three-layer antireflection coating with similar properties using films of arsenic selenide and sodium fluoride.

Figure 1-32 depicts the $R(\lambda)$ of germanium covered with a single layer of *hard carbon*, which is a mixture of hydrocarbons and carbon. This manifests 3% to 4% absorption in the FLIR band, which is approximated by the spectral region 7.7 μm to 12.3 μm. Its chief virtues are its robustness and resistance to mechanical abrasion.

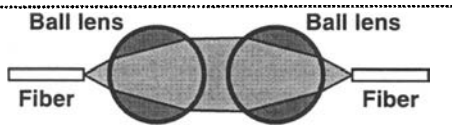


Fig. 1-36 Transfer of the radiant flux from one fiber core to another fiber core by a train of ball lenses. After Deposition Sciences [99].

1.3.2.2 IR antireflection coatings with broad bandwidth

Antireflection coatings are used in imaging systems functioning in the FLIR band. There are essentially two classes of antireflection coatings: Those that contain layers of rare earth fluorides (with its relatively low refractive index) and those that do not. The use of a rare earth fluoride, such as yttrium fluoride, in an antireflection coating results in lower reflectance, but at the expense of considerable mechanical stress. §4.1.3.3 shows that the average reflectance of germanium coated with an antireflection coating containing no rare earth fluoride materials usually hovers about 1.5%, whereas the $R(\lambda)$ of a coating with rare earth fluoride layers is at least a half, if not one-third, of that value.

Figure 1-33 shows the $R(\lambda)$ of FLIR antireflection coatings on several substrates. Figure 1-34 depicts the $R(\lambda)$ of an antireflection coated silicon slab. Its spectral bandwidth covers a 5:1 ratio, from 2.2 μm to 13 μm.

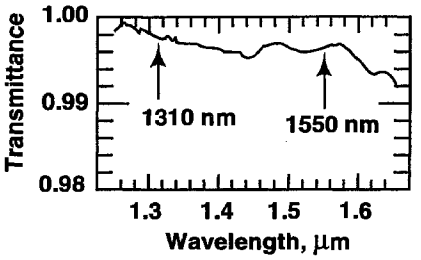


Fig. 1-37 Measured $T(\lambda)$ of an individual ball lens, as in Fig. 1-36. After Deposition Sciences [99].

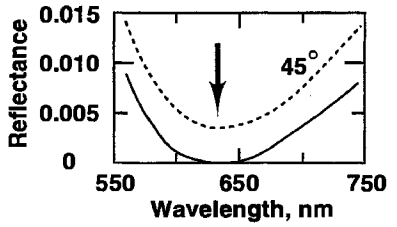


Fig. 1-38 $R(\lambda)$ of antireflection coatings at 633 nm at normal incidence (*solid*) and $R_a(\lambda)$ at 45° (*dashed*). After Rodenstock [90].

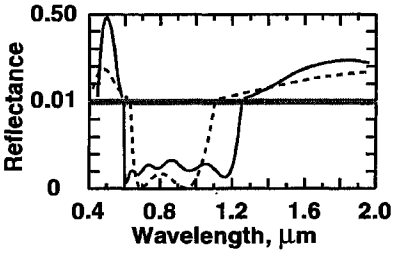


Fig. 1-39 $R(\lambda)$ of an antireflection coating (*dashed*)

air 0.533**F** 1.07**M** 0.533**S** *glass*

and antireflection coating with 2:1 bandwidth (*solid*)

air 0.540**F** 1.18**H** 0.29**L** 0.238**H** 0.256**L** 1.26**H** 0.191**L** 0.141**H** *glass*

where indices of **F**, **L**, *glass*, **S**, **H** and **M** are 1.38, 1.45, 1.517, 1.63, 1.95 and 2.10, respectively. Optical thicknesses are multiples of $\lambda_0/4$ at λ_0 of 1.5 μm. Ordinate changes scale at 0.01.

polarized light) at 45° incidence.

1.3.3 Antireflection coatings for fiber-optic communication devices

At the time of this writing (year 2000), communication by fiber optics centered at two spectral regions: 1310 nm and 1550 nm. Four-layer antireflection coatings are deposited upon the ends of single mode fibers that reduce the reflectance to below a few tenths of a percent in those spectral regions. Often it is necessary to transfer the signal from one fiber to another. Figure 1-36 shows that a train of ball lenses of LaSFN9 glass (refractive index 1.88) images one fiber onto another. The measured transmittance of the antireflection coated ball in Fig. 1-37 shows that the surface reflection is below 0.25%. This coating is deposited by a CVD process [Deposition Sciences 99].

1.3.4 Antireflection coatings at nonnormal incidence

Notwithstanding the danger of sweeping generalizations, the author's experience in designing and producing optical coatings has taught him that: (1) It is straightforward to design an antireflection coating for a relatively narrow bandwidth for a *given* polarization (i.e., either *p* or *s*), regardless of incidence angle; (2) it is quite difficult to design an antireflection coating — even over a small bandwidth — that functions well for both *p* and *s* polarized flux, especially when the incidence angle exceeds 55°; (3) at angles in excess of 55°, it is usually easier to design an antireflection coating for *p* polarized flux, as compared to *s* polarization.

For example, how does the coating shown in Fig. 1-34 perform at higher angles of incidence? Figure 1-35 shows that at 50° average reflectance approaches 9% — in contrast to 4% at normal incidence. Figure 1-38 shows a vee-coat antireflection coating (for 633 nm) manifesting R close to zero at normal incidence and 0.5% R (for randomly

1.3.5 Out of bandwidth reflectance

The performance of a single-layer antireflection coating can be summarized: At *any* angle of incidence and at *any* wavelength its $R(\lambda)$ (for randomly polarized light) *never* exceeds that of the uncoated substrate. The same is usually not true of multiple-layer antireflection coatings. Figure 1-39 shows $R(\lambda)$ of a three-layer antireflection coating. Its out-of-band reflectance is 20% at wavelengths in the IR and UV. The Fig. also depicts $R(\lambda)$ of an antireflection coating with a 2:1 bandwidth. Its reflectance is 50% at certain out-of-band wavelengths.

1.3.6 Specification of an antireflection coating

If pertinent, a specification of an antireflection coating could mention:

(a) Type of substrate and incident medium.
(b) Maximum reflectance in a given spectral band.
(c) Average reflectance in a given spectral band.
(d) Transmittance and/or absorption.
(e) Angle of incidence (is the flux linearly polarized or randomly polarized?).
(f) Environmental stability — 50 rub cheese cloth, eraser rub, 24 hour humidity and salt spray.
(g) Clear aperture (see §1.9.1).

A specification might read

> "The antireflection coating is deposited on customer supplied BK-7 glass. At any angle of incidence between 0° and 20° the average reflectance for randomly polarized light shall not exceed 0.3% in the spectral region 420 nm to 680 nm. The reflectance cannot exceed 0.5% in the aforementioned spectral region and range of angles."

1.3.6.1 Wavelength ratio

In defining bandwidth of low reflectance, the important factor is the *ratio* of wavelengths. An antireflection coating that functions from 400 nm to 800 nm covers a 2:1 span of wavelength. It is just as difficult to cover the spectral region 1000 nm to 2000 nm — this, too, is a 2:1 span of wavelength.

1.3.6.2 Use of a witness piece

A manufacturer deposits on lenses the antireflection coating specified in §1.3.6. How does the producer certify that those requirements have been met? In most cases, $R(\lambda)$ is measured on a *witness piece*, rather than on the lens surface itself. This witness is a slab of material of the same physical composition as a lens and is treated in every way like a lens. It is cleaned in the same manner and inserted in the hole in the tooling, as shown in Fig. 9-25. It receives the same coating as the lenses. After deposition, the witness is inserted into the sample compartment of a spectrophotometer and its spectral reflectance measured.

A witness should meet several requirements. First, it should have approximately the same thermal mass as the lenses. In this way, the surface temperatures of the lenses and witnesses are approximately the same when they are heated inside the vacuum system and coated with thin films. This is important because temperature changes during coating can alter the properties of thin films, which, in turn, alter the spectral reflectance. Another requirement is that the witness should be neither too large nor too small. If too

large, it will not fit into the spectrophotometer. If too small, it will fail to intercept the beam in that instrument.

There are instances, however, in which the spectral reflectance on the surface of a lens itself is measured, as, for example, the curves in Fig. 1-30. This requires a special instrument that is time consuming to operate.

1.3.6.3 Staining and its avoidance

It is quite costly to deposit antireflection coatings on certain optical glasses that are stainable. During polishing of such glasses in an aqueous slurry, small areas are damaged by chemicals in the slurry. These areas are often not visible until the optical coating process heats the glass to several hundred degrees.

A mottled pattern on the surface, termed *stain*, appears. Such problems are prevalent in glasses listed in catalogs of manufacturers [Schott 88; Hoya 86] that indicate a *resistance to staining* or *resistance to climatic variations* of 4 or 5. Stainable glasses must be cleaned laboriously and also stored in dry nitrogen while awaiting coating.

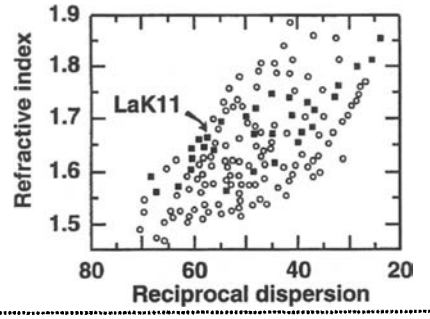


Fig. 1- 40 Refractive index at 589 nm vs reciprocal dispersion ν of optical glasses. Glasses marked as black squares manifest a tendency to stain — their ability to resist attack by chemicals is rated 4 or 5. After Matteucci [80] and Schott [88].

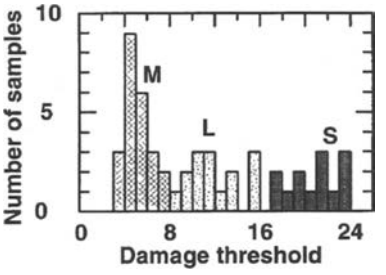


Fig. 1-41 Histogram of threshold for laser damage at 1.06 μm of BK-7 covered with multilayer antireflection coatings deposited by vacuum evaporation (*M*), etched surfaces (*L*) and sol-gel deposited single layers (*S*).

Figure 1-40 depicts the refractive index n_d and reciprocal dispersion of optical glasses [Matteucci 80]. Those identified by a solid square stain easily. Substitution of a nonstainable glass (with comparable index and dispersion, of course) substantially reduces the cost of the coating. As an illustration, Schott glass LaKN22 and LaK11 have a *resistance to climatic variations* of 2 to 3 and 4, respectively. Given that the dispersion and n_d are reasonably close, the former is preferable.

1.3.7 Alternatives to vacuum-deposited antireflection coatings

Most antireflection coatings in previous sections of §1.3 are deposited by vacuum deposition or by sputtering. Two alternatives to these coatings are: (1) etched surfaces; and (2) sol-gel coating.

1.3.7.1 Etched surfaces

Immersion of glass into an HF etch [Nicoll 55] has been used extensively as an antireflection coating. This etched glass often appears in picture frames sold in North American discount department stores. Minot [76] improved the process by etching glass that had been phase separated by heat treatment. Schröder [53] developed neutral solution etching of glass. In each process, chemical leaching produces a porous, optically inhomogeneous region at a glass surface. Physical optics explains the reduction of reflectance by modeling an etched surface as a *refractive index graded profile*. Its index gradually changes from that of air ($n = 1.00$) to that of bulk glass.

Advantages of this process are its low cost and higher laser damage threshold. For example, Figure 1-41 shows that an antireflection coating produced by neutral solution etching of BK-7 has a higher laser damage threshold than an antireflection coating produced by conventional vacuum deposition [Cook 82].

Disadvantages of the etched surface are its poor mechanical properties and its diffuse scattering of light. An etched surface is porous and diffusely scatters light. Its reflectance is increased by ingestion of oil, dirt or water into that surface. Engineers who quantify performance of a visual display terminal (such as a faceplate of a cathode ray tube of a computer terminal) call the scattered light *veiling glare*. This glare is not objectionable in a photothermal energy collector, however, where the relatively low cost of this coating makes it ideal [Jurisson 75].

1.3.7.2 Sol-gel coatings

Another alternative is a chemical process that deposits a thin layer with a graded refractive index profile. An article reviewing this technology [Lowdermilk 87] uses the term *sol-gel* generically to describe such a coating. Porous sol-gel silica films have been used extensively on surfaces of BK-7 glass used in high-power lasers. Figure 1-41 shows that these antireflection coatings have been successful in increasing laser damage threshold at 1.06 μm.

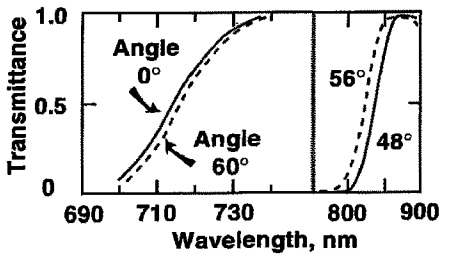


Fig. 1-42 Neglecting surface reflection losses, $T(\lambda)$ of Schott RG-715 colored glass (3 mm thickness) at incidence angles of 0° and 60°. $T_p(\lambda)$ (measured) of a multilayer at angles of 48° and 52°. Abscissa changes scale at 750 nm. Multilayer after Coherent [92].

1.4 Spectral filtering and narrowband rejection

1.4.1 Comparison of interference and absorption filters

There are numerous absorption filters that isolate broad spectral regions. In many cases, they are more effective than optical coatings. A Schott UG-11 glass filter, for example, transmits in excess of 80% at 330 nm in the UV, yet attenuates to at least 4 absorbance units throughout most of the visible spectrum. This coating would be more expensive to produce using thin films. Absorption filters have several advantages: (1) lower cost; (2) immense attenuation; and (3) as compared to multilayers, lower *angle shift*.

As an example of attenuation, an absorbing organic chemical known as *Cation-x* produces an attenuation of 6 absorbance units at $\lambda = 320$ nm, while yet transmitting 73% at $\lambda = 250$ nm [Kasha 48; McBride 62]. The angle shift (see §1.2.4) of the cutoff or cuton wavelength of absorption filters is considerably less than that of interference devices. As an example, Fig. 1-42 shows $T(\lambda)$ of Schott RG-715 glass and a multilayer polarizer containing layers of titania and silica. A path of a ray through the glass *increases* by 21% as the angle of incidence in air grows from zero to 60°. This, in turn, causes the cuton wavelength to increase by 1.3 nm. In other words, its $T(\lambda)$ curve shifts to *longer* wavelengths.

By comparison, the same Fig. depicts $T(\lambda)$ of a multilayer for *p* polarized flux. $T(\lambda)$ shifts to *shorter* wavelengths by an amount 16 nm when the angle of incidence increases by a mere 4° — from 48° to 52°. Wavelength shifts of a bulk absorber and multilayer are in opposite directions. §2.5.2.2 and §1.2.4 discuss *angle shift* of multilayers.

1.4.2 Selective absorbers

There are data on $T(\lambda)$ of solutions and gases in Landolt-Börnstein [Hellwege 62; Geffcken 57]. A text of Wood [34] and article of Greenler [58] contain useful information on absorbers. The transmittance of UV cements was studied by Pellicori [64, 66]. Dobrowolski [74, 77, 78a] examined filter glass of various manufacturers [Corning 53; Hoya 88; Schott 62] and concluded that many are equivalent and thus interchangeable.

Gelatin sheets manufactured by Eastman [70] are useful filters, although some organic dyes degrade under UV radiation. IR absorption filters with passbands in the spectral region from 20 μm to 100 μm are concocted by mixing powders of alkali halides in a matrix of polyethylene [Yamada 62; Mitsuishi 63]. Other filters for that spectral region are reviewed by Holah [89].

1.4.3 Reflection filter

Four uses of a reflection filter are: (1) attenuation of the long-wavelength spectral region; (2) "notch" — narrow-band attenuator; (3) reflector in a laser cavity (see §1.7.7); and (4) optical variable device (see §1.7.7).

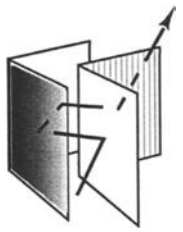


Fig. 1-43 Reflection filter consists of four plano mirrors coated with multilayer dielectric reflectors. After Schott [72].

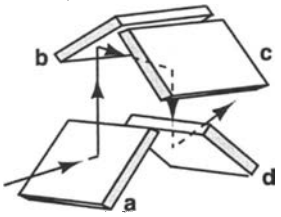


Fig. 1-44 Modification of reflection filter in Fig. 1-43. The inner pair of mirrors is rotated by 90° about an axis perpendicular to the beam.

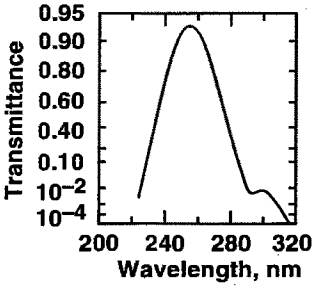


Fig. 1-45 $T(\lambda)$ of reflection filter, construction of which appears in Fig. 1-43. After Schott [72].

1.4.3.1 Long-wavelength attenuator

This filter transmits shorter wavelengths (such as UV) and attenuates longer wavelengths in the visible and near IR. Figure 1-43 shows four plano multilayer-coated mirrors inclined at 45° [Schott 72]. The mirror configuration is modified by rotating the inner pair of mirrors by 90° about an azimuthal angle perpendicular to an axial ray. The reflection filter in Fig. 1-44 is less polarization sensitive because each ray makes two *p* reflections and two *s* reflections.

Figure 1-45 shows that the overall transmittance of the filter in Fig. 1-43 exceeds 85% in the UV and is below 10^{-4} at wavelengths of 320 nm and longer. There is nothing sacred about an angle of incidence of 45°. An angle of incidence of 30° has been suggested [Sarna 74].

1.4.3.2 Notch filter

This coating rejects a narrow wavelength band in the vicinity of λ_0. It transmits at wavelengths shorter and longer than λ_0. Its design is similar to that in Fig. 1-43. Each mirror has a low reflectance at λ_0; R rises abruptly at wavelengths adjacent to λ_0. Figure 1-46 shows that its $T(\lambda_0)$ is below 6 absorbance units. Such filters are used in Raman spectroscopy to reject light from an argon ion laser at 515 nm.

1.4.4 Narrowband attenuator

A transmissive notch filter is expensive to fabricate because many thin layers must be deposited in order to achieve a relatively narrow bandwidth and a reasonable level of rejection. Figure 1-47 shows $T(\lambda)$ of a notch filter composed of 61 layers of alumina and silica. It has a transmittance of 0.3% at 507 nm and a luminous transmittance in excess of 80%.

Figure 1-48 graphs $T(\lambda)$ of a rugate filter — the refractive index vs distance has a quasi-sinusoidal profile [Barr 97]. The spectral bandwidth of its rejection region is 18 nm and its minimum transmittance is 4.5 absorbance units. Figure 1-49 shows $T(\lambda)$ of another rugate type of filter. This coating, which is composed of alternate layers of oxynitrides of silicon, has a rejection bandwidth of 6% and minimum transmittance of 4×10^{-5} [Edmonds 90]. Figure 1-50 depicts $T(\lambda)$ of a rejection filter produced by more conventional means. Figure 1-51 graphs the transmittance of a relatively narrow-band rejection filter for the middle infrared spectral region [Barr 97].

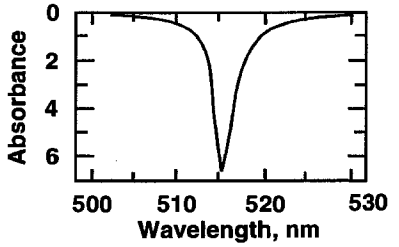


Fig. 1-46 $T(\lambda)$ (in absorbance units) of notch rejection filter, construction of which is similar to Fig. 1-43. After Omega [91].

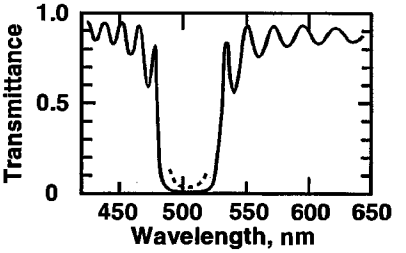


Fig. 1-47 Measured $T(\lambda)$ of rejection filter consisting of 61 alternate layers of silica and alumina. *Short dash* — ordinate is from 0 to 0.01. After Coherent [92].

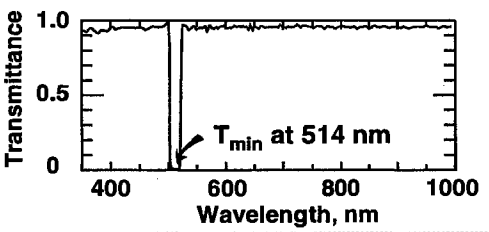


Fig. 1-48 Measured $T(\lambda)$ of a rejection filter designed to attenuate at the wavelength 514.5 nm. After Barr [97].

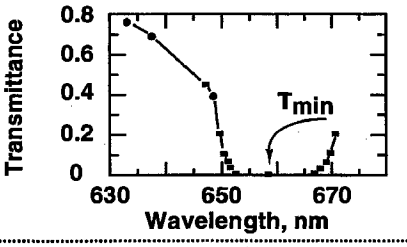


Fig. 1-49 Measured $T(\lambda)$ (squares and circles) of a rejection filter. T_{min} is 4×10^{-5}. After Edmonds [90]

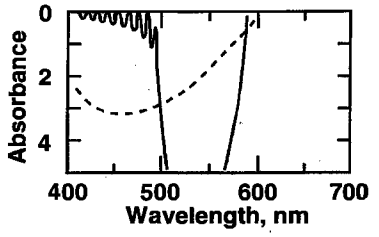


Fig. 1-50 $T(\lambda)$ (in absorbance units) of several rejection filters. After Omega [91].

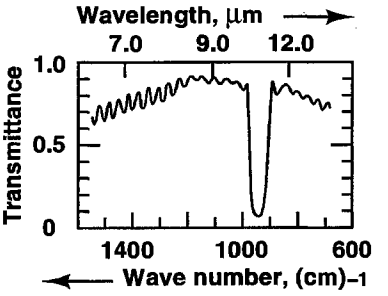


Fig. 1-51 Measured $T(\lambda)$ of a narrow-band rejection filter that attenuates the principal emission lines from a carbon dioxide laser. After Barr [97].

1.4.4.1 Cascading of attenuators

Suppose that two plano sheets of glass S are each coated with the same all-dielectric attenuator. Figure 10-27 shows that their faces are parallel. Assume that the transmittance of each coating is 1% and that the light is at normal incidence. The overall transmittance T_{oa} of the cascaded slabs is then approximately only 0.5%. Why this reduction of T by a factor of two? Why is the transmittance not $(10^{-2})^2 = 10^{-4}$?

The explanation lies in the fact that since the coatings are all-dielectric, they are essentially nonabsorbing and therefore have a reflectance in excess of 99%. One percent of the light penetrates the first slab and this light bounces many times between the two surfaces, due to the high reflectance. One-half of the light bouncing between the two surfaces propagates to the left and one-half to the right. Hence, the factor of two in the reduction of T_{oa} .

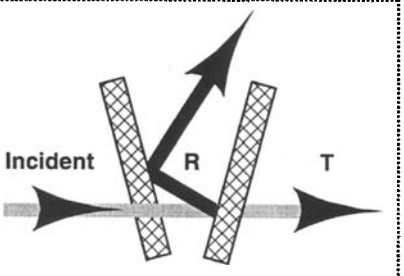


Fig. 1-52 Two attenuators (cross hatched) are tilted at 15° so that the reflected component (*solid black, R*) exits from the system.

Inspection of Fig. 10-27 hints at a method of increasing the attenuation. The slabs are tilted so that the rays "walk off" out of the system. Figure 1-52 shows that a mere 15° tilt of the slabs permits the reflected ray (*solid*) to exit from the system.

When should this technique of "wedging" the attenuators be used? It is purely a question of the economics. Consider the attenuator whose $T(\lambda)$ appears in Fig. 1-47. Its minimum transmittance is below 10^{-3}. A consequence is that two wedged attenuators in tandem produce an attenuation of 10^{-6}. Suppose that this transmittance of 10^{-6} is produced in one coating. Several questions arise. First, since the coating now contains at least double the original number of layers, namely 240, there is a danger that mechanical stress will cause the coating to crack and delaminate. A second issue is the measurement of a transmittance of 10^{-6}. The scattered light in most spectrophotometers precludes such a measurement. A remaining possibility is to measure T at specific laser lines. The additional cost of building a mechanical structure containing the two slab attenuators must be balanced against the advantages mentioned earlier in this paragraph.

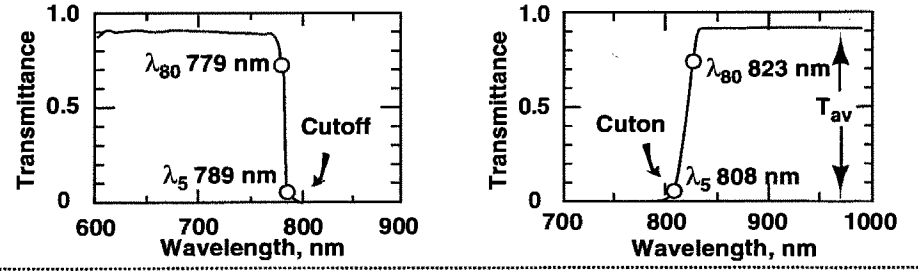


Fig. 1-53 $T(\lambda)$ of edge filters. For each is shown its *cutoff* (*left panel*) or *cuton* (*right panel*) wavelength λ_5. λ_{80} is the point at which the transmittance is 80% of passband average T_{av}. T_{av} will increase after cementing. After Barr [89].

1.5 Filters with broad spectral bandwidth

1.5.1 Introduction

A broadband filter transmits or reflects over a relatively wide range of wavelengths. Unwanted wavelengths are reflected or absorbed. The term *edge filter* generically

describes short-wave pass or long-wave pass filters, such as those in Fig. 1-42. *Dichroic mirrors* are edge filters that function at nonnormal incidence (see §1.8.4).

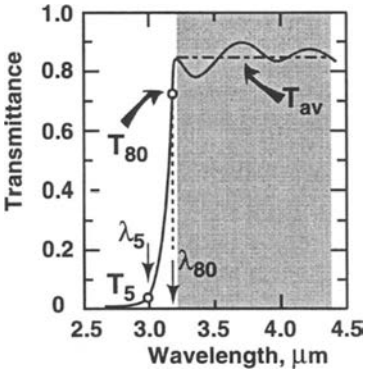


Fig. 1-54 $T(\lambda)$ of an edge filter showing its *cuton* wavelength λ_5, 5% transmittance level T_5 and passband average transmittance T_{av} (*long-short dash*). T_{80} is 0.8 of T_{av} and λ_{80} is the wavelength at which T_{80} occurs. After Eastman [60]. The passband spectral region is shaded.

Other coatings are classified in terms of their final usage. There are *thermal control* coatings that either increase or decrease the heat in (or near) a projection lamp or other radiation source. Examples include the *cold mirror*, *hot mirror* and *regenerative* coatings — they are the focus of §1.5.8 and §1.5.9.

1.5.2 Edge filters

Figure 1-53 depicts $T(\lambda)$ of long-wave pass and short-wave pass edge filters. §1.5.2.1 presents terminology that quantifies their performance.

1.5.2.1 Cuton and cutoff wavelengths

The *cuton* wavelength λ_5 of a long-wave pass edge filter occurs when its *absolute* transmittance is 5%. This is 808 nm in Fig. 1-53 (*right panel*). Cuton implies that this is the *on*set of a spectral region of high transmittance and *absolute* means that its transmittance is as actually measured on a spectrophotometer. It is not *relative* to any other transmittance nor is it corrected in any manner. For the short-wave pass filter in the *left panel* of Fig. 1-53, T drops below 5% at the cut*off* wavelength λ_5 of 789 nm.

1.5.2.2 Slope

Slope is inversely proportional to the abruptness with which transmittance at an edge rises from relatively low values (in its region of attenuation) to high values in the passband region. Slope is computed as follows:

(1) Determine the *cuton* wavelength (λ_5 in Fig. 1-53 or Fig. 1-54);
(2) determine the mean passband transmittance T_{av};
(3) compute the transmittance T_{80} that is 0.8 times T_{av}

$$T_{80} \equiv 0.8\, T_{av}\,; \tag{1-8a}$$

(4) find the wavelength λ_{80} at which T_{80} occurs (this is 823 nm in Fig. 1-53); and
(5) compute percentage slope from

$$\text{slope} \equiv 100\,\frac{\lambda_{80}-\lambda_5}{\lambda_5} = 100\,\frac{823-808}{808} = 1.9\%\,. \tag{1-8b}$$

There is a procedure for finding the slope that uses T_0 (the first transmittance maximum in the passband) rather than T_{av}. It is coincidental that the results would be the same for the filters in Figs. 1-53 and 1-54. The slope of the short-wave pass is

$$\text{slope} \equiv 100\,\frac{\lambda_5-\lambda_{80}}{\lambda_5} = 100\,\frac{789-779}{789} = 1.3\%\,. \tag{1-8c}$$

For the IR filter in Fig. 1-54, the *cuton* λ_5 of this filter is 2.97 μm; T_{av} is 85%. From Eq. (1-8a) T_{80} is 0.8 X 0.85 = 0.68; whence λ_{80} = 3.14 μm. The slope computes to 5.7%.

Slopes of 5% and greater are usually not difficult to achieve. The production of an edge filter with a 2% slope demands exceedingly precise control of the process.

1.5.3 Dichroic coating

A dichroic is an edge filter that functions at nonnormal incidence and is the subject of §1.8.4.

1.5.4 Coatings for the IR

An edge filter for the IR is often used under different conditions than filters for the visible and UV. Quite often an IR filter is inserted into a highly convergent beam of flux (see Fig. 1-17). A consequence is that its angle shift is quite important. An IR filter is frequently cryogenically cooled, which means that its $T(\lambda)$ is measured at low temperatures. Figure 1-55 shows that its $T(\lambda)$ shifts to shorter wavelengths by a few percent as it is cooled to 77 K. Because such a filter is often in proximity to a detector, it is important that its emissivity should be quite low.

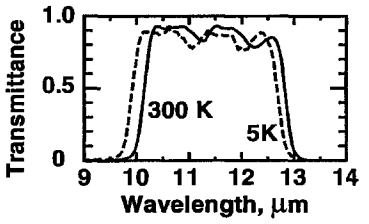


Fig. 1-55 $T(\lambda)$ of a bandpass at ambient temperature (*solid*) and at 5 K (*dashed*). After Stierwalt [91].

1.5.4.1 Long-wavelength pass edge filter for erbium doped fiber amplifier

An edge filter used in an erbium-doped fiber amplifier (depicted in Fig. 1-240) is designed to transmit the longer wavelength band that ranges from 1530 nm to 1565 nm. It should attenuate strongly at shorter wavelengths. The rejected wavelengths originate mostly from the pump laser at wavelengths of 0.98 μm and 1.48 μm. As depicted in Fig. 1-56, its transmittance is well below 0.1% at wavelengths shorter than 1485 nm.

1.5.5 Effect of humidity

It many instances optical coatings are porous and adsorb water vapor from ambient atmosphere. Such adsorption typically shifts a coating's $T(\lambda)$ to longer wavelengths; desiccation shifts it to shorter λs. This is termed "wet-dry shift" of coatings in the trade. Figure 1-57 shows the shift of a coating exposed alternately to air (40% humidity) and dry nitrogen [Pellicori 88]. The shift occurred in a few minutes and was reversible. Table 3-3 summarizes the refractive index shift with humidity of other thin films and multilayers.

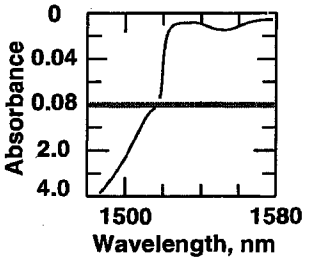


Fig. 1-56 $T(\lambda)$ of an edge filter that transmits the 1550 nm spectral region. It is used for optical fiber communication. The scale of the ordinate changes at 0.08. After Optical Coatings Japan [99].

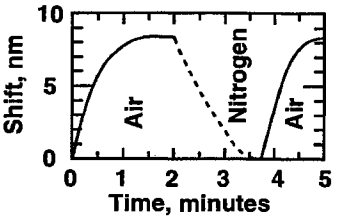


Fig. 1-57 Shift to shorter wavelengths of the cuton of an edge filter when exposed to humid air (*solid*) and dry nitrogen (*dashed*). After Pellicori [88].

A typical shift is 1% and is particularly noticeable in a coating containing oxide and fluoride materials (see §3.4.3.2). After a coating is removed from the vacuum environment of a coating tank, a standard procedure is to let it sit in ambient atmosphere (thereby permitting it to adsorb water vapor) before its spectral transmittance is measured.

1.5.6 Specification of an edge filter

A specification for an edge filter could include:

- cuton (or cutoff) wavelength,
- wavelength at which $T = 50\%$
- slope
- passband — wavelength range
- passband — average transmittance
- out-of-band — wavelength range
- out-of-band — average transmittance
- out-of-band — maximum transmittance
- cone angle of incident flux
- operating temperature
- emissivity
- allowed amount of optical absorption
- chromaticity coordinates

It is quite important that parameters be *toleranced*. To illustrate, a specification should read "the cuton wavelength is 7.8 μm ± 0.1 μm" or that "the slope is 5% ± 0.5%." Optical absorption is important if the coating is inserted into a high irradiance beam, such as a laser. Clear aperture should also be mentioned (see §1.9.1).

It is also important that one does not specify parameters that either cannot be measured or are quite costly to measure. An example is a transmittance of 10^{-7}.

A specification of an IR filter might read:

> Average transmittance exceeds 90% in the spectral region of 7.7 μm to 12.2 μm. The 50% transmittance point is between 7.85 μm and 7.9 μm. Average transmittance at wavelengths below 7.2 μm is less than 0.1%.

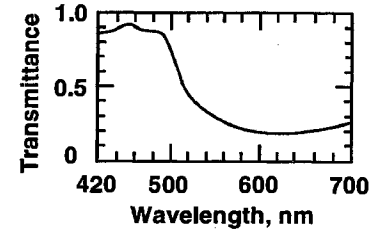


Fig. 1-58 Transmissive multilayer corrects color temperature of an incandescent light from 3200 K to 5500 K. After Barkley [70].

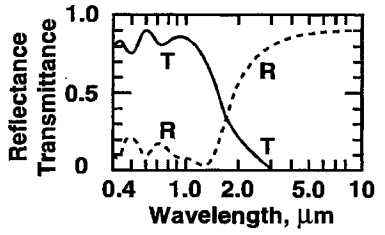


Fig. 1-59 $T(\lambda)$ and $R(\lambda)$ of a single layer deposited on glass. The composition of the film is a mixture of tin oxide and indium oxide.

1.5.7 Edge filter for color control

Dielectric multilayers are deposited upon the interior of colored spotlights that transmit at least 85% in the blue but less than a few percent in the red and green. This multilayer not only produces a more saturated hue than glass absorption filters, but, more important, a glass filter placed in proximity to a 150 watt tungsten filament would absorb heat and thermally fracture.

Multilayers are used in illuminators that correct the color temperature of tungsten (3200 K) to 4400 K [Barkley 70]. A manifestation of 3200 K color temperature is its higher output of red light, as compared to blue. A remedy is to add a color correction filter (see Fig. 1-58) that transmits more blue than red. Illuminators in surgical rooms correct to an even higher color temperature than 4400 K.

1.5.8 Thermal control coatings

Thermal control generically describes coatings that selectively absorb, reflect or transmit IR in preference to visible or other wavelengths. They are used principally in illuminators in projection systems.

Figure 1-59 shows $T(\lambda)$ of a thermal control coating consisting of a mixture of oxides of tin and indium. Vossen [77] describes how such coatings are mass-produced by sputtering or chemical deposition.

1.5.8.1 Cold mirrors for projection systems

Radiation from a tungsten lamp contains at least six times as much heat radiation (i.e., IR) as useful light in the visible. The term *cold light* was coined by H. Schröder [49] to describe radiation in which the heat — the IR portion — is removed. A producer of such *cold light* is an all-dielectric multilayer, called a *cold mirror.* It reflects visible light and transmits IR [H. Schröder 49; Koch 51; Carlson 55; Schroeder 60; Scoledge 64; Koch 65b].

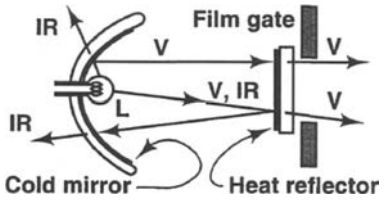


Fig. 1-60 In a projection system, a cold mirror reflects mostly visible light *V* from lamp filament *L* to film gate. The heat reflector reflects the IR and transmits visible light.

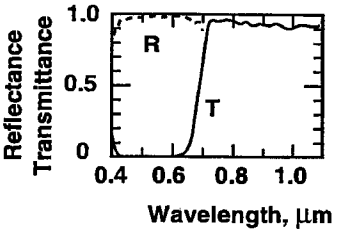


Fig. 1-61 For a cold mirror at 45° angle of incidence, reflectance *R* and transmittance *T* vs wavelength. After Balzers [92].

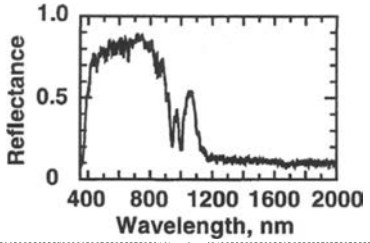


Fig. 1-62 $R(\lambda)$ of a multilayer cold mirror consisting of hundreds of layers of extruded plastic [Alfrey 69, 73; Dow 93].

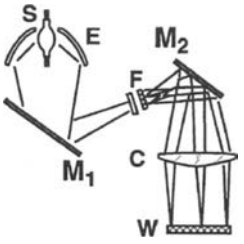


Fig. 1-63 Elliptical reflector *E* focuses a deep UV radiation source *S* onto fly's eye lens *F* and condenser *C* to irradiate wafer *W*. $R(\lambda)$ of mirrors M_1 and M_2 is in Fig. 1-64. After Sawamura [82].

Fig. 1-60 shows a cold mirror reflector behind a light source in a film projector. That mirror reflects visible light towards the film gate, but permits heat in the IR to escape. In typical systems, the heat reaching the film gate is reduced by a factor of three. Figure 1-61 shows $R(\lambda)$ and $T(\lambda)$ of a cold mirror that functions at 45° angle of incidence.

Paragraph (a) of §1.1 mentions a multilayer reflector produced by co-extrusion of plastic materials [Alfrey 69, 73]. An advanced version of this process produced a broadband reflector whose $R(\lambda)$ appears in Fig. 1-62 [Dow 93]. It could easily contain 800 layers.

Figure 1-63 delineates a projection system that extracts the 254 nm wavelength from a mercury arc lamp and directs it onto a semiconductor wafer. M_1 and M_2 are multilayer dielectric "UV cold mirrors" that reflect strongly at 254 nm and, as shown in Fig. 1-64, reflect much less at longer wavelengths.

1.5.8.2 Heat reflectors for projection systems

A cold mirror becomes even more efficient when a *heat reflector* is inserted at the film gate, as shown in Fig. 1-60 [Dimmick 52a]. This reflector, which is sometimes called a *hot mirror*, has a high transmittance in the visible spectral region and a substantial IR reflectance. Its efficiency depends upon the number of layers. An illustration is Fig. 1-65, which shows that the wavelength span of the rejection region of a reflector containing at least 50 layers exceeds that of a coating containing fewer layers.

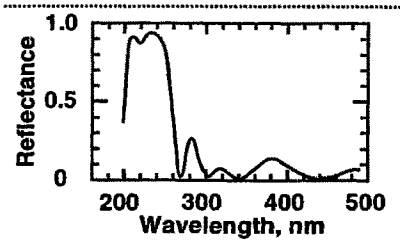


Fig. 1-64 $R(\lambda)$ of a multilayer reflecting 254 nm and transmitting longer wavelengths. After Sawamura [82].

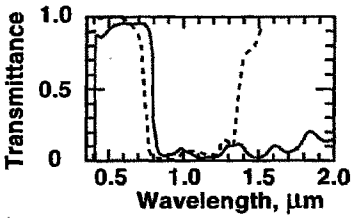


Fig. 1-65 $T(\lambda)$ of multilayer dielectric heat reflectors containing an estimated 25 layers (*dashed*, Thelen [66a]) and 50 layers (*solid*, Deposition Sciences [99]).

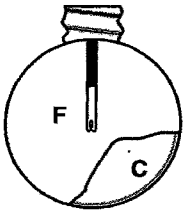


Fig. 1-66 Regenerative coating *C* (*shaded*) on interior of spherical glass shell reflects IR back onto incandescent filament *F*. After Hoffman [22].

1.5.9 Regenerative coatings

Regenerative generically describes a coating that increases the efficiency of a source of radiation. It is assumed that there is a coated, transparent shroud surrounding the radiation source (i.e., filament, flash lamp).

1.5.9.1 Incandescent tungsten source

As depicted in Fig. 1-66, Hoffman [22] deposited a regenerative coating on the interior of a quasi-spherical light bulb with an incandescent tungsten filament at its center. The IR is retro-reflected to the filament and visible radiation escapes.

A coating deposited on an envelope of a quartz-halogen tungsten lamp is a more modern version (see Fig. 1-67). Figure 1-68 depicts $R(\lambda)$ of this coating and the spectral radiance of a 2700 K tungsten filament. The gain of this system is typically 1.35. That is, 35% less electrical input power is consumed to produce the same amount of visible light. Coatings are also deposited on the exterior of cylindrical quartz-halogen lamps [Rancourt 87]. This coating reflects *unwanted* radiation back into the source and permits *wanted* radiation to escape. Table 1-3 lists radiation sources with regenerative coatings.

1.5.9.2 Flashlamp coating

A tubular gas discharge flash lamp in Fig. 1-69 pumps a Nd:glass laser. Figure 1-70 depicts those pump bands and $T(\lambda)$ of a regenerative coating — an all-dielectric multilayer that reflects unwanted UV into the plasma [Baranov 78]. This coating improves lamp efficiency by about 60% at high current densities.

1.5.9.3 IR beacon

The IR beacon emits a time-modulated signal that confuses and thwarts heat-seeking missiles, such as the Sidewinder and Stinger. This beacon emits at wavelengths longer than 2 µm. A regenerative coating reflects shorter wavelengths into the lamp, thereby increasing its efficiency.

1.5.10 Coatings for solar cells

Several coatings improve the electrical efficiency of solar cells that power space vehicles. A thin glass or fused silica plate (i.e., cover) is bonded with organic cement to the exterior of the cell. If needed, the exterior of the cover is coated with a mixture of

magnesium fluoride and fused silica that enhances its emissivity, thus permitting it to radiate more efficiently to outer space [Krisl 88]. Silicon photodiodes operate more efficiently at a cooler temperature.

Table 1-3 Regenerative coatings

	Wavelength range of:	
Radiation source	Wanted radiation	Unwanted radiation
Tungsten incandescent	400 nm to 700 nm	700 nm to 2000 nm
Xenon flash lamp	550 nm to 950 nm	300 nm to 420 nm
IR beacon	2 μm to 3 μm	0.4 μm to 2 μm

Regenerative coatings are applied to radiation sources that transmit *wanted* radiation and reflect (back into the source) the *unwanted*.

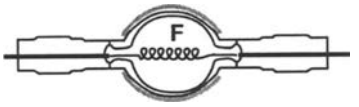


Fig. 1-67 Regenerative coating (*shaded*) on the exterior of a fused silica shell reflects the infrared onto an incandescent filament *F*. After Bergman [91].

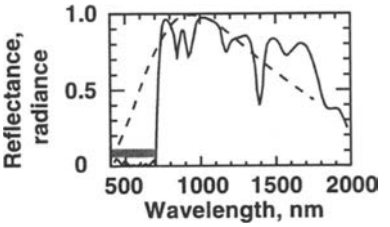


Fig. 1-68 (*Solid*) Measured $R(\lambda)$ of regenerative coating applied to exterior of incandescent tungsten light source. (*Dashed*) Relative radiance per unit wavelength interval of tungsten at 2700 K. After Bergman [91].

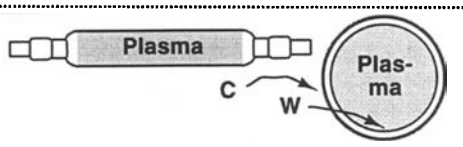


Fig. 1-69 Plasma (*shaded*) forms in the interior of ignited flash lamp. A regenerative multilayer coating *C* deposited on exterior of wall *W* reflects unwanted radiation into the plasma.

A dielectric multilayer is deposited on the interior surface of the cover, i.e., contiguous to the organic cement. This coating, the spectral transmittance of which appears in Fig. 1-71, serves to both reflect and attenuate UV at wavelengths below 350 nm. The UV is undesirable because it produces little electrical power and decomposes the organic cement. A "red-blue" coating rejects both UV and wavelengths longer than 1.2 μm, where a silicon cell responsivity is zero [Thelen 61]. The responsivity of a gallium arsenide solar cell cuts off at shorter wavelengths than a silicon cell. Figure 1-72 shows a "blue-red" cover for such a cell [Deposition Sciences 99].

1.5.11 Thermal control for satellites

A thermal control coating enables an orbiting satellite to operate at a lower temperature. This coating (shown in Fig. 1-73) has two purposes: (1) It should reflect as much solar radiation as possible; and (2) its emissivity (weighted for a 300 K Planckian radiator) should approach 1.00.

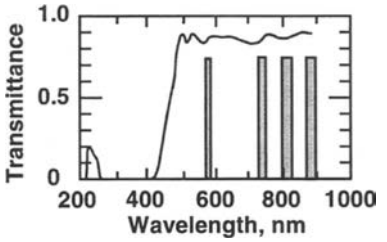


Fig. 1-70 (*Solid curve*) $T(\lambda)$ of multilayer dielectric UV-reflecting stack. *Shaded bars* — pump bands of Nd-doped glass. After Baranov [78].

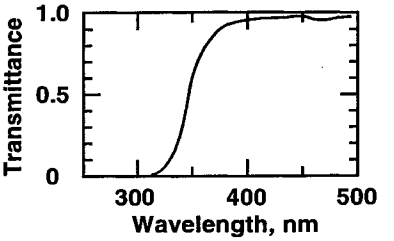


Fig. 1-71 $T(\lambda)$ of UV-reflecting (i.e., "blue") solar cell cover. After Rancourt [81].

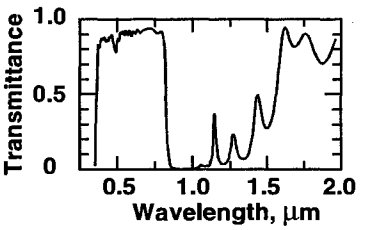


Fig. 1-72 $T(\lambda)$ of "red-blue" solar cell cover for a gallium arsenide solar cell. After Deposition Sciences [99].

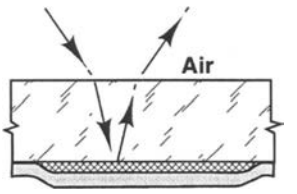


Fig. 1-73 Thermal control coating consists of second-surface silver layer (*cross-hatched*) and protective nickel film (*dotted*). Reflections at the outer interface are not shown.

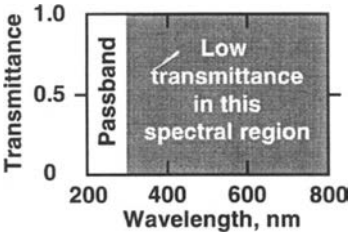


Fig. 1-74 Transmittance of a coating that transmits 200 nm to 300 nm and attenuates at wavelengths longer than 300 nm.

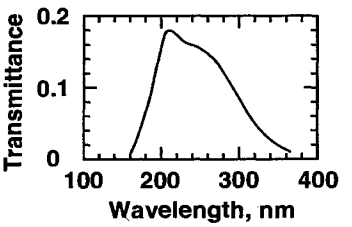


Fig. 1-75 Measured $T(\lambda)$ of short-wavelength pass filter. After Acton [90].

The rear of a glass substrate is coated with a highly reflective silver film, which is overcoated with a protective layer of a nickel alloy. Because the silver is opaque, the nickel does not alter properties of the silver, which reflects at least 90% of the solar energy. Glass has an emissivity in excess of 85%.

1.5.12 Short-wavelength pass — a case study

The problem posed is isolation of the passband spectral region — from 200 nm to 300 nm. Given the bounds in Fig. 1-74, T should be below 1% at wavelengths longer than the passband. There are three possible coatings: (1) All-dielectric multilayer; (2) bandpass containing metal layers; or (3) reflection filter.

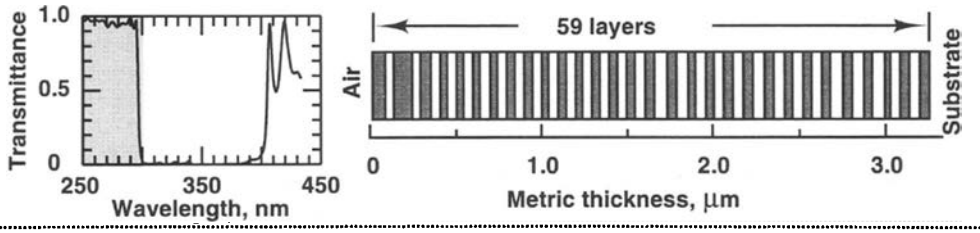


Fig. 1-76 Computed $T(\lambda)$ and layer thicknesses of an edge filter whose design appears in §1.10.2.4. The passband spectral region is shaded.

Figure 1-75 shows $T(\lambda)$ of a short-wave pass coating containing 59 layers. The high index material is yttria, which is non-absorbing at wavelengths longer than 240 nm. A problem is that the rejection region barely reaches to 400 nm. Extension of the rejection region beyond 400 nm mandates the addition of even more layers.

The bandpass in Fig. 1-76, which contains layers of aluminum, probably manifests the required rejection at wavelengths longer than 400 nm. A problem is that absorption of the aluminum severely limits the passband transmittance.

A conclusion is that a reflection filter (see §1.4.3, Figs. 1-45 and 1-64) has the greatest chance of producing the required passband transmittance and offband rejection.

1.6 Bandpasses

1.6.1 Applications

Bandpasses with passband widths of 5 to 10 nm are widely used in analytical instruments. Narrowband filters are used in conjunction with gallium arsenide lasers in distance measuring devices and interferometers, as well as surveillance systems. A detector in a laser-driven point-of-sale scanner is protected with a bandpass that suppresses ambient light. Filters are widely used in instruments that analyze combustion products and atmospheric pollutants. Wavelength division multiplexing filters are important components of fiber-optic communication systems.

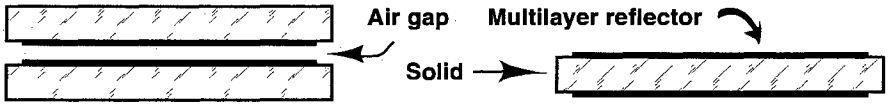


Fig. 1-77 In a single-cavity bandpass, reflectors surround the spacer, which is either an air gap or solid material.

1.6.2 Single-cavity bandpass filters

The simplest bandpass is fabricated by arraying reflectors at both ends of a resonant cavity, which is also called a *spacer*. Fig. 1-77 shows that the cavity may be filled with air or a solid material, such as glass, mica or an evaporated layer. Those reflectors are typically dielectric multilayers or semitransparent metal films such as silver or aluminum. Most of our attention focuses on coatings containing an evaporated layer for a spacer.

1.6.2.1 Filters with non-evaporated cavities

The bandpass that does not contain an evaporated spacer layer traces its development to the interferometer of Fabry and Perot [Born 75, §7.6.2; Stone 63]. Its advantage is that

an exceedingly narrow bandwidth can be produced. A disadvantage is that it is expensive because of the extreme flatness of the surfaces. It is usually difficult to isolate a selected narrow transmission band such that other (i.e., spurious) transmission bands are suppressed.

Figure 1-77 shows a single-cavity bandpass in which the cavity is an air gap. Each plate is optically polished to a flatness of at least $\lambda/50$ and is coated with multilayer dielectric reflectors with the following attributes: (1) Absorption and scatter are minimal; (2) coating thickness is extremely uniform; and (3) coatings must have minimal mechanical stress. Coating thickness nonuniformity broadens the spectral bandwidth of this bandpass (see § 1.6.6.3), as does warpage of the plates by mechanical stress [Greet 86].

Figure 1-77 depicts a cavity consisting of a solid material, such as fused silica [Austin 71, 72a, 73] or a thin sheet of mica [Dobrowolski 59, 62, 68; Holmes 64]. An example is a bandpass transmitting the atomic hydrogen line at 653 nm with a bandwidth as narrow as 0.05 nm [Daystar 88]. An advantage of the solid spacer is that its *angle shift* (see §1.6.4.1) is at least two times smaller, as compared with the air-spacer configuration mentioned above.

A bandpass has been constructed that uses a thin fiber bundle as a spacer [Schineller 71]. Each fiber functions as a dielectric waveguide that limits the propagation of the light to a single mode. The advantage is that there is no shift (see §1.6.4.1) of the wavelength of the passband with increasing angle of incidence of the light that illuminates it. The reason is that the phase delay of the wave propagating through the spacer is determined by the mode of the waveguide. Since only one mode is permitted, changing the angle of incidence does not alter that mode.

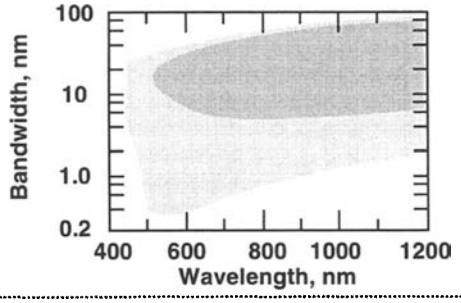


Fig. 1-78 Achievable bandwidth [see Eq. (1-9)] at various wavelengths for all-dielectric bandpasses with high index spacers (heavy shading) and low index spacers (light shading). After Barr [89].

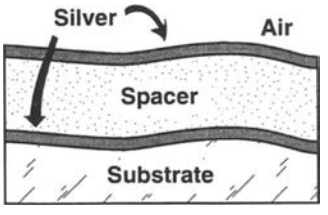


Fig. 1-79 Cross section of metal-dielectric-metal bandpass. Lack of planeness of the substrate is exaggerated for purposes of illustration.

1.6.2.2 Bandwidth of bandpasses

Figure 1-78 depicts the achievable bandwidth of bandpasses at various wavelengths in the visible and near IR. The bandpasses contain only vacuum deposited layers.

1.6.2.3 MDM filters Figure 1-79 shows a **MDM** bandpass consisting of a dielectric spacer layer **D** sandwiched between silver layers **M**. Prior to invention of the **MDM** [Geffcken 42, 48, 54] in the late 1930s, absorption filters (see §1.4.1) were the widely

used means of isolating a relatively narrow spectral region. Hence the **MDM** was called an *interference filter*. This terminology is no longer used.

Unlike the bandpasses of §1.6.2.1, both the metal **M** layers and the spacer are deposited by evaporation in a vacuum. In Fig. 1-79, substrate waviness is exaggerated for purposes of illustration. The coatings contour the substrate and hence the relative spacing of the **M** layers remains constant, even though the substrate may deviate many wavelengths from planeness. Hence filters can be deposited upon fire polished glass, provided its passband is not particularly narrow and the optical system is nonimaging.

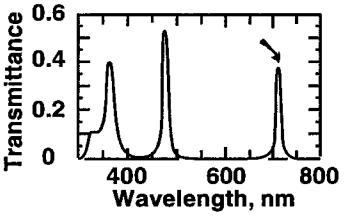


Fig. 1-80 $T(\lambda)$ of **MDM** type bandpass with second-order transmission passband (*arrowed*) at 710 nm. After Bausch [60].

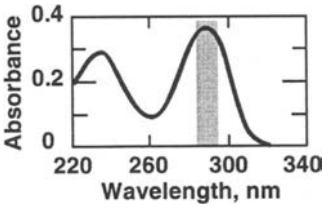


Fig. 1-81 Absorbance of uric acid in aqueous solution. The shaded area is its absorption band.

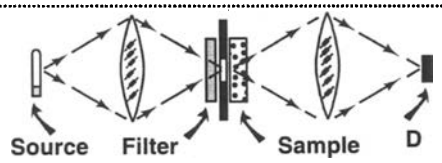


Fig. 1-82 Use of a bandpass filter in a system to measure concentration of a chemical component in the sample, showing transfer lenses, source of radiation and detector of radiation *D*.

Figure 1-80 delineates $T(\lambda)$ of a **MDM** produced in the 1950s. It has been supplanted by all-dielectric bandpasses. The **MDM** has evolved into designs such as **MDMDMDM** that contain not just two but many silver or aluminum layers in order to attenuate in the offband part of the spectrum. §1.9.2 describes the technique of protecting these coatings with an optical cement.

1.6.3 Classes of bandpass filters

There are two broad categories of bandpass filters: (1) Conventional filters; and (2) Wavelength Division Multiplexing (WDM) filters. Because their attributes are quite different, conventional filters are described in §1.6.4 and WDM filters in §1.6.7.

1.6.3.1 Conventional bandpass filters

The descriptor *conventional* is used for lack of a better word. An example of the use of this type of filter is to measure the concentration of a chemical component in a sample, such as uric acid. Figure 1-81 illustrates that uric acid absorbs radiation in a band centered at 290 nm. Figure 1-82 shows schematically the apparatus that measures the concentration of uric acid in a solution. A lens passes radiation at 290 nm through both a bandpass filter and sample, while another lens conveys the radiation to a detector. Beer's Law predicts that the logarithm of the output of the detector of radiation *D* is proportional to the concentration of uric acid. As mentioned in §1.6.5, it is quite important that the bandpass filter in Fig. 1-82 be adequately *blocked*, so that virtually no radiation other than the narrow 290 nm band is transmitted through the sample.

1.6.3.2 Bandpass filters for wavelength multiplexing or wavelength demultiplexing

The objective is use multilayer bandpass filters to increase the information carrying capacity of a fiber by sending two or more wavelengths over the same fiber. This is termed wavelength division multiplexing (WDM). The measured $T(\lambda)$ of one of those filters appears in Fig. 1-83. Figure 1-84 shows a demultiplexer that separates the signals into the ITU communication channels that, at the time of this writing (in the year 2000), are separated by 100 GHz. This translates to approximately 0.8 nm in the spectral region near 1552 nm.

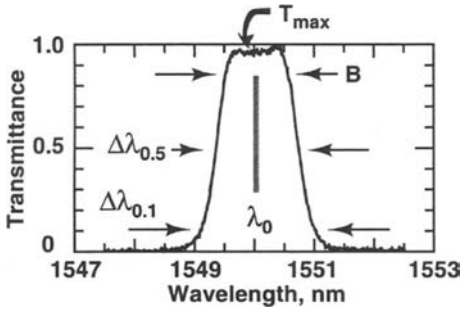


Fig. 1-83 $T(\lambda)$ of a bandpass, showing its maximum passband transmittance T_{max}, passband width B as a WDM, semiwidth passband width $\Delta\lambda_{0.5}$, deciwidth of passband $\Delta\lambda_{0.1}$ and location of center of passband λ_0. After [Matl 98].

An input fiber, containing signals in many channels, impinges upon a bandpass filter that transmits channel 31 (CH 31). The bandpass is an all-dielectric multilayer that has a quite high reflectance, except in the narrow spectral region where channel 31 lies. The computed bandpass transmittance appears in Fig. 1-85. The radiant flux in channels 32, 33, 34, 35 and so on, is reflected from that bandpass. When that reflected radiant flux impinges upon the filter at channel 32, the flux in the spectral region of channel 32 propagates into the fiber (CH 32) and the remaining flux in channels 33, 34, 35 and so on is reflected. The foregoing process is repeated in order to isolate each of the channels.

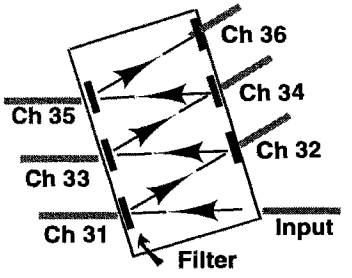


Fig. 1-84 Use of bandpass filters (*solid black rectangles*) for wavelength division demultiplexing, showing the input fiber and output fibers containing the demultiplexed signals.

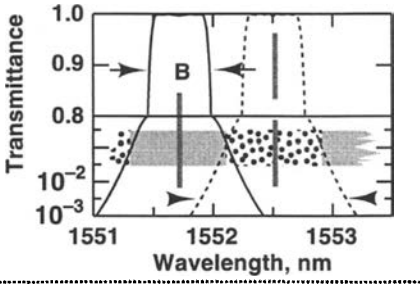


Fig. 1-85 Computed $T(\lambda)$ of a bandpass BP that transmits ITU channel 32 (*dashed*). Another BP transmits channel 31 (*solid*). Alternating dotted and shaded areas along the abscissa represent ITU communication channels 33 to 30. B is the bandwidth at the –0.5 dB level.

The description above of the demultiplexing "box" is rather oversimplified. For example, omitted from the diagram are relay lenses (see Fig. 1-36) that transfer the signal from fiber to fiber and filter to filter. Further information related to these filters appears in §1.6.7.

1.6.4 Attributes of conventional bandpass filters

The attributes that describe a conventional bandpass include:

(a) **Passband transmittance** If the spectral width of the passband is relatively narrow, as shown in Fig. 1-83, maximum transmittance T_{max} at the peak is a useful parameter. The passband average transmittance T_{av} is more meaningful for a relatively broad bandwidth, as depicted in Fig. 1-86.

(b) **Wavelength of the passband** The wavelength λ_0 at which T_{max} occurs can be specified, as depicted in Fig. 1-83. Alternatively, *cuton* and *cutoff* wavelengths may be specified — see Fig. 1-86. *Cuton* and *cutoff* have the same meaning as in §1.5.2.1.

(c) **Bandwidth** For filters with relatively narrow passbands, it is customary to specify *semiwidth* $\Delta\lambda_{0.5}$, which is full passband width where the transmittance is 0.5 T_{max}, as illustrated in Fig. 1-83. This parameter is also known as FWHM (Full Width at Half Maximum) or simply the bandwidth BW. The "B" label in Fig. 1-83 pertains to a bandpass for WDM and is not applicable here.

Other useful parameters are its *deciwidth* $\Delta\lambda_{0.1}$ (full passband width where T is one-tenth of T_{max}) or its *centiwidth* $\Delta\lambda_{0.01}$ (full passband where T is one-hundredth of T_{max}) and so on.

Semiwidth, deciwidth and so on are measured *relative* to peak passband transmittance, whereas spectral bandwidth between the *cuton* and *cutoff* wavelengths is measured where T is five percent, regardless of peak transmittance.

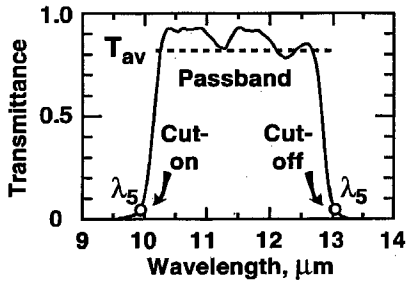


Fig. 1-86 Attributes of a relatively broad bandwidth bandpass, showing its *cuton* and *cutoff* wavelengths λ_5 (where $T = 0.05$) and average passband transmittance T_{av}.

(d) **Fractional bandwidth** This dimensionless parameter quantifies spectral width. Fractional bandwidth FBW is

$$\text{FBW} \equiv \Delta\lambda_{0.5} / \lambda_0 \qquad (1\text{-}9)$$

and is proportional to the bandwidth of a bandpass.

(e) **Resolution** Resolution Q is the reciprocal of the fractional bandwidth

$$Q \equiv \lambda_0 / \Delta\lambda_{0.5} = (\text{FBW})^{-1} \qquad (1\text{-}10)$$

and is a merit function. An illustration is the $T(\lambda)$ in Fig. 1-83, which shows that $\lambda_0 = 1550$ nm and $\Delta\lambda_{0.5} = 1.4$ nm; hence $Q = 1500/1.4 = 1107$. The filters in Fig. 1-85 have narrower passbands and manifest an even larger Q. If this bandpass is used for fiber-optic communication, however, the definition of bandwidth in paragraph (b) of §1.6.7 should be used.

(f) **Shape factor** The shape factor S_1 is the ratio of the deciwidth to the semiwidth

$$S_1 \equiv \Delta\lambda_{0.1} / \Delta\lambda_{0.5} \qquad (1\text{-}11a)$$

and is an inverse measure of how precipitously $T(\lambda)$ at the passband edges plunges to low values. In other words, a bandpass with a relatively small S_1

manifests a passband with a rectangular shape. Other shape factors for conventional bandpasses are:

$$S_2 \equiv \Delta\lambda_{0.01} / \Delta\lambda_{0.5} \quad \text{and} \tag{1-11b}$$

$$S_3 \equiv \Delta\lambda_{0.001} / \Delta\lambda_{0.5} \ . \tag{1-11c}$$

For bandpasses used for fiber wavelength division multiplexing, S_3 is defined as

$$S_{25} \equiv \Delta\lambda_{-2.5dB} / \Delta\lambda_{-0.5dB} \tag{1-11d}$$

where, as discussed in §1.6.7, $\Delta\lambda_{-0.5dB}$ is the full spectral width of the passband at –0.5 dB (i.e., 0.05 absorbance units or 89.1%) below T_{max}.

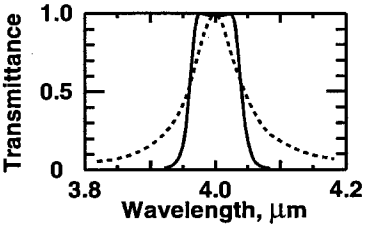


Fig. 1-87 $T(\lambda)$ of a single-cavity bandpass (*dashed*) and triple-cavity bandpass (*solid*). The latter design is in §1.10.2.16.

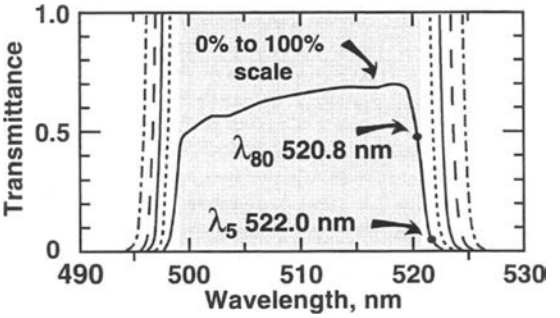


Fig. 1-88 Measured $T(\lambda)$ of a bandpass on a scale of 0 to 100% (*solid*), 0 to 10% (*short dash*), 0 to 1% (*solid*), 0 to 0.1% (*long dash*), 0 to 0.01% (*long-short dash*). After Barr [89]. The shaded area is the passband.

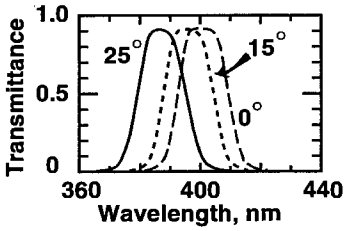


Fig. 1-89 Measured $T(\lambda)$ of a bandpass at incidence angles of 0° (*long dash, n*), 15° (*short dash*) and 25° (*solid*). After Coherent [92].

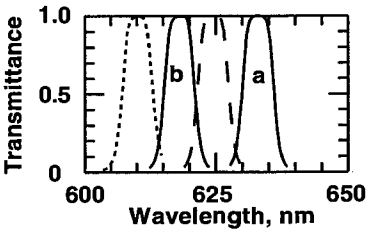


Fig. 1-90 Computed $T(\lambda)$ at incidence angles of zero (*a, solid*), 15° (*long dash*), 20° (*b, solid*) and 25° (*short dash*) Its design appears in §1.10.2.16.

Table 1-4 Bandwidth and shape factor of bandpasses

Type	Semiwidth	Deciwidth	Shape factor	Curve
Single cavity	83 nm	256 nm	3.1	Dashed
Triple cavity	80 nm	110 nm	1.4	Solid

Semiwidth, deciwidth and shape factor of bandpasses whose $T(\lambda)$ and designs appear in Fig. 1-87.

[(f) **Shape factor continued**] Figure 1-87 compares $T(\lambda)$ of single and triple-cavity bandpasses with identical semiwidth. Table 1-4 shows that the deciwidths are quite different, however. The shape factors of the single and triple cavities are 3.1 and 1.4, respectively. Figure 1-87 illustrates that the smaller shape factor of the triple cavity is manifested as a passband that is much "squarer" than that of the single cavity. Figure 1-88 depicts $T(\lambda)$ of a bandpass whose passband spectral shape is, indeed, much "squarer" than that of any bandpass in Fig. 1-87. This is discussed in §7.6.1.2. Figures 7-40 and 7-41 present additional illustrations.

Table 1-5
Angle shifts of bandpasses

Type of bandpass	C_a (degrees)$^{-2}$
UV, low index spacers	-67×10^{-6}
Visible, low index spacers	-59×10^{-6}
Visible, high index spacers	-45×10^{-6}
IR, low index spacers	-23×10^{-6}
IR, high index spacers	-16×10^{-6}

Typical values of C_a in Eq. (1-12) for all-dielectric multiple-cavity bandpasses.

(g) **Angle shift, temperature shift and étendue.** These parameters are discussed subsequently in §1.6.4.1 and §1.6.6.3.

(h) **Clear aperture** This determines the usable area of the filter (see §1.9.1).

1.6.4.1 Angle shift

The wavelength of the passband center shifts by an amount $\Delta\lambda_a$ to *shorter* wavelengths as the angle of incidence increases from zero to a maximum angle of θ_m. Figure 1-89 shows a measured $T(\lambda)$ at several angles of a bandpass composed of hafnia and silica. Computed $T(\lambda)$ curves appear in Fig. 1-90. At high angles, such as 40°, there are two passbands — one for each linear polarization. At lower angles, however, this polarization splitting may be neglected.

The fractional angle shift $\Delta\lambda_a / \lambda_0$ of a commercially produced bandpass appears in Fig. 1-91. There are two regions: (1) Larger angles (*shaded* in Fig. 1-91) — the dependence is relatively linear; and (2) when θ is close to zero (region Q in Fig. 1-91), a quadratic dependence is evident. For angles ranging from 0° to 20°, the angle shift is approximated by

$$\Delta\lambda_a \approx C_a \lambda_0 \theta^2 \qquad (1\text{-}12)$$

where the constant C_a depends upon the construction of the bandpass. Table 1-5 lists this constant for several types of filters. In the IR, coating materials of quite large refractive index are used which, in turn, reduces angle shift.

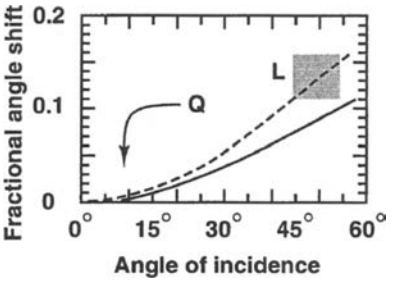


Fig. 1-91 Fractional angle shift $\Delta\lambda_a/\lambda_0$ vs θ_0 of all-dielectric bandpass with low index spacers (*upper, dashed*) and high index spacers (*lower, solid*). After Baird-Atomic [65].

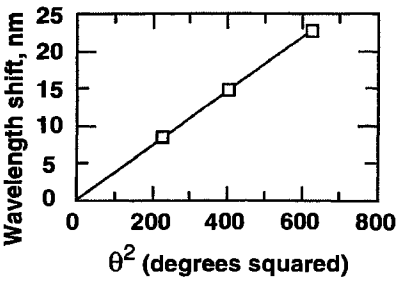


Fig. 1-92 Shift $\Delta\lambda_a$ of passband wavelength of the bandpass captioned in Fig. 1-90 vs square of incidence angle θ_0.

One method of finding C_a is to plot the magnitude of the angle shift vs "theta squared," as shown in Fig. 1-92. The slope of the straight line is proportional to C_a. The points in Fig. 1-92 are gleaned from Fig. 1-90. The slope of the line is 0.0365 nm per reciprocal degree squared. Dividing the number in the foregoing sentence by λ_0, we obtain 58×10^{-6} for C_a, which is close in magnitude to the value of 59×10^{-6} in Table 1-5 for "Visible, low index spacers." The angle shift is reexamined when the transmittance of the bandpass in convergent flux is computed in §1.6.6. The angle shift is often used to "tune" a bandpass. That is, a bandpass centered at longer wavelengths may be tilted to shift its passband wavelength into specification.

1.6.4.2 Temperature shift

The shift with temperature of the cuton and cutoff wavelengths of an edge filter is mentioned in §1.5.4 (see Fig. 1-55). It is also discussed in §3.4.4. The temperature shift of a bandpass is increasingly important as its passband width shrinks. This shift is attributed to several causes:

(a) Irreversible chemical and physical changes. For a noncemented filter, this is illustrated by the dashed curves in Fig. 3-36. Baking a cemented bandpass, however, has a different effect. This causes organic cement to penetrate porous films; an irreversible shift is observed [Meaburn 66; Title 74a, 74b]. Levina [73] attributed drift of the center wavelength to diffusion of water vapor through the optical cement and into the coating. This typically shifts the passband to longer wavelengths.

(b) Reversible changes of refractive indices of the layers due to adsorption of water vapor or other vapors. This is similar to the effect shown in Fig. 1-57 for an edge filter. Heating a filter makes it lose water vapor, thereby *reducing* the refractive index, which in turn shifts its $T(\lambda)$ to *shorter* wavelengths. Effects (c) and (d) below have the opposite effect and shift the passband to *longer* wavelengths.

(c) Change of refractive index with temperature. In most materials (with the exception of lead telluride) this is a positive change, as illustrated in Table 3-4. A fractional shift of $\Delta\lambda/\lambda_0$ of 7.0×10^{-5} / °C is observed for bulk germanium.

(d) Change of physical thickness of a layer with temperature.

Effects (c) and (d) combine to produce a positive temperature coefficient in most filters — a higher temperature shifts λ_0 to longer values. A value of 5×10^{-5}/°C for $\Delta\lambda/\lambda_0$ for filters in the visible has been reported [Roche 76; Pelletier 74]. Bifford [66] measured a shift of 2.5×10^{-5}/°C for a bandpass at 1 μm. Mauer [59] reported a $\Delta\lambda/\lambda_0$ of 1.2×10^{-4}/°C for germanium-silicon monoxide filters in the IR. A conclusion is that ultra-narrowband bandpasses should be inserted into a temperature-stabilized oven.

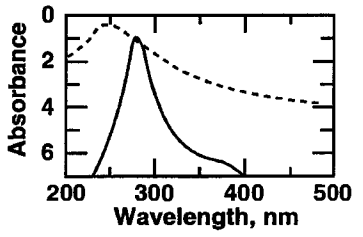


Fig. 1-93 Absorbance vs λ of UV bandpasses. (*Solid*) after Barr [68] and (*dashed*) after Acton [90].

1.6.5 Blocking of a bandpass

Blocking is slang for the art of attenuating unwanted flux that a bandpass transmits in offband spectral regions. $T(\lambda)$ curves of Fig. 1-93 show, for example, that one bandpass is blocked more thoroughly — at the price of lower peak passband transmittance. $T(\lambda)$ of the bandpass captioned in Fig. 1-90 appears on an expanded scale in Fig. 1-94, which shows that T drops below 6 absorbance units at wavelengths ±10% on both sides of the passband. At wavelengths ±15% of the center, however, T rises rapidly; additional blocking is necessary. Blocking filters invariably decrease the peak transmittance in the passband.

1.6.5.1 Level of blocking required

Suppose a bandpass is used in a photometer that determines quantitatively the concentration of a chemical that absorbs at λ_0. Beer's Law predicts that concentration is proportional to absorbance at λ_0. A photometer consists of an incandescent tungsten filament, absorber, bandpass and silicon photodiode detector. The "wanted signal" is the flux Φ_s collected by the detector:

$$\Phi_s = \int_{\lambda_2}^{\lambda_3} S(\lambda)\, T_{max}\, d\lambda \propto T_{max}\ \Delta\lambda_{0.5} \tag{1-13}$$

where $S(\lambda)$ is the product of its responsivity and the radiance of tungsten; T_{max} is peak passband transmittance; and λ_2 and λ_3 are spectral limits of the passband, as depicted in Fig. 1-95. The unwanted offband flux Φ_u is obtained by a similar integration:

$$\Phi_u \propto \int_{\lambda_1}^{\lambda_2} S(\lambda)\, T(\lambda)\, d\lambda + \int_{\lambda_3}^{\lambda_4} S(\lambda)\, T(\lambda)\, d\lambda \tag{1-14}$$

where λ_1 and λ_4 are the limits of nonzero responsivity of the photodetector. As an approximation, the variation of $S(\lambda)$ is assumed to be relatively constant and transmittance in the offband region is at a level $T_{offband}$. Equation (1-14) is written

$$\Phi_u \propto T_{offband}\ (\lambda_2 - \lambda_1 + \lambda_4 - \lambda_3) \propto T_{offband}\ (\lambda_4 - \lambda_1), \tag{1-15}$$

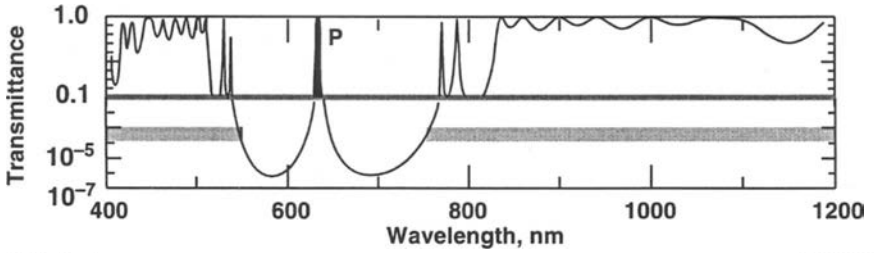


Fig. 1-94 $T(\lambda)$ of the bandpass in Fig. 1-90 showing its passband P (*solid area*) and regions (*shaded*) needing additional blocking. Ordinate changes from linear to log at 0.1

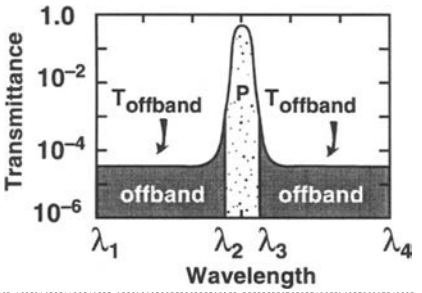


Fig. 1-95 $T(\lambda)$ (on an absorbance scale) of a hypothetical bandpass showing its passband P and offband spectral regions. The passband spectral width is exaggerated for purposes of illustration.

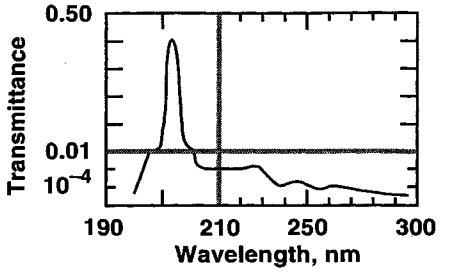


Fig. 1-96 $T(\lambda)$ of a bandpass. Its offband rejection is provided by a reflection filter. The ordinate changes from linear to log at 0.01. The abscissa changes scale at 210 nm. After Seavom [75].

which is valid because $\lambda_2 \approx \lambda_3$. A comparison of Eqs. (1-13) and (1-15) shows that the ratio of wanted flux to unwanted is

$$\text{ratio} \propto \frac{\Phi_s}{\Phi_u} \propto \frac{\Delta\lambda_{0.5}}{\lambda_4 - \lambda_1} \tag{1-16}$$

and is inversely proportional to the spectral bandwidth of the rejection region.

As an example, suppose flux is collected at 500 nm with a bandpass of 5 nm bandwidth. The rejection band extends to 1200 nm in the near IR, which is the limit of the silicon photodiode.

$$\text{ratio} \propto \frac{\Phi_s}{\Phi_u} \propto \frac{5}{1200 - 500} \propto \frac{5}{700} \propto \frac{1}{140} \; . \tag{1-17}$$

In other words, there is 140 times as much unwanted flux, as compared to signal. A conclusion is that if the unwanted flux is to be less than 1/1000 of signal, an offband transmittance of 10^{-6} is mandated.

1.6.5.2 Achievement of blocking

Blocking is achieved by several methods: (1) absorption filters; (2) multilayer dielectric stacks; (3) multilayer filters containing metallic layers; (4) reflection filters; or (5) combinations of the aforementioned methods.

As an example, consider a bandpass transmitting the 515 nm laser line used with a silicon photodiode detector. Glass absorption filters (see §1.4.2) provide an offband

transmittance of 8 absorbance units at wavelengths shorter than 400 nm. Multilayer blockers containing silver layers (see Chap. 8) attenuate at longer wavelengths.

Table 1-6 lists filters for various spectral regions and methods of blocking.

Table 1-6 Blocking filters and bandpass filters

		Blocking for the spectral region	
Wavelength range	Possible materials used	at short wavelengths	at long wavelengths
120 nm - 200 nm	Al* plus dielectrics	Substrate absorption	Self-blocking
200 nm - 400 nm	PbF_2, AlF_3, cryolite, LaF_3, MgF_2	Al* plus dielectrics or substrate absorption	Al* plus dielectrics
280 nm - 380 nm	PbF_2, AlF_3, cryolite, HfO_2, Y_2O_3, ZrO_2, SiO_2, Al_2O_3	Schott UG-11 glass or substrate absorption	Schott UG-11 glass
400 nm - 1200 nm	ZnS, cryolite, PbF_2, AlF_3, HfO_2, Y_2O_3, ZrO_2, SiO_2, Al_2O_3	Absorbing glass	Ag‡ plus dielectrics
1.2 μm - 2.5 μm	Si, SiO, ZnS, SiO_2, CdS	Dielectric stacks plus substrate absorption	Dielectric stacks
2.4 μm - 6 μm	Si, SiO, ZnS, ZnSe, CdS, Ge	Dielectric stacks plus substrate absorption	Dielectric stacks
6 μm - 12 μm	ZnS, ZnSe, CdS, CdSe, Ge	Dielectric stacks plus substrate absorption	Dielectric stacks

Construction of bandpasses and blocking filters in various parts of the spectrum.
*Aluminum ‡Silver

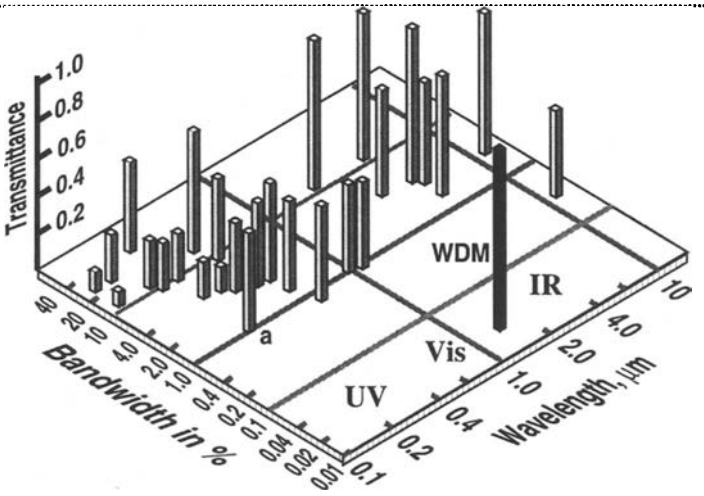


Fig. 1-97 For a single WDM bandpass and many conventional bandpasses, peak transmittance in the passband, wavelength of the passband and fractional bandwidth. Data collected from sales literature in the late 1980s. The out-of-band transmittance is below 10^{-3} in most instances.

Figure 1-96 graphs $T(\lambda)$ of a UV bandpass of 2 nm bandwidth. Its superb offband rejection is provided by a four-bounce reflection filter (see §1.4.3.1).

Figure 1-97 shows bandwidth, center wavelength of the passband and maximum passband transmittance of bandpasses produced by commercial vendors. In most filters, only aluminum layers and dielectric layers are used when the passband is below 200 nm. Only all-dielectric blockers are used in most IR filters.

1.6.6 Bandpass in convergent flux

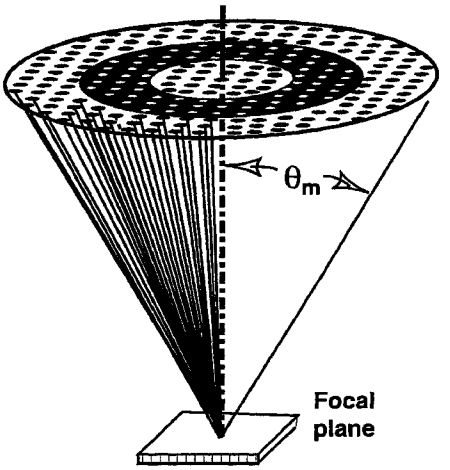


Fig. 1-98 Rays in an incident parallel bundle descend vertically and impinge upon the entrance pupil of a lens (not shown). The refracted rays converge upon the focal plane at the maximum semi-angle θ_m. The angles are exaggerated for purposes of illustration.

1.6.6.1 Introduction

Suppose a bandpass is inserted at the focus of a convergent beam, as at the focal plane in Fig. 1-98. Several questions are: (1) What is the maximum allowable angular subtense θ_m of the convergent beam? (2) At what wavelength should the passband be centered? (3) How is the transmittance in the convergent beam found? (4) Can an optical system be redesigned to reduce the effects of the *angle shift* of the bandpass?

Questions (1) and (2) are addressed in §1.6.6.1.1 and §1.6.6.2, respectively. The remaining questions involve the computation of the biconical transmittance, which is discussed in §1.2.2.7 and §10.3.7.

1.6.6.1.1 Maximum cone angle

As an example, Fig. 1-98 shows a parallel bundle of rays focused by an f/1.9 lens, which corresponds to θ_m of 15°. Given the 6.1 nm full width of the passband, is this cone angle excessive?

The foregoing question may be answered in several ways. One is to insert that bandpass on a rotary table in the sample compartment of a spectrophotometer. Transmittance at λ_0 (633 nm in this example) is determined from the scans at several angles. As an illustration, Fig. 1-99 shows scans at 0°, 8° and 11°. In the right panel is a curve of $T(\theta)$ at 633 nm. A conclusion is that when θ is 11°, T plummets to 9%. Hence, there is little motivation to extend θ to 15°.

An estimate of the maximum cone angle may also be found by equating $\Delta\lambda_a$ in Eq. (1-12) of one-half the bandwidth BW and then solving forθ. This is

$$0.5 \times \mathrm{BW} = 0.5 \times 6.1 = C_a\ \lambda_0\, \theta^2 \quad \text{or} \tag{1-18a}$$

$$58\times10^{-6} \times 633 \times \theta^2 = 3.05. \tag{1-18b}$$

Solution of Eq. (1-18b) for θ yields 9.1°. Figure 1-100 depicts biconical transmittance of this coating — the *solid* curve obtains when the bandpass is centered at 633 nm. The problem is that T at 633 nm is only 39%. The *peak* biconical transmittance of 72% at a shorter wavelength leads us to shift its wavelength center, as outlined below.

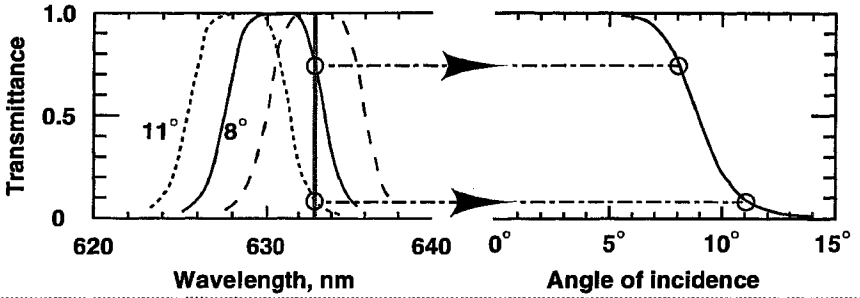


Fig. 1-99 (*Left panel*) For the bandpass cited in the caption to Fig. 1-90, $T(\lambda)$ at 0° (*long dash*), 8° (*solid*) and 11° (*short dash*). Shaded vertical line is 633 nm. (*Right panel*) $T(\theta)$ at 633 nm.

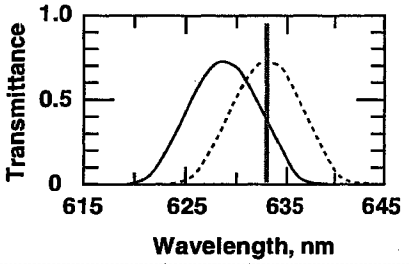


Fig. 1-100 (*Solid*) Biconical transmittance of the bandpass captioned in Fig. 1-90 for θ_m of 15°. (*Dashed*) bandpass design is the same, but centered at 637 nm. The shaded vertical line is 633 nm.

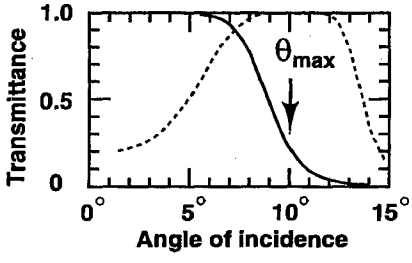


Fig. 1-101 (*Solid*) $T(\theta)$ at 633 nm of the bandpass captioned in Fig. 1-90. (*Dashed*) The bandpass design is the same, but centered at 637 nm.

1.6.6.2 Shifting of the center wavelength

An assumption is that the bandpass is to transmit at a relatively narrow wavelength (such as a laser) λ_1. Its biconical transmittance is improved by shifting the wavelength center of the bandpass, as follows:

> The wavelength λ_0 of the passband center at normal incidence should be shifted to a wavelength *longer* than λ_1 by an amount of one-half the maximum angle shift. Use Eq. (1-12) to find the maximum angle shift.

The foregoing rule is attributed to Linder [67] and Pidgeon [64]. The bandpass cited in §1.6.6.1 serves as an example. When illuminated at a 15° angle, its maximum fractional angle shift $\Delta\lambda_a/\lambda$ is 0.014 = 1.4%. Thus, if the bandpass were to collect flux in a convergent cone (with a 15° semi-angle) from a He-Ne laser at λ_1 = 633 nm, the maximum angle shift of its passband center would be 0.014 × 633 = 8.5 nm. The wavelength of that center should be shifted to longer wavelengths by one-half of that

8.5 nm. This amounts to 4.3 nm and hence $\lambda_0 = 633 + 4.3 = 637.3$ nm. The dashed curve in Fig. 1-100 obtains when the wavelength center is shifted longer by 4 nm.

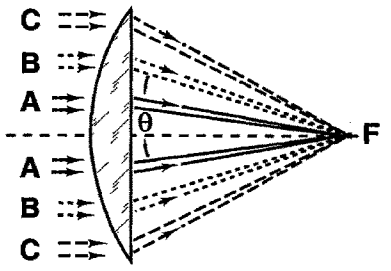


Fig. 1-102 Rays in a parallel bundle are refracted by a lens and converge at angle θ upon the focus *F*. The angles are exaggerated for purposes of illustration. *A*, *B* and *C* refer to the zones in Figs. 1-98 and 1-103.

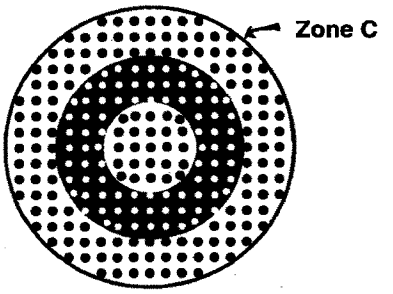


Fig. 1-103 The rays captioned in Fig. 1-102 impinge upon the entrance pupil of a lens. The angle θ at which the rays converge on the focus are:
(*Inner black dots* — Zone *A*) $0° < \theta < 5°$.
(*White circles* — Zone *B*) $5° < \theta < 10°$.
(*Black circles* — Zone *C*) $10° < \theta < 15°$.

1.6.6.2.1 Why this rule is successful

Figure 1-101 depicts $T(\theta)$ (at 633 nm) of two bandpasses. The first (*solid curve*) has its passband centered at 633 nm at normal incidence. The other (*dashed*) follows the prescription of §1.6.6.2 — its wavelength center is at 637 nm, which is 4 nm longer then 633. Figure 1-101 shows that at normal incidence its T at 633 nm is a rather low value of 20%. So why is there any improvement? The reason is that there are *few* rays that impinge on the focus at this low angle of incidence — but many, many more at relatively high angles.

Figure 1-102 depicts a parallel bundle of meridional rays that refract and converge on the focus at an angle θ. Bundles *A*, *B* and *C* correspond to $\theta = 5°$, $\theta = 10°$ and $\theta = 15°$, respectively. Thus the circular entrance pupil is divided into three zones. As shown in Fig. 1-103 the inner zone *A* includes those rays converging on the focus at angles from zero to 5°. The intermediate zone *B* includes rays whose convergence angles are between 5° and 10° and the outer zone *C*, angles of 10° to 15°.

The number of rays emerging from each aforementioned zone is shown in Fig. 1-103. There are 25 in the inner, 84 in the intermediate and 132 in the outer zone. Figure 1-104 shows that these ray counts correspond to 10%, 34% and 54% of the total, respectively. This makes it clear that most of the rays converge upon the bandpass, not at normal incidence, but at higher angles. In conclusion, there are more than five times as many rays in the outer zone as compared to the inner zone. Inspection of Fig. 1-101 shows that this deliberate shifting of the wavelength center of the passband (at normal incidence) to longer wavelengths increases T at higher angles, where there are relatively more rays.

1.6.6.3 Étendue (throughput)

Étendue, also called *throughput*, is a measure of the ability of a bandpass of area $\mathcal{A}$ to gather radiant flux. Étendue is proportional to $\mathcal{A} \sin^2\theta_m$, where θ_m is its maximum acceptable angle (i.e., semi-angle) of the cone of convergent flux transmitted by the bandpass, as discussed in §1.6.6.1.

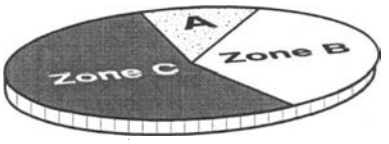


Fig. 1-104 For the rays cited in the caption to Fig. 1-103, fraction of the total in Zone *A*, Zone *B* and Zone *C*.

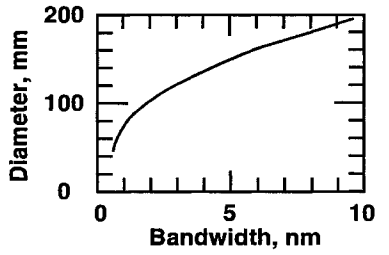


Fig. 1-105 For bandpasses in the visible spectral region, smallest attainable bandwidth vs maximum diameter of bandpass. After Spectrofilm [88].

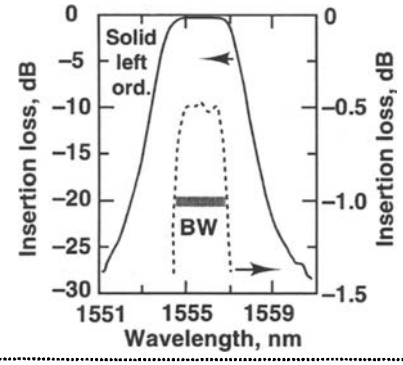


Fig. 1-106 Measured $T(\lambda)$ (on a decibel scale) of a bandpass for WDM. *Dashed* curve for the right ordinate. After Barr [97].

Thickness uniformity of the layers in the bandpass is quite important because this determines the usable area. As an illustration, suppose a bandpass has a fractional bandwidth of 1%. A question to ask is: Over what area of the bandpass are layers uniform in thickness to a reasonable fraction of that 1% — say, 0.2%? If the answer is "a region 400 mm^2 " then this area of 400 mm^2 is the useful area $\mathcal{A}$, no matter what the physical dimensions of the filter may be. As an example, Fig. 1-105 shows the diameter over which a bandpass of a given bandwidth may be produced. The diameter of a bandpass of 1.0 nm bandwidth is less than 65 mm. Small areas of a bandpass have been scanned spectrally and the results presented either as a contour map [Bickford 66] or arrays of scans [Elliott 77].

1.6.7 Attributes of a bandpass for wavelength multiplexing or wavelength demultiplexing

A warning to the reader is that the specifications (written in the year 2001) are continually changing. Information stated here today may be completely altered a few years hence. The topic of wavelength division multiplexing WDM is introduced in §1.6.3.2. Figure 1-83 depicts the spectral transmittance on a linear scale and Fig. 1-106 depicts $T(\lambda)$ on an decibel scale of filters for a WDM system. Typical filter specifications are:

(a) **Passband transmittance** Its maximum transmittance (T_{max}) cannot fall below D = 0.03 absorbance units (–0.3 dB), which corresponds to T = 93.3%.

(b) **Bandwidth** The bandwidth is its full spectral width of the passband at $D = 0.05$ absorbance units (i.e., –0.5 dB) below T_{max}. The measured bandwidth of the bandpass in Fig. 1-106 is 2 nm and the computed bandwidth of the filter in Fig. 1-85 is 0.55 nm.

(c) **Offband rejection**. The spectral width of the passband at an absorbance level of 2.5 (i.e., –25 dB) is important, because this determines the "cross talk" between adjacent communication channels. For the filter in Fig. 1-106 this is 8 nm and for the filters in Fig. 1-85 it is 1.1 nm. In other instances, a lower transmittance of 3 absorbance units (i.e., –30 dB) may be required in lieu of the –25 dB cited earlier.

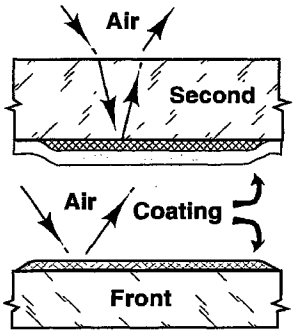


Fig. 1-107 Reflecting coating (*cross-hatched*) forms either second-surface or front-surface mirror.

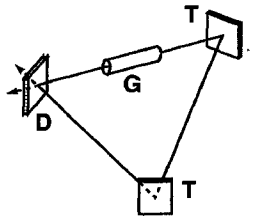


Fig. 1-108 Ring laser consisting of gain tube G, totally reflecting mirrors T and beamdivider D (output mirror). After Walker [74].

(d) **Shape factor, merit function** It is desirable that these WDM filters manifest a small shape factor [defined in §1.6.4(f)], because this allows the communication channels to be packed closely together. The reciprocal of the *shape factor* is often termed the *merit function*.

(e) **Blocking**. Low transmittance, i.e., blocking, (see §1.6.5) at wavelengths over a wide spectral region is not required.

(f) **Temperature shift**. Due to the narrowness of the spectral width of the passband, a quite small temperature shift of 2.3×10^{-6}/°C is required for a WDM bandpass. At λ_0 of 1550 nm, this translates into 1.4 pm/°C. This is 30 times smaller than the temperature shift of 5×10^{-5}/°C (for a conventional bandpass) cited in the final paragraph of §1.6.4.2.

1.7 Reflectors — used principally at normal incidence

1.7.1 Introduction

The *normal incidence* modifier alludes to an arbitrary decision to relegate the discussion of reflectors/transmitters that function at nonnormal incidence to §1.8, which is devoted to beamdividers and dichroics.

1.7.1.1 Front-surface reflector or second-surface reflector

There are several ways to classify reflectors. First, a mirror can either be a front-surface or second-surface reflector, depending upon whether the front side or reverse side of the substrate is coated (see Fig. 1-107). The second-surface mirror is used in nonimaging systems, where optical quality of the slab is of no concern.

1.7.1.2 Total and partial reflectors

This terminology is borrowed from laser technology. All lasers contain a cavity consisting of a gain medium surrounded by two or more mirrors. In most cases, at least one of the mirrors is totally reflecting — this is termed a *high reflector* or *total reflector* in the trade. One of those mirrors is a partial reflector that permits flux internal to the laser cavity to leak out, as shown in Fig. 1-108. This is termed an *output coupler,* an *output reflector* or *partial reflector.*

As another example, suppose a multilayer reflector with R = 99% is used as a turning mirror in a high irradiance laser system. Only its R is important; no effort is made to optimize T. This coating may leak a modicum — say, 0.7% — of the flux. The transmitted flux is not wasted, however. It is collected by a detector that monitors the power level of the laser. Although this 0.7% leakage might be acceptable for a multilayer used outside the laser cavity, the coating would be unusable if it were used as the high reflector of a laser cavity.

1.7.1.3 Substrates for reflectors

Table 1-7 lists substrates upon which reflectors are deposited. The coating must also be essentially opaque when a substrate is opaque and the irradiance level is high.

1.7.1.4 Types of coatings

§1.7.3 surveys types of coatings ranging from a single 50 nm thick layer of metal to a dielectric multilayer that is 60 times as thick. Figure 1-109 delineates thickness of coatings, which is a hint at their complexity.

1.7.2 Criteria for the selection of reflectors

Table 1-8 lists criteria for the selection of a reflector. Cost is important for mass-produced mirrors. Spectral bandwidth is discussed in §1.7.2.1. Reflectors in a laser gyro cavity demand the highest R of 99.99%. Often mirrors must survive mechanical abrasion and immersion in a humid atmosphere. The phase of a reflected wave front may be distorted if the reflector is not deposited with uniform thickness. Differential phase shift upon reflection Δ is of concern when the polarization state of the reflected beam is controlled.

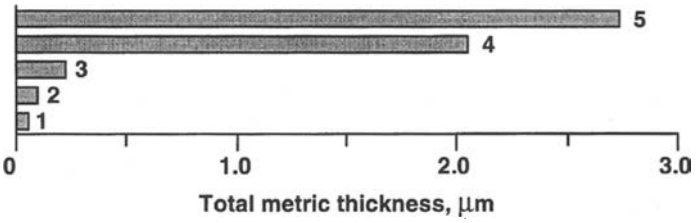


Fig. 1-109 In the visible spectrum, metric thicknesses of reflectors: (1) Opaque metal; (2) opaque metal with one-layer overcoat; (3) opaque metal with enhanced reflectance; (4) quarterwave stack all-dielectric reflector; (5) broadband all-dielectric reflector.

Table 1-7 Substrates for reflectors

Type of substrate	Comment
Stretched plastic	Low physical mass — used in aircraft*
Soda-lime glass	Possibly the least expensive
Zerodur™	Lowest thermal expansion coefficient
Nickel	Electroformed nickel is used for reflectors
Molybdenum	Used for high power lasers
Copper	Used for high power lasers
Silicon	High thermal conductivity — may be polished quite smoothly.
Fused silica	May be polished quite smoothly.

Substrates for mirrors. *See Jeffree [65] and Benzies [71].

1.7.2.1 Bandwidth and level of reflectance

Consider a front surface coating on a mirror of an astronomical objective — i.e., a "telescope" mirror. It should reflect from 300 nm in the UV to 30 μm in the IR. A single layer of aluminum is a reasonable compromise — its reflectance hovers near 90% over much of that bandwidth.

Figure 1-110 shows that the UV reflectance of all-dielectric multilayers is well above that of aluminum cited above — but only over a much narrower bandwidth. Figure 1-111 depicts $R(\lambda)$ of a broadband reflector. This all-dielectric coating contains 15 layers.

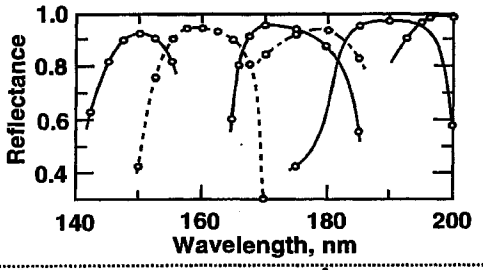


Fig. 1-110 Measured $R(\lambda)$ of all-dielectric UV reflectors. After Marling [77].

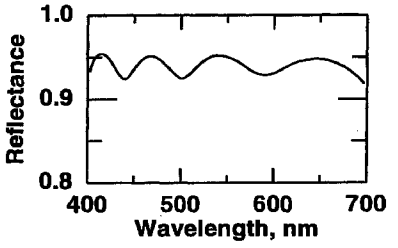


Fig. 1-111 Computed $R(\lambda)$ of a 15 layer all-dielectric reflector. Its design appears in §1.10.2.1.

1.7.2.2 Effects of thickness nonuniformity

Effects of thickness nonuniformity are illustrated by the comparison of two coatings: Each reflects in the spectral region of 400 nm to 700 nm. The coatings are: (1) A single opaque layer of aluminum of 50 nm thickness; and (2) a dielectric multilayer (see Fig. 1-111). The mean R of aluminum is approximately 89% in the visible, whereas R of the multilayer hovers about 94%. In §10.7.3.3 and Fig. 10-63, it is evident that a simultaneous altering of the thickness of each layer by a mere 1% produces a distortion in the reflected wavefront of 0.04 waves at 440 nm. For the aluminum layer, substitution of a 1% thickness change into Eq. (10-20a) shows that the wavefront distortion produced by

a thickness change is 0.002 waves, which is seventeen times smaller than that of the dielectric multilayer. One of the reasons that a thickness error in the multilayer dielectric produces a greater wavefront distortion (as compared to the aluminum) is that its metric thickness is 1100 nm — 22 times that of the aluminum. This is discussed in more detail in §10.7.3.

Table 1-8 Attributes of reflectors

Parameter	Comment
Low cost	Use as few layers as possible.
Large spectral bandwidth	A metallic layer has widest BW.
Low mechanical stress	Single layer
Laser damage threshold	All-dielectric stack
Controlled Δ^*	Dielectric stack or metal + two layers

Parameter	Comment
Best environmental stability	Rhodium layer or all-dielectric multilayer
Low scatter	Single metallic layer or all-dielectric multilayer
Highest reflectance	All-dielectric stack
Minimal phase distortion	Use as few layers as possible

Attributes of reflectors. *Differential phase shift upon reflection Δ.

1.7.2.3 Scatter and other losses

The *loss* of a laser mirror is that component of the flux absorbed or scattered outside of the mode (see §1.2.2.3). Absorbed flux heats a laser mirror and may thermally distort it. A defect function (i.e., an inverse of a merit function) for an output mirror of a laser is its *loss over T ratio*. A perfect output mirror has a *loss over T ratio* of zero. A mirror could be fabricated with an adequately high reflectance, but it could be unsuitable as an output mirror because its *loss/ T* is excessive.

1.7.2.4 Mechanical stress

Mechanical stress of a coating can warp a substrate, thereby distorting the wavefront. In one coating house, small mirrors were deposited on relatively thin plano substrates used in a laser scanning system. Coating warped the substrate — the mirror was convex. The remedy was to polish the substrates slightly concave. They warped to a plano optical figure after coating.

1.7.3 Overview of reflectors

Table 1-9 lists commonly used reflectors, which range in complexity from a single layer to multilayers with 50 films.

Table 1-9 Construction and usage of reflectors

Coating	Wavelength range	Layer count	Range of reflectance	Disadvantages	Advantages
Aluminum	120 nm to 99 μm	1	87% to 98%	Mechanically soft	Easy to apply, inexpensive.
*Aluminum + MgF_2	120 nm to 99 μm	2	87% to 98%	—	Best UV reflector.
*Aluminum + SiO	450 nm to 7 μm	2	87% to 98%	Slightly absorbing at $\lambda < 500$ nm	Mechanically durable coating.
*Aluminum + multilayer	450 nm to 600 nm	2 to 7	87% to 98%	Relatively narrow bandwidth	Easy to mass produce.
‡Silver + multilayer¶	400 nm to 3 μm	5 to 7	87% to 98%	Often degrades in humid atmosphere	Highest R in visible and near IR
‡Silver + multilayer§	10 μm to 11 μm	5 to 7	$R > 99.7\%$ at λ_0	Wavelengths range > 400 nm.	Survives in humid atmosphere
Gold	1 μm to 99 μm	1	$R > 99.8\%$ in IR	Mechanically soft	Can be electroplated
All-dielectric QW stack	$\lambda_0 \pm 10\%$	9 to 25	$R > 99.99\%$ at λ_0	Relatively narrow bandwidth	High laser damage threshold
All-dielectric broadband	2:1 bandwidth in visible or near IR	31 to 50	$R > 99\%$	Mechanical stress, non-uniform thickness	Mechanically durable coating.

Reflectors containing metal layers, metal and dielectric layers and all-dielectric coatings.
¶ Multilayer optimized for visible part of spectrum.
§ Multilayer optimized to reflect at $\lambda_0 = 10.6$ μm.
QW is an abbreviation for quarterwave.
*Aluminum with overcoat.
‡Silver with overcoat.

1.7.4 Metallic reflectors

Although most coatings of this section are metallic, there are oxide semiconductors (such as lanthanum hexaboride or indium tin oxide) that have a reflectance of at least 80% in the IR.

Advantages of a mirror consisting of a single metallic layer are threefold: First, it is relatively inexpensive to prepare. A 50 nm thick layer is virtually opaque in all wavelengths, save the far UV. Another advantage is that the spectral region over which metals reflect is broad, indeed. Aluminum reflects well from 120 nm to 100 μm. §1.7.2.2 mentions that thickness nonuniformity is less of a problem because of the thinness of the layer.

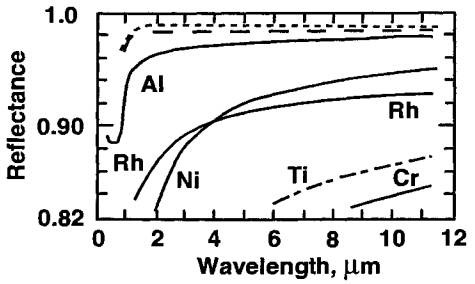


Fig. 1-112 At an angle of incidence of 20°, $R(\lambda)$ of opaque metallic films. (*Short dash*) silver. (*Long dash*) gold. After Gates [55].

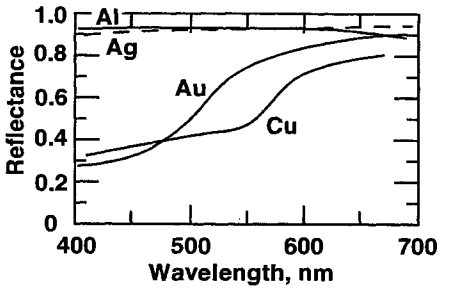


Fig. 1-113 $R(\lambda)$ of opaque layers of aluminum, silver, gold and copper. After Colbert [46].

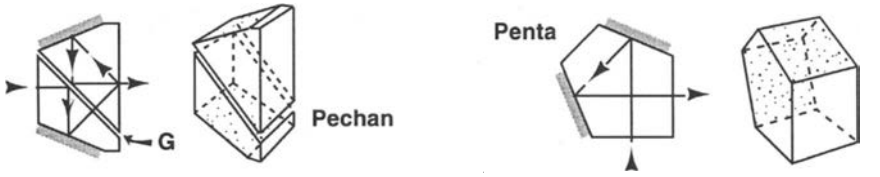


Fig. 1-114 Shaded (or dotted) surfaces of prisms are coated with reflectors. G is an air gap. After Hopkins [62, 65].

Although the R of rhodium is not as spectacular as the aforementioned materials, it has the advantage of being a hard, durable, noble metal. It resists chemical attack, as from salt spray. Nickel and chromium are also hard and durable. Most rear view mirrors on automobiles used for night driving are chrome coated. Hass [65] reviews techniques of producing these mirrors.

Figure 1-112 shows $R(\lambda)$ in the IR of commonly used metals. Silver has the highest IR reflectance and gold is a close second. Gold is often used in preference to silver because silver is attacked by contaminants such as water vapor in conjunction with sulfur and chlorine. Both metals are relatively soft and mechanically fragile. Figure 1-113 shows $R(\lambda)$ in the visible spectrum, where gold and copper are no longer strongly reflecting. R of silver drops precipitously to low values at 390 nm.

In contrast to noble metals, aluminum has a thin oxide passivation layer that protects its surface from further oxidation and makes it somewhat more durable. Its substantial R of 98% at $\lambda = 10$ μm decreases steadily until it reaches a minimum of 88% at 800 nm. R increases in the visible and UV until it peaks again at $\lambda = 200$ nm. Most telescope mirrors are "front surface aluminum" because of its relatively high R in the visible and near UV.

1.7.4.1 Second-surface coatings

Metallic coatings are applied to exterior surfaces of prisms for two reasons: (1) Enhancement of R when there is no total internal reflection; and (2) control of the differential phase shift upon reflection.

1.7.4.1.1 Coatings on prisms

Angle of incidence on faces of Pechan and penta prisms shown in Fig. 1-114 does not exceed the critical angle. Hence, those faces must be coated with either silver or

aluminum. Figure 1-115 shows that internal reflectance (with glass as the incident medium) is lower than when reflection takes place in air. Although the calculation is made for opaque aluminum, the foregoing sentence is true for all metals.

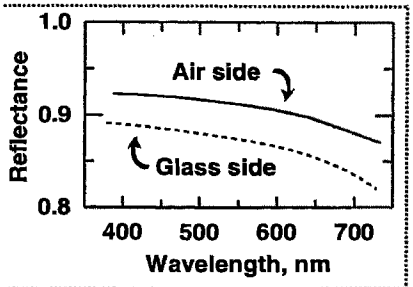


Fig. 1-115 At θ_0 of 45°, computed $R_a(\lambda)$ of opaque aluminum. (*Solid*) Incident medium is air. (*Dashed*) Incident medium is glass (index 1.52).

1.7.4.1.2 Control of differential phase

It is surprising that in certain high quality imaging systems it is necessary to metalize a prism surface even when total reflection occurs. This is requisite when a beam reflects from different halves of a prism, as illustrated by an Amici (alias roof) prism, penta-roof and Abbe, as shown in Fig. 1-116. If the surfaces were not coated, R is 100%, but there is a different phase shift upon reflection for the light that impinges on the two halves of the roof, designated as A and B in Fig. 1-116. This phase difference introduces *phase apodization* into the beam, which in turn broadens its point spread function because of diffraction [Mahan 50]. Metal coatings reduce the phase difference and thus improve image quality.

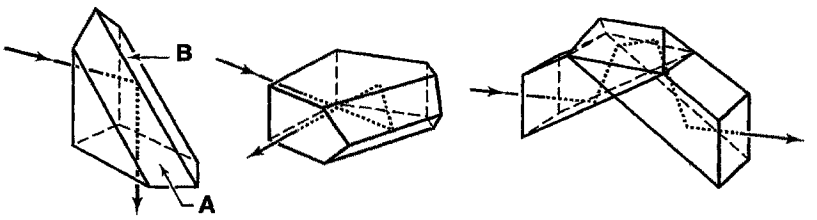


Fig. 1-116 Ray paths in an Amici prism (*left*), penta-roof (*center*) and Abbe (*right*). Surfaces A and B may be aluminized to reduce the differential phase shift

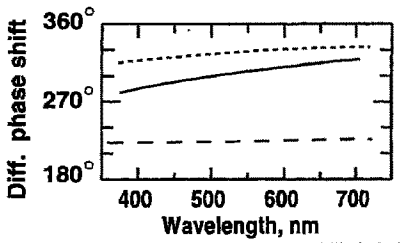


Fig. 1-117 At θ_0 of 54.7° in BK-7, differential phase shift upon reflection of aluminum (*short dash*), silver (*solid*) and bare glass (*long dash*).

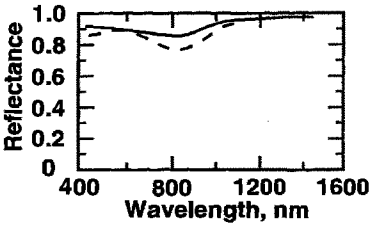


Fig. 1-118 $R(\lambda)$ of opaque aluminum (*solid*) and aluminum overcoated with a silicon monoxide layer (*dashed*).

Figure 1-117 shows that a metal-coated prism surface produces a differential phase shift closer to 360° (i.e., zero) than bare glass. Silver has a larger differential phase shift than aluminum. Thus the latter is often used to minimize diffraction, although silver would produce a higher reflectance. Similar diffraction effects occur in cube corner reflectors (see §1.7.6.4).

1.7.5 Overcoated metals

One or more layers are frequently added to a metallic mirror to: (a) enhance its reflectance; and/or (b) increase its mechanical robustness; and/or (c) protect it from the environment.

1.7.5.1 Single layers

Figure 1-118 shows $R(\lambda)$ (*dotted*) of an aluminum mirror with a single-layer silicon monoxide (SiO) overcoat. It protects the metal, but only at the expense of lower reflectance over most of the spectrum. Only in the IR, where the absorption of the SiO layer is reduced and its thickness is small compared to a wavelength, does its reflectance rival that of uncoated aluminum. Nevertheless, a SiO protective layer enables a mirror to pass the salt spray environmental test, which naked aluminum cannot.

The decrease of reflectance in the UV due to absorption of SiO can be avoided by replacing it with an evaporated layer of silica [Cox 64a]. Single layers of lithium fluoride and magnesium fluoride [Hunter 71; Hutcheson 72] are used to both protect and enhance the R of aluminum in the far UV, as shown in Fig. 1-119. The dielectric overcoating layer that overcoats a metal must be free of pinholes and have low porosity. A layer of aluminum oxide manifests a relatively high physical density and thus is quite effective as an overcoat for silver [Hass 75].

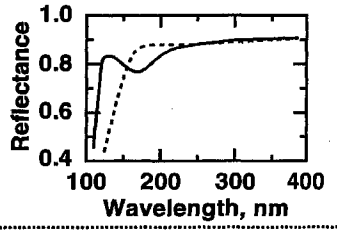


Fig. 1-119 $R(\lambda)$ of opaque aluminum overcoated with magnesium fluoride that enhances the reflectance at 120 nm (*solid*) and at 200 nm (*dashed*). After Acton [90].

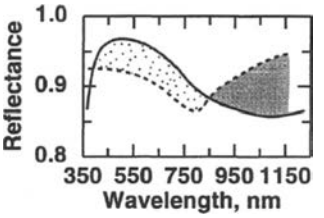


Fig. 1-120 Computed $R(\lambda)$ of bare aluminum (*dashed*) and (*solid*)

air **H** 0.8**L** *aluminum*

where optical thickness of **H** and **L** is $\lambda_0/4$ at λ_0 of 550 nm. Refractive indices of **H** and **L** are 2.30 and 1.45, respectively.

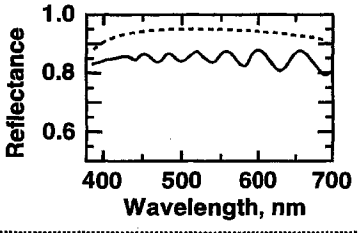


Fig. 1-121 (*Solid*) $R(\lambda)$ of rolled aluminum sheet. (*Dashed*) $R(\lambda)$ of evaporated opaque aluminum overcoated with two dielectric layers. After Dielectric Coating [91].

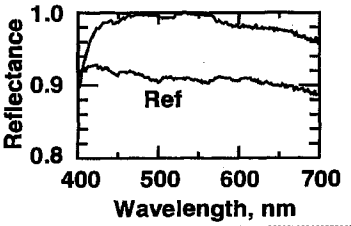


Fig. 1-122 (*Solid*) $R(\lambda)$ of "enhanced silver" and a reference reflectance. After Broomer [83].

1.7.5.2 Reflectance enhancement

Dielectric layers on top of an opaque metal can enhance its reflectance — usually over a relatively narrow spectral region. Over a broad spectral region, an enhanced coating produces a *net loss* in reflectance, as compared to R of a bare metal. The dotted area in Fig. 1-120 is the gain in reflectance for a two-layer overcoat. The shaded area, which extends farther into the IR, is the loss.

The dashed curve in Fig. 1-121 is R of opaque aluminum overcoated with two dielectric layers. The solid curve is $R(\lambda)$ of rolled aluminum sheet that has a thick oxide layer on its surface. This coating is installed in a lighting fixture to enhance its reflectance and, hence, its light output.

Figures 1-122 and 1-123 depict $R(\lambda)$ of "enhanced silver" reflectors. R exceeds 98% in the visible and 99% in the IR. Figure 1-124 plots $R(\lambda)$ of several coatings with enhanced UV reflectance. Office copiers with variable magnification also use many of these enhanced mirrors.

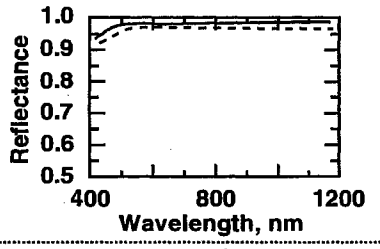


Fig. 1-123 (*Solid*) $R(\lambda)$ of "front surface enhanced silver." (*Dashed*) $R(\lambda)$ of "second-surface silver." After Denton [90].

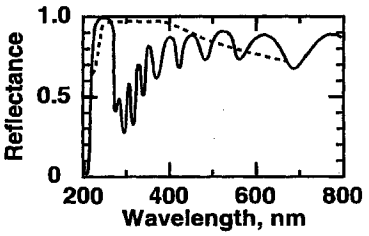


Fig. 1-124 $R(\lambda)$ of "enhanced UV reflectors." (*Dashed*) — after Acton [90]. (*Solid*) — after Nitto [98].

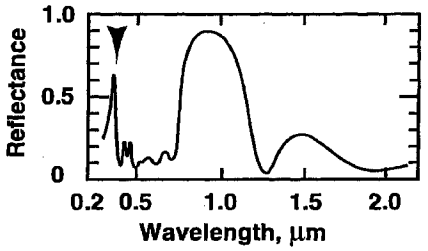


Fig. 1-125 $R(\lambda)$ of seven alternating layers (of equal optical thickness) of $PbMoO_4$ and cryolite. After Edwards [69].

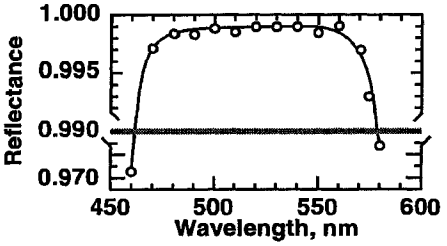


Fig. 1-126 Measured $R(\lambda)$ (*dots*) of 27 alternate layers of SiO_2 and Ta_2O_5. The ordinate changes scale at 0.99. After Arnon [78b].

1.7.6 All-dielectric reflectors

All-dielectric coatings are used for several reasons: (1) An *output mirror* is required (see §1.7.1.2); (2) its required R is higher than a metal coating provides; (3) coating must have high laser damage threshold and/or low absorption; (4) phase shift upon reflection is to be controlled. As mentioned in §1.7.2.1, consequences of substituting an all-dielectric coating for a metallic layer are greater cost and reduced spectral bandwidth of high reflectance.

The coating captioned in Fig. 1-125 contains seven alternating high index and low index layers and illustrates the limited reflection bandwidth of an all-dielectric multilayer [Edwards 69]. A R in excess of 85% is attained in a spectral band 800 nm to 1100 nm. Its reflectance could be increased to over 95% and its bandwidth increased only by increasing the number of layers. An arrow highlights a reflection peak at 300 nm, which is a third harmonic of the peak centered at 900 nm. Unless considerable effort is made to eliminate them, such harmonics are produced by all multilayer reflectors.

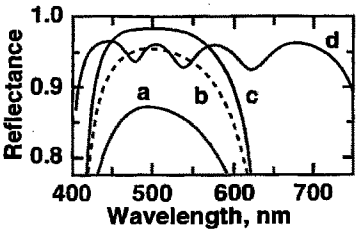


Fig. 1-127 $R(\lambda)$ of mirrors of alternate layers of ZnS and cryolite: (*a*) five layers; (*b*) seven layers; (*c*) nine layers; (*d*) 15 layer broadband reflector. After Stone [56].

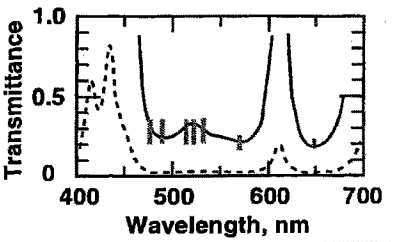


Fig. 1-128 Measured $T(\lambda)$ of an output coupler for a large frame krypton-argon ion laser. Scale 0 to 100% (*dashed*) and 0 to 10% (*solid*). After Coherent [92].

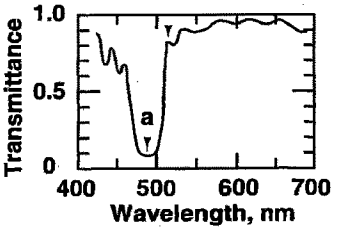


Fig. 1-129 For an argon ion laser, measured $T(\lambda)$ of an output coupler permitting only the 488 nm line (*arrowed, a*) to lase. After Microcoatings [90].

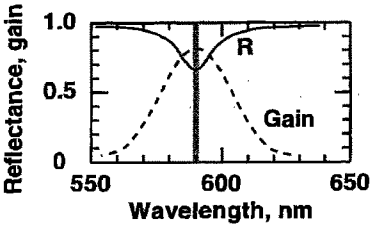


Fig. 1-130 $R(\lambda)$ of an output coupler for a dye laser (*solid*). (*Dashed*) relative gain of Rhodamine 6G dye.

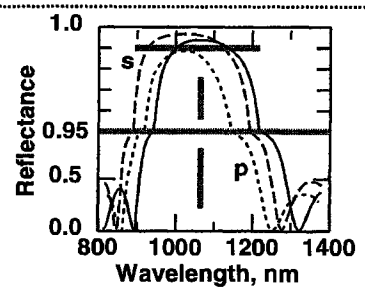


Fig. 1-131 Computed $R(\lambda)$ of an all-dielectric reflector at 1064 nm at normal incidence (*solid*), at 30° incidence angle, *p* polarization (*short dash*) and *s* polarization (*long* dash). Air is the incident medium. The design appears in §1.10.2.8.

As another example, Fig. 1-126 depicts $R(\lambda)$ of a mirror with a peak reflectance of 99.9% at λ = 520 nm [Arnon 78a]. Addition of more layers can raise R to 99.96% [Arnon 78b]. The limited bandwidth of the region of relatively high reflectance extends from 470 nm to 570 nm. This coating contains alternating high index layers of tantala and low index layers of silica. Because each layer has an optical thickness of a quarterwave, this is termed a *quarterwave stack* in the trade.

The bandwidth of an all-dielectric coating is extended by piling on the same substrate two or more quarterwave stacks — each "tuned" to reflect at a different wavelength. An illustration is Fig. 1-127, which depicts one minus transmittance "1 – $T(\lambda)$" of a broadband reflector used in a Fabry-Perot interferometer [Stone 56]. $R(\lambda)$ of quarterwave stacks is shown for comparison.

1.7.6.1 Laser reflectors

As pointed out in §1.7.1.2, these are usually classed as *total reflectors* (with as high a reflectance as possible) or *output couplers* that permit a prescribed amount of flux to leak out of a laser cavity. For a given level of gain in a laser medium, T should be within prescribed limits in order to extract the optimum amount of laser power. The vertical bars in Fig. 1-128 show, for example, upper and low transmittance limits of an output mirror for a "white light" krypton-argon ion laser in which the shortest wavelength is 476 nm and the longest is 647 nm. Figure 1-129 depicts $T(\lambda)$ of an output coupler that permits the 488 nm line in an argon ion laser to lase. Its R is low enough at 515 nm so that this relatively high gain line does not lase.

Table 1-10 Optical properties of a reflector

Incidence	Peak R_p	Peak R_s	λ_{peak}	R_p at λ_0	R_s at λ_0
Normal	0.9938	0.9938	1064 nm	0.9938	0.9938
30°	0.9888	0.9967	1020 nm	0.9862	0.9959

Computed spectral properties of the stack captioned in Fig. 1-131 and §1.10.2.8. λ_0 is 1064 nm and λ_{peak} is the wavelength at which R is a maximum.

Figure 1-130 shows the gain of a dye (curve G) and $R(\lambda)$ of a reflector tailored to couple the proper amount of power from a laser cavity at each wavelength. When gain is relatively high, its R is lower. R is higher at the "wings" of the gain curve because less power is permitted to leak from the cavity. Costich [87] reviews laser coatings.

1.7.6.2 Reflectors at nonnormal incidence

Table 1-10 summarizes the spectral reflectance of an all-dielectric multilayer reflector that was designed to function at normal incidence. Figure 1-131 graphs these data. What happens when this stack is tilted? Some general observations are:

- There is an *angle shift* (see §1.2.4) of the $R(\lambda)$ curves — the peak reflectances of both the *p* and *s* polarized curves shift 44 nm to shorter wavelengths.
- As compared to normal incidence, the peak reflectance of the *s* polarization increases.
- As compared to normal incidence, the peak reflectance of the *p* polarization decreases.
- The spectral width of the region of high reflectance grows larger for *s* polarization.
- The spectral width of the region of high reflectance shrinks for *p* polarization.

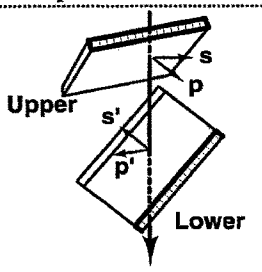


Fig. 1-132 A transmitted beam (heavy vertical line) penetrates two identical tilted coatings, such that, respectively, *p* and *s* polarizations at the upper plate translate into *s'* and *p'* at the lower plate.

The objective of the YAG reflector coating in Fig. 1-131 is to reflect in excess of 99% at 1064 nm. The thirteen layer stack of alternating "high-low" layers accomplishes this. But when the light is incident at 30°, the *p* reflectance drops below the specified 99%. In conclusion, if a stack were required to reflect randomly polarized light incident at 30°, the 13 layer stack must grow to 15 layers. The *s* reflectance would then increase well above the required 99%. There is little that can be done to mitigate this "*s*-reflectance overkill," however, since the additional layers are added to enhance the *p* reflectance.

1.7.6.2.1 Polarization-independent transmissive system

The problem is to monitor the power of a pulsed laser. A laser beam hits a turning mirror at a 45° angle of incidence and the transmitted beam (the transmission level is only a fraction of a percent) is collected by a detector. The problem is that, according to §1.7.6.2, the transmitted flux depends upon the linear polarization state of the laser beam, which may vary from pulse to pulse.

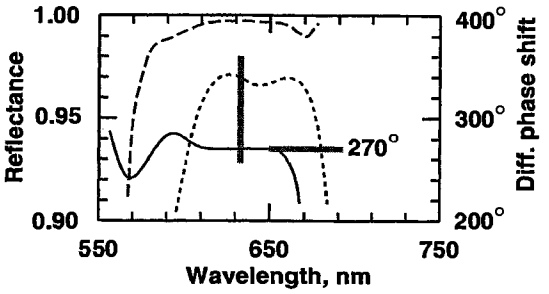


Fig. 1-133 At an angle of 45° in air, $R_p(\lambda)$ (*short dash, left ordinate*), $R_s(\lambda)$ (*long dash, left ordinate*) and differential phase shift upon reflection (*solid, right ordinate*) of a dielectric multilayer. Its design is in §1.10.2.6.

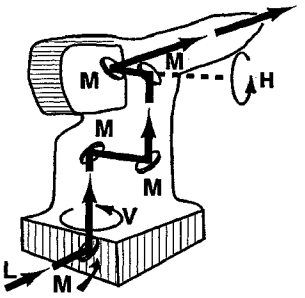


Fig. 1-134 "Gun" with horizontal axis of rotation *H* and vertical axes of rotation delivers laser beam *L* to target. *M* is a turning mirror.

The transmittance of the coated slabs in Fig. 1-132 is insensitive to the polarization state of the laser. In addition to the *upper* turning mirror, an identical turning mirror (i.e., *lower*) is also at an angle of incidence of 45°, but rotated 90° about an azimuthal angle relative to the *upper* mirror. The net effect is that the *p* and *s* polarizations, which are defined with reference to the plane of incidence upon the slab, are interchanged. That is, the *p* polarization transmitted by the upper mirror becomes *s* polarization at the lower mirror. Suppose that the mirrors increase in temperature, which causes the spectral curves to shift. Because the mirrors are from the same coating deposition, the shifts are identical and the overall transmittance for randomly polarized flux is independent of temperature.

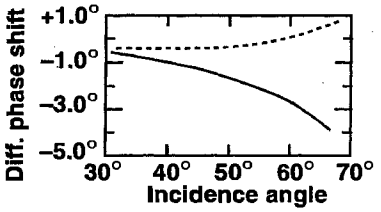


Fig. 1-135 At a wavelength of 10.6 μm, differential phase shift upon reflection vs θ_0 (measured in air) of bare silver (*solid*) and silver with a halfwave optical thickness layer of germanium (*dashed*).

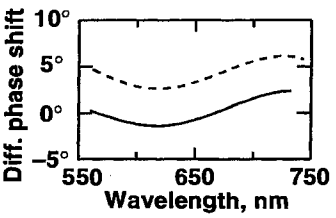


Fig. 1-136 At θ_0 in glass of 54.7° (*solid*) and 55.7° (*dashed*), differential phase shift upon reflection vs wavelength of a five-layer coating whose design appears in §1.10.2.7.

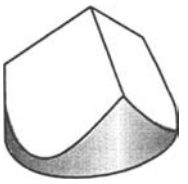


Fig. 1-137
Cube corner, nonshaded portions of which are coated with the multilayer captioned in Fig. 1-136.

1.7.6.3 Reflectors with phase control

Two examples illustrate that in addition to specifying radiant reflectance, control of differential phase shift upon reflection Δ is possible. The all-dielectric reflector whose Δ and $R(\lambda)$ appear in Fig. 1-133 converts linearly polarized light from a He-Ne laser into a circularly polarized beam. Its R is nearly the same for both linear polarizations and there is a Δ of 270°.

A "zero differential phase" coating is applied to turning mirrors of a system that delivers a multi-kilowatt circularly polarized CO_2 laser beam to a sheet of metal that is to be welded. Delivery systems are much more complex than the simple two-axis system appearing in Fig. 1-134 that contains five turning mirrors. Each mirror preserves the state of polarization of the beam that reflects from it. In other words, if circularly polarized radiant flux is incident upon a mirror, then a circularly polarized beam should emerge.

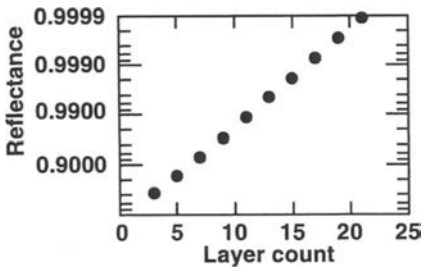


Fig. 1-138 Versus number of layers, maximum R of a quarterwave stack containing layers of silica and titania.

The coating consists of opaque silver overcoated with a halfwave of germanium. Each mirror has a reasonably high R and Fig. 1-135 shows that Δ changes by less than 1° as the angle of incidence ranges from 30° to 65°. In contrast, Δ of bare silver fluctuates by 3.5° over the same span of angles.

1.7.6.4 TIR — phase control

Total internal reflection TIR occurs when the angle of incidence exceeds the critical angle. R is 100% and addition of any thin layers (that may absorb and scatter) can only reduce this R. The intent is to reduce the differential phase shift upon reflection Δ from 135° (for an uncoated prism face) to zero. At 633 nm Mauer [66] accomplished this with a three-layer stack. The coating whose Δ vs wavelength appears in Fig. 1-136 was designed by optimization. It is applied to three faces of the cube corner shown in Fig. 1-137. Figure 1-136 illustrates that Δ increases by 4° when the angle of incidence (in the glass) grows by 1°.

1.7.6.5 Specification of the level of reflectance

There are several reasons why an optical systems engineer should resist any temptation to specify R as 99.99%, when 99% would be adequate: (a) As compared to 99%, a 99.99% coating requires additional layers, which adds to its cost; (b) A 99.99% R is difficult to achieve because of residual absorption and scatter losses in the coating. In fact, R of 99.99% is needed (and is achieved on a routine basis) for mirrors used in a ring laser gyro. Gyro mirrors are by no means inexpensive. Figure 1-138 shows R of a quarterwave stack of silica and titania vs number of layers. Eleven layers produce a R of 99%. Six additional layers must be deposited to achieve 99.9% R.

1.7.6.5.1 Reflector in an immersed system

§1.8.1.1 discusses the meaning of the term "immersed." An example of an immersed coating appears in Fig. 1-157. Suppose that it is to have a high reflectance (such as 99%) for both the *p* and *s* polarizations. At certain angles of incidence, an immersed (i.e., a coating within a cemented "cube") multilayer dielectric reflector requires an immense number of layers to achieve the 99% *R* cited above.

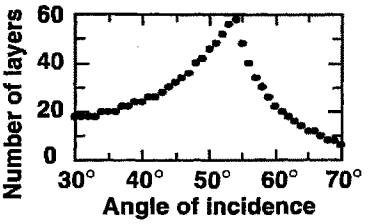


Fig. 1-139 For quarterwave layers of alternate high and low refractive index, layer count needed to produce R_p in excess of 99%. Coating is in cemented cube of index 1.52. θ_0 is measured in that medium.

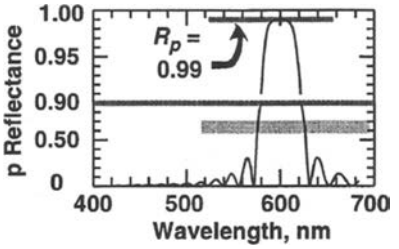


Fig. 1-140 $R_p(\lambda)$ of a reflector inside a cemented "cube" at an angle of 53° in the glass of index 1.52. The design appears in §1.10.2.9. The ordinate changes scale at 0.90.

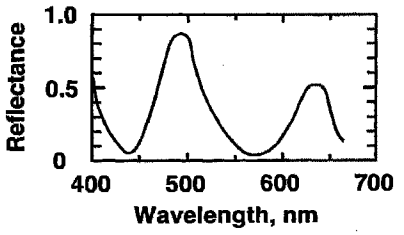


Fig. 1-141 $R(\lambda)$ of
air **M L** *aluminum*
where **L** is a dielectric layer and **M** a semitransparent metal. The *aluminum* is opaque. After Tripp [52].

The reflectance for *s* polarized light is always quite high. The difficulty is in producing high R_p. Consider a 45° angle of incidence. Although the data in Fig. 1-138 pertain to normal incidence, we can estimate that 15 layers are needed to attain 99% *R*, when the incident medium is air. Figure 1-139 shows that 29 layers are required for a reflector inside a cube. At 53° incidence angle the situation is even worse — 56 layers are needed. Figure 1-140 shows that the spectral bandwidth of the reflector is relatively narrow. An estimate is that at least 400 layers are needed to produce at least 99% reflectance in the spectral region 400 nm to 700 nm. A further complication is that there are angles at which this reflector is difficult to produce because the required refractive indices are not readily available.

1.7.7 Reflectors containing metal layers

These coatings are used in a variety of applications. An absorbing coating can have a dissimilar reflectance from its front and back sides (see §1.2.2). So-called "one way glass" has the same *T* in both directions, but dissimilar *R* (see Fig. 1-13). A single chromium layer is widely used. Its *R'* from its reverse (i.e., "glass") side ranges from 5% to 10%; *R* from its air side is 30% to 40%.

Figure 1-141 shows $R(\lambda)$ of opaque aluminum overcoated with two layers — a metal and a dielectric. Its reflected color shifts as the viewing angle changes — a property that makes it useful as an "anticounterfeiting" coating [Berning 87, 88; Dobrowolski 91]. This is also termed an "optical variable device."

A similar design (but with a much thicker dielectric layer) is used for the laser reflector shown in Fig. 1-142, which depicts its $R(\lambda)$ and also axial modes of an injection laser [Anthony 86]. Its reflectance is large enough at one mode (*arrowed*) so that it lases. Lower R at other modes suppresses their lasing.

Figure 1-143 depicts $R(\lambda)$ of a coating reflecting more than 40% in the red from its reverse (i.e., glass) side. Its R is relatively neutral from its air side [Apfel 72p].

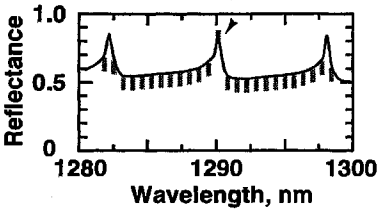


Fig. 1-142 $R(\lambda)$ of reflection filter used on an injection laser — axial modes of which are the vertical shaded lines. After Anthony [86].

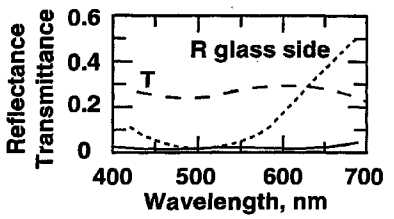


Fig. 1-143 $T(\lambda)$ (*long dash*) and $R(\lambda)$ from air side (*solid*) and $R'(\lambda)$ from reverse side (*short dash*) of a coating containing nickel. After Apfel [72p].



Fig. 1-144 In an optical waveguide, a meridional ray reflects from wall (*shaded*) at an angle θ' to that wall.

Fig. 1-145 At 10.6 μm, computed loss [see Eq. (1-4b)] for *p* polarization (*short dash*) and *s* polarization (*long dash*) vs optical thickness (in waves at 14.3 μm) of a layer of refractive index 1.40 that overcoats a substrate of optical constant $11.5 - j63$. After Trost [94].

1.7.8 Reflector for an optical waveguide

Figure 1-144 depicts a meridional ray reflecting from the walls of an optical waveguide (of circular cross section) at an angle of incidence θ' to its surface (i.e., *not* to the normal) of a few degrees. Its walls should be highly reflecting for both polarizations at 10.6 μm. They are covered with opaque silver overcoated with a dielectric layer of relatively low refractive index. Figure 1-145 graphs loss (see Eq. (1-4b) for both polarizations vs optical thickness of the layer. A conclusion is that naked silver reflects *s* polarization superbly but is quite "lossy" for *p* polarization. The converse is true when layer thickness approaches a quarterwave — loss for *s* is astronomical and loss for *p* approaches a few parts per thousand. An arrow in Fig. 1-145 designates a compromise solution. An optical thickness of 0.17 waves (after correction is made for angle of incidence in the layer) provides a relatively low loss for both polarizations. An overcoat of a higher refractive index results in increased loss.

1.7.9 A checklist for the specification of reflector

The specified parameters for a reflector could include:

Radiant reflectance or radiant transmittance*	Clear aperture (see Fig. 1-206)
Allowed level of mechanical stress	Substrate composition and smoothness
Angle of incidence	Laser damage threshold
State of polarization of the incident light	Phase shift upon reflection
Differential phase shift upon reflection	Optical figure

Some comments relating to the specifications are: *Always provide tolerance levels. Examples: "*R* to exceed 98%." or "*R* must exceed 95% but must not exceed 97%." "The differential phase shift upon reflection should not exceed 20 degrees nor be less than 5 degrees."

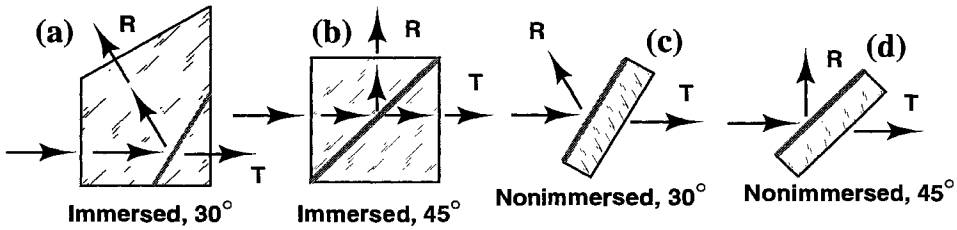


Fig. 1-146 At angles of incidence of 30° and 45°, showing the beams reflected *R* and transmitted *T* of immersed and nonimmersed coatings (*shaded*).

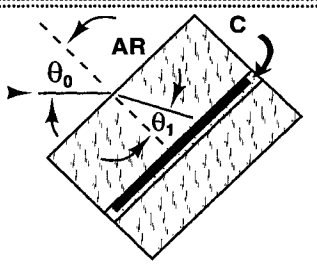


Fig. 1-147 Ray incident at $\theta_0 = 45°$ (in air) refracts in plane parallel plate upon which a coating *C* (*solid black*) is deposited. The angle of refraction in the plate is $\theta_1 = 27.7°$.

1.8 Beamdividers, dichroics and polarizers

1.8.1 Introduction

The functions of these coatings are quite different. A beamdivider (BD) splits a beam, a dichroic produces colored light and linearly polarized flux emerges from a polarizer. Yet what they have in common is that they must function at nonnormal incidence. For this reason they are grouped together in this section.

1.8.1.1 Immersed and nonimmersed coatings

Angle of incidence θ_0 and refractive index n_0 of the incident medium determines the immersion constant L

$$L \equiv n_0 \sin\theta_0 \ . \qquad (1\text{-}19)$$

A rule of thumb is that a coating is immersed if L exceeds 0.95. In physical optics and lens design, the product "$n \sin\theta$" is known as *numerical aperture*. A coating is immersed by encasing it in a cemented prism, as shown in Fig. 1-146. Only by immersion can L exceed 1.00. Table 1-11 lists L for coatings in Fig. 1-146.

A coating may have glass (or optical cement) for both its incident medium and emergent medium and yet not be immersed. Figure 1-147 shows a coating deposited on the reverse side of a plane parallel slab, the opposite side of which is coated with an

antireflection coating. The incident medium for this coating is glass. This coating is not immersed, however, because the reference medium is air (n_{ref} = 1.0003), where the reference angle θ_{ref} is 45°. Equation (1-19) is written

$$L = n_{ref} \sin\theta_{ref} = 1.0003 \sin(45°) = 1.52 \sin(27.7°) = 0.71 \ . \qquad (1\text{-}20)$$

Addition of a plane parallel plate (whatever its refractive index) in no way alters L, since θ_{ref} and n_{ref} determine the immersion constant.

1.8.1.1.1 Effects of immersion

Consider the coatings at the 45° angle of incidence in Fig. 1-146. The difference between the immersion constants, with values of 1.08 and 0.71 (for immersed and nonimmersed designs, respectively), profoundly affects its performance. Unless the coating materials are selected with great care, immersed multilayers usually reflect the *s* polarization quite strongly and the *p* polarization weakly. This is termed *polarization splitting*.

Table 1-11 Immersion constants of coatings

Identification in Fig. 1-146	Index n_0	Angle θ_0	L	Other Figs.
Immersed, 30°	1.85	30°	0.93	1-183
Immersed, 45°	1.52	45°	1.08	1-151, 1-168, 1-170, 1-183,1-180
Nonimmersed, 30°	air 1.00	30°	0.50	1-152
Nonimmersed, 45°	air 1.00	45°	0.71	1-153, 1-178

Immersion constant [see Eq. (1-19)] L of slabs and cemented prisms in Fig. 1-146. Immersed coatings and coatings on slabs also appear in other Figs.

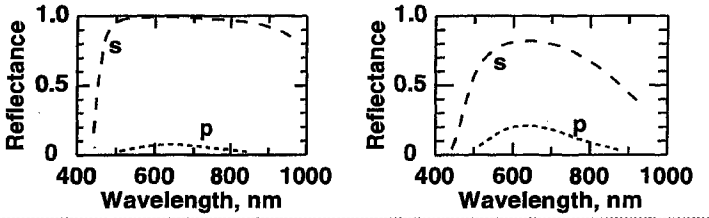


Fig. 1-148 At θ_0 of 58°, $R_s(\lambda)$ (*long dash*) and $R_p(\lambda)$ (*short dash*) of
(*left panel*) *glass* **H L H L H** *glass* and
(*right panel*) *glass* **L M L M L** *glass* .
Refractive indices of *glass*, **L**, **M** and **H** are 1.517, 1.45, 1.822 and 2.30, respectively. Optical thicknesses of **L**, **M** and **H** are $\lambda_0/4$ at λ_0 of 1372 nm, 894 nm and 763 nm, respectively.

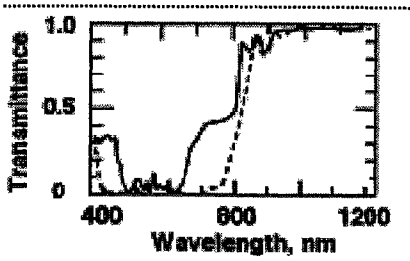


Fig. 1-149 At θ_0 of 45°, $T_a(\lambda)$ of cold mirror coatings on a slab (*dashed*) and immersed in a cemented cube (*solid*). After Balzers [92].

A comparison of immersed coatings illustrates this splitting. The left panel of Fig. 1-148 depicts R_p and R_s of a five-layer quarterwave stack of alternate layers of titania and silica. Its polarization splitting is quite large — R_s is 99% and R_p is only a few percent. This is a superb coating if the goal is to produce a polarizer. The right panel of Fig. 1-148 shows R of the same five layers when yttria is substituted for the titania. Although the polarization splitting has not vanished, it is not nearly so severe.

The contrast between immersed and nonimmersed coatings is evident in Figs. 1-149 and 1-150. The coatings in Fig. 1-149 should have low transmittance in the visible spectral region at an angle of incidence of 45°. A nonimmersed multilayer (*dashed*) which is coated on a slab does a reasonable job. An immersed coating (*solid*) has undesirable transmission leaks in the blue and red parts of the spectrum. The angle of incidence is also 45° in Fig. 1-150, which depicts $T(\lambda)$ of "fifty-fifty beamdividers." There is a 2:1 ratio of R_s to R_p for the slab configuration. That ratio for the immersed coating is 9:1.

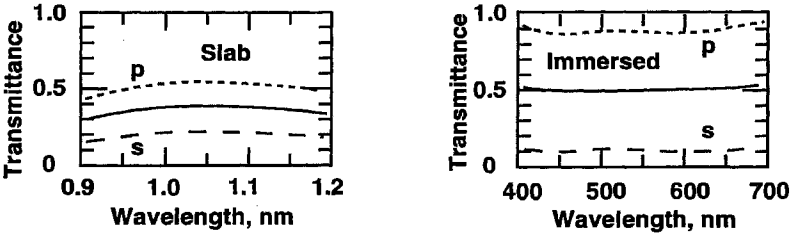


Fig. 1-150 At θ_0 of 45°, $T_a(\lambda)$ (*solid*), $T_s(\lambda)$ (*long dash*) and $T_p(\lambda)$ (*short dash*) of coatings on a slab (*left panel*) and immersed in a cemented cube (*right panel*). After Continental [90].

Table 1-12 lists attributes of nonimmersed and immersed coatings. Notwith-standing its lower cost, often a slab cannot be used because it introduces astigmatism when a convergent beam impinges upon it.

Table 1-12
Comparison of immersed and nonimmersed coatings

	Type of coating	
Attribute	Slab (nonimmersed)	Immersed
Cost	Less	Greater, especially in large sizes
Optical performance	Introduces astigmatism when illuminated by a convergent beam	Free of astigmatism
Polarization splitting	Moderate	Severe
Used with a pulsed laser	High laser damage threshold	Damages easily because of cement

Attributes of coatings on a slab and immersed (i.e., cemented-in-glass).

1.8.1.2 Properties of beamdividers, dichroics and polarizers

Attributes of these coatings are:

(a) *Reflectance and transmittance for randomly polarized flux* These terms are defined in Eqs. (1-5a) and (1-5b). If randomly polarized flux is reflected at nonnormal incidence only once in an optical system and is collected by a detector insensitive to polarization, then R_a and T_a are meaningful. Figure 1-150 shows that although T_a may be nearly identical for two coatings, the polarization splitting is considerably different. Consider the immersed coatings in Fig. 1-151 that combine light from four lamps. T_a and R_a are meaningful for cubes A and B. Unless Herculean efforts are made to avoid it, light hitting prism C is highly polarized. Hence, T_a and R_a are often meaningless in specifying coatings for such a cemented cube.

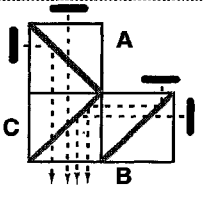


Fig. 1-151 Three immersed beamdividers combine flux from lamps in an illuminator. After Weiss [78].

(b) *Nonpolarizing* This implies that R_p and R_s are the same — or nearly so.

(c) *Phase shift upon reflection* A coherent system is one in which the phase of a light wave is preserved and two or more waves combine to interfere constructively or destructively. This phase shift determines the relative phase of a wave after it reflects from a beamdivider. Figure 1-108 shows a ring laser gyro in which two total reflectors T and a beamdivider output mirror D form a cavity [Walker 74]. If there is a different phase shift upon reflection for s and p polarized components, then the apparent cavity lengths will not be the same for the two polarizations.

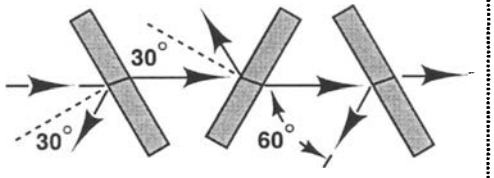


Fig. 1-152 Light incident of 30° upon slab beamdividers provides complete separation of reflected beams.

(d) *Susceptibility to laser damage* A cemented coating (illustrated in Fig. 1-147) self-destructs at relatively low irradiance levels and, therefore, finds little use in a high power laser system.

(e) *Angle of incidence* An important consideration in integrating a beamdivider or dichroic into an optical system is its angle of incidence — especially if polarization effects are to be avoided. It is noteworthy that 30° incidence provides a complete separation of adjacent beams, as shown in Fig. 1-152. An advantage of 30° incidence is that polarization effects are considerably less than at 45° or higher angles.

(f) *Angular acceptance cone* This is a measure of the amount the angle of incidence can be varied and yet not degrade coating performance. An illustration is that the transmittance of the slab polarizer in §1.8.3.2 is fairly sensitive to angle.

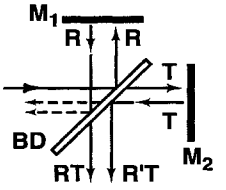


Fig. 1-153 Relative irradiances of beams traversing two paths of a Michelson interferometer. R and R' are reflectances from obverse and reverse sides of the beamdivider. T is its transmittance. Mirrors M_1 and M_2 are totally reflecting.

(g) *RT product* This is a measure of coating efficiency when beams both reflect from and are transmitted through the same coating. Fig. 1-153 shows a Michelson interferometer in which a coating on the beamdivider is possibly absorbing. Therefore, reflectance R from its incident side and R' on its reverse side are nonidentical (see §1.2.2.3). If a beam emerging from a light source has an irradiance $\mathcal{E}$, this irradiance is $\mathcal{E}\,R\,T$ after reflecting from the beamdivider, reflecting from mirror M_1 and transversing the beamdivider. Loss at mirror M_1 is neglected, since its R can be as high as 99%. Irradiance in the other path is $\mathcal{E}\,R'\,T$.

A conclusion is that RT product is an important measure of system efficiency. In the optimum case, the beamdivider is nonabsorbing and its R is 0.5 (i.e., 50%), whence RT product is 0.25. It is surprising, however, how much R can stray from 50% without permitting RT product to degrade seriously. An illustration is that if $R = 0.3$ and $T = 0.7$, RT product is 0.21. This is only 16% below the optimum of 0.25.

Finally, note that *any* nonabsorbing beamdivider — regardless of its RT product — still produces fringes of 100% contrast. Any absorption of the beamdivider reduces its efficiency, but does not alter contrast as long as RT products in *both* arms of the interferometer are identical.

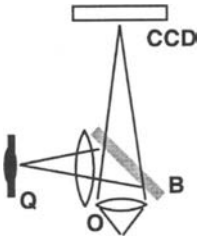


Fig. 1-154 In a reflective microscope, a light from the source *Q* reflects from the slab beamdivider *B* and thence into the objective *O*. The light reflected by the sample is focused onto a CCD array.

(h) *One-sided reflectance or two-sided reflectance?* Light reflects from *both* sides of the beamdivider in the interferometer in Fig. 1-153. This requirement of a reflectance on both sides makes the coating design more complex.

Figure 1-154 depicts a beamdivider in which the light reflects from only *one* side. A sample that is inspected in a reflection microscope is illuminated by reflecting the light from the beamdivider and then into the objective. The beamdivider coating is simplified because the *RT* product is specified for the reflectance from only one side.

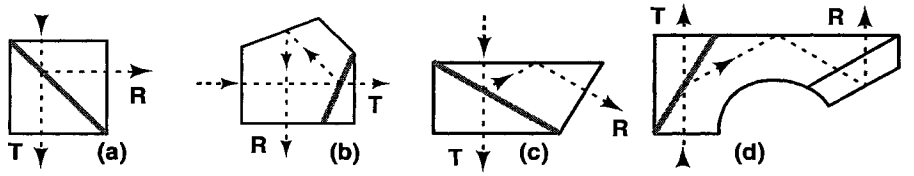


Fig. 1-155 Coated surfaces (*shaded*) are added to prisms that adapt them as beamdividers: (a) Porro; (b) penta prism; (c) 30° deviation prism; (d) Leman.

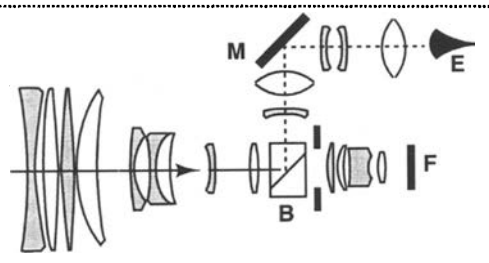


Fig. 1-156 Inclusion of an immersed beamdivider *B* in a video camera objective lens directs light towards the viewfinder and the human eye *E*.

1.8.1.3 Adaptation of prisms as beamdividers

It is tempting to consider only the cemented cube in panel (a) of Fig. 1-155 as a configuration for an immersed beamdivider. The penta prism (panel *b*) has an advantage of 22.5° incidence angle (see Fig. 1-185). §1.8.4.2.1 emphasizes that this relatively small angle reduces the polarization splitting. The prism in panel (c) of Fig. 1-155 is used in binocular microscopes to deviate the image direction from the vertical — angle of incidence at the coating is 30°. Incidence angle in the Leman prism in panel (d) of Fig. 1-155 is 60°. There are instances in which the polarization splitting of the beamdivider may be tolerated, as in the viewfinder in the video camera in Fig. 1-156

Figure 1-157 depicts trapezoidal prisms in which the deviation angle is 90°, yet the incidence angle on the coating is 55°.

1.8.2 Beamdividers

A convention is that reflectance and/or transmittance of a beamdivider is relatively neutral, whereas a dichroic reflects (and transmits) certain colors in preference to others.

1.8.2.1 Slab beamdividers

In this section, no effort is made to constrain the coatings to be nonpolarizing.

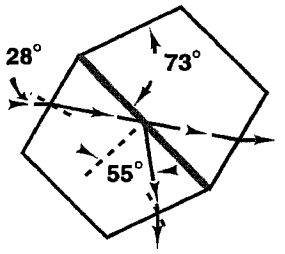


Fig. 1-157 A coating (*shaded*) is cemented between trapezoidal prisms (of index 1.517) that deviate a beam by 90°.

1.8.2.1.1 Single-layer and multilayer beamdividers

A simple slab beamdivider coating is a single film of zinc sulfide or titania deposited on a glass substrate. This divider is relatively neutral in color and has a R_a of approximately 40% at 45°, as depicted in Fig. 1-158. An antireflection coating is coated on the opposite side of the slab. Hass [52] describes how a simple titania divider is fabricated by depositing a titanium layer and then oxidizing it by baking. Figures 1-150 and 1-159 depict $T(\lambda)$ of multilayer beamdividers that span the visible spectral region.

1.8.2.1.2 Pellicles

A pellicle is a relatively thin sheet of organic material that ranges in thickness from 80 nm to 50 μm. It can be coated on both sides; because of thinness of the sheet, image doubling is sometimes unnoticeable.

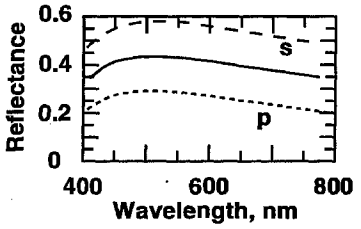


Fig. 1-158 At angle of incidence θ_0 of 45° in air, $R_a(\lambda)$ (*solid*), $R_s(\lambda)$ (*long dash*) and $R_p(\lambda)$ (*short dash*) of a rutile layer coated on a BK-7 slab.

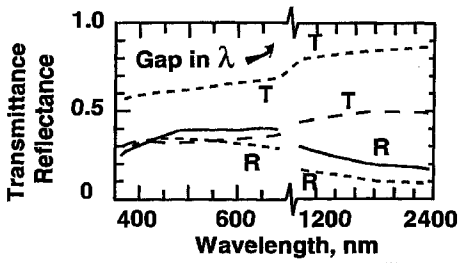


Fig. 1-159 At θ_0 of 45° in air, $T_a(\lambda)$ (*solid*), $T_s(\lambda)$ (*long dash*) and $T_p(\lambda)$ (*short dash*) of a beamdivider on a slab. After Matra [89].

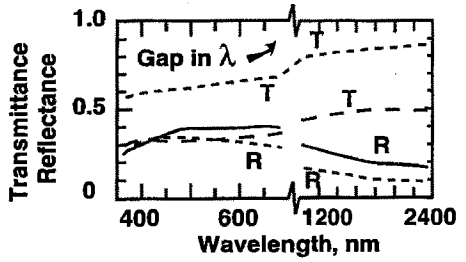


Fig. 1-160 At θ_0 of 45° in air, R and T of two coated pellicles: (a) $T_a(\lambda)$ and $R_a(\lambda)$ (*short dash*). (b) $T_a(\lambda)$ (*long dash*) and $R_a(\lambda)$ (*solid*). Wavelength scale changes at 1000 nm. After National Photocolor [90].

In all but exceedingly thin sheets, however, interference effects between coatings on two sides are easily observed if illuminated with a laser or other narrowband coherent source of illumination. Due to its thinness, minute air currents easily distort a pellicle, thus contorting the reflected image.

Often the plastic is associated with trade names such as Kapton, Tedlar or Mylar [Dupont 78]. The last can be obtained in sheets 2500 nm thick. But even thinner sheets can be produced if a nitrocellulose solution is deposited on a water surface. The solvent evaporates and the film is picked up on a ring. With skill and patience, films 80 nm thick and 50 mm ϕ can be produced [Harris 67].

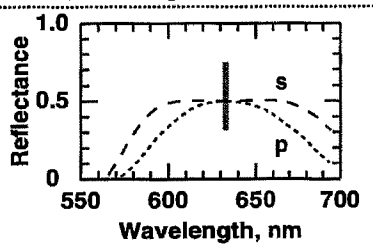


Fig. 1-161 At θ_0 of 45°, in the incident medium of air, $R_s(\lambda)$ (*long dash*) and $R_p(\lambda)$ (*short dash*) of an all-dielectric nonpolarizing beamdivider. The design is in §1.10.2.10.

Coating a pellicle is no trivial task. Not only is it difficult to adhere the coating to plastic, but mechanical stress in the coating can buckle it. One solution is to coat the sheet on both sides to compensate for stress. Figure 1-160 depicts $T(\lambda)$ of pellicle coatings [National Photocolor 90]. Layton [71] observed that uncoated polyethylene reflected as much as 27% at 10.6 μm.

1.8.2.1.3 Slab beamdividers with metallic layers

Prior to the advent of vacuum deposited films in the 1930s, beamdividers of a typical Michelson interferometer consisted of a chemically deposited layer of silver. In the visible part of the spectrum, it reflected about one-third and transmitted one-third — the remainder was absorbed. A semitransparent aluminum layer is even more absorbing; it is used only in desperation. A metallic layer has several advantages, notwithstanding its absorption. First, its polarization splitting (see §1.8.1.1.1) is usually considerably less than that of a dielectric coating. Another plus is that its phase shift upon reflection varies relatively little with wavelength, thus making it ideal for interferometers used in metrology.

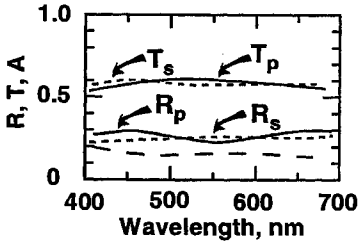


Fig. 1-162 At θ_0 of 45°, absorption (*long dash*), $R_p(\lambda)$, $R_s(\lambda)$, $T_p(\lambda)$ and $T_s(\lambda)$ of a cemented cube beamdivider. After Refermat [71].

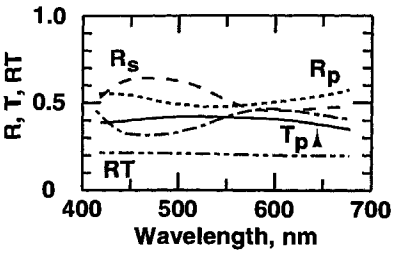


Fig. 1-163 At θ_0 of 45°, computed $R_s(\lambda)$ (*long dash*), $R_p(\lambda)$ (*short dash*), $T_s(\lambda)$ (*solid*), $T_p(\lambda)$ (*long-short*) and *RT* product for *p* polarization (*lowest curve*) of an absorbing beamdivider whose design is in §1.10.2.11

1.8.2.1.4 Nonpolarizing

Figure 1-161 depicts computed $R_p(\lambda)$ and $R_s(\lambda)$ of a 50%-50% beamdivider that is nonpolarizing at λ_0 of 633 nm. Its band of 50% *R* extends only a few percent on either side of λ_0. Splitting ratios other than 50%-50% can be produced. Widening the bandwidth is more difficult.

1.8.2.2 Immersed beamdividers

Figure 1-150 depicts $T(\lambda)$ of an immersed beamdivider. Its polarization splitting is quite large. This splitting can be reduced with metallic layers (see §1.8.2.2.1) or with all-dielectric designs over a relatively narrow spectral region (see §1.8.2.2.2).

An air gap between two prisms can function as a beamdivider [Land 34]. A variant of this is to use a high index glass prism (n = 1.60) and a relatively low index film (n = 1.35) in which the incidence angle exceeds the critical angle [Turner 57].

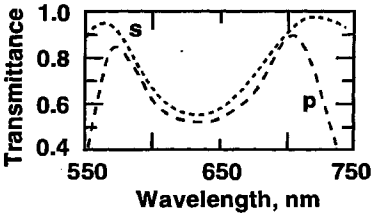


Fig. 1-164 At angle of incidence of 45° in glass, $T_s(\lambda)$ (*short dash*) and $T_p(\lambda)$ (*long dash*) of nonpolarizing immersed beamdivider at 633 nm. After Broomer [83].

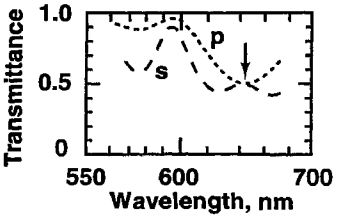


Fig. 1-165 At θ_0 of 45°, $T_p(\lambda)$ (*short dash*) and $T_s(\lambda)$ (*long dash*) of nonpolarizing immersed beamdivider at 633 nm. After Spectra-Physics [73].

1.8.2.2.1 Immersed beamdividers with metallic layers

Figure 1-162 shows a beamdivider with 30% *R* and 60% *T* that is relatively nonpolarizing and neutral it its color. This coating consists of a silver film sandwiched between layers of zinc sulfide [Refermat 71]. There is 10% to 15% absorption over the visible spectrum. Are other splitting ratios available? Figure 1-163 depicts *R* and *T* of a more complex design produced by computer optimization. Its *RT* product is close to 20%. Its angular

bandwidth is not as wide as that of Fig. 1-162 because it contains layers of relatively lower refractive index.

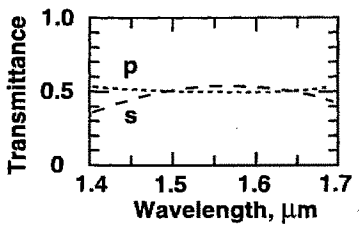


Fig. 1-166 At θ_0 of 45°, $T_p(\lambda)$ (*short dash*) and $T_s(\lambda)$ (*long dash*) of a nonpolarizing immersed beamdivider for the near IR. After Nitto [98].

1.8.2.2.2 Nonpolarizing beamdividers

Figures 1-164 and 1-165 depict the spectral transmittance of beamdividers (at 45° incidence) that manifest minimal polarization splitting at 633 nm. Both show a 50%-50% splitting ratio. Figure 1-166 shows a similar coating that functions in the near IR. The bandwidth of such coatings is usually narrow. The refractive indices in such coatings must be controlled with utmost precision, which makes them not easy to fabricate.

Table 1-13
Attributes of slab and MacNeille linear polarizers

	Type of coating	
Attribute	Slab	MacNeille
Cost	Less	Greater, especially in large sizes
Typical angle	57°	45° to 55°
Optical bandwidth	Usually less than 5% of λ_0	Can exceed 1.5:1
Reflectance of *p* polarization	Below 1%	Usually a few percent
Used with a pulsed laser	High laser damage threshold	Damages easily because of cement

1.8.2.3 Specification of a beamdivider

This topic is discussed in §1.8.6.3.

1.8.3 Linear polarizers

The goal of these devices is the ultimate polarization splitting — zero reflectance for *p* polarization and 99.9% for R_s. Figures 1-167 and 1-168 show slab and immersed configurations of such polarizers. The immersed is termed a MacNeille [46] polarizer. Table 1-13 compares their attributes.

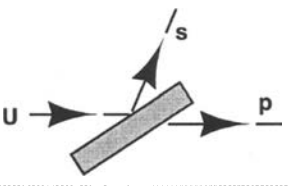


Fig. 1-167 Randomly polarized light *U* is incident upon slab polarizer that transmits *p* linear polarization and reflects *s*.

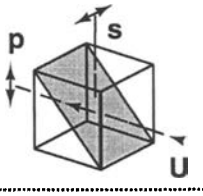


Fig. 1-168 Randomly polarized light *U* illuminates a MacNeille polarizer. The *p* component is transmitted and *s* is reflected.

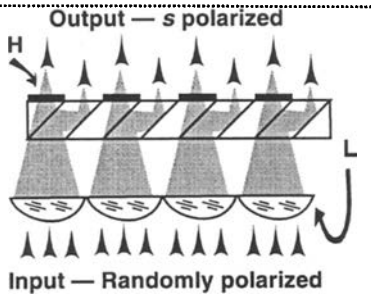


Fig. 1-169 Randomly polarized light impinges upon a lenslet array *L* followed by a polarization converter consisting of an array of MacNeille polarizers. Alternate MacNeilles are covered with a halfwave plate. After [Nitto 98].

1.8.3.1 Applications

The slab polarizer is used widely in laser systems. An illustration is that a slab linear polarizer and Faraday rotator are inserted between amplifiers in a pulsed neodymium glass laser to prevent amplification of reflected beams.

The polarization converter in Fig. 1-169 is widely used in color projection systems. It consists of an array of MacNeille polarizers. Randomly polarized light impinges upon a lenslet array, which concentrates the light upon alternate polarizers. The transmitted light from the polarizer is *p* polarized, which is converted into *s* polarized light by a halfwave linear retarder, also termed a halfwave plate. Since the coated interfaces strongly reflect s polarized light, after two bounces *s* polarized light emerges. As much as 80% of the light emerges *s* polarized.

The illuminator in Fig. 1-170 also produces linearly polarized light. Randomly polarized light *U* from lamp *L* impinges upon a MacNeille polarizer *M*. The *s* polarization is reflected and is converted to *p* polarization when it passes through a halfwave plate (alias halfwave linear retarder) *H*. The *p* is transmitted by the next MacNeille *M'* that it hits. The light is reflected by mirror *D*.

A system that writes upon and reads out a signal from a magneto-optical disk (see Fig. 1-171) contains three MacNeille polarizers.

1.8.3.2 Slab polarizers

These devices are useful when the range of wavelengths is not excessive — say, 5% of λ_0. They are used in preference to cemented beamdividers when: (a) inserted into a high irradiance laser beam; (b) the astigmatism introduced by the slab (see §1.8.1.1.1) can be tolerated; and (c) a lower cost is requisite. A slab is less expensive because it uses less glass and does not require cementing.

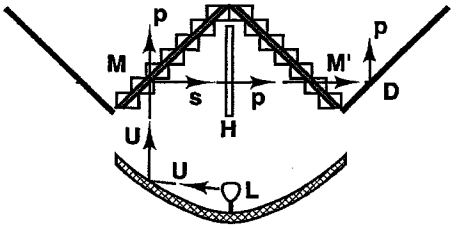


Fig. 1-170 Randomly polarized light U is converted into linearly polarized light p:
M — MacNeille polarizer.
H — halfwave plate. D — total reflector.
L — light source.
After Schwarzmüller [75].

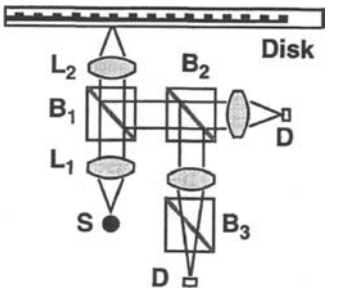


Fig. 1-171 Optical system that reads and writes on a magneto-optical disk. Source S is focused by lenses L_1 and L_2 onto the disk. The signal is collected by the detectors D. Each beamdivider B is a MacNeille polarizer. After Nitto [98].

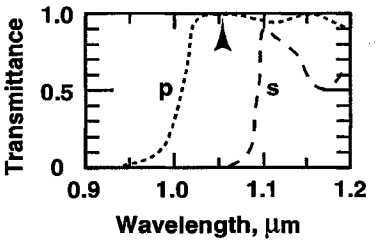


Fig. 1-172 At θ_0 of 57°, $T_p(\lambda)$ (*long dash*) and $T_s(\lambda)$ (*short dash*) of a slab linear polarizer. After Continental [90]. 1.06 μm is arrowed.

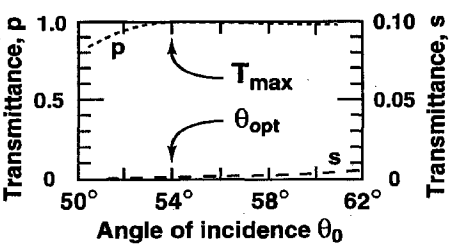


Fig. 1-173 For a slab polarizer at 1.06 μm, measured $T_p(\theta_0)$ (*short dash*, left ordinate) and $T_s(\theta_0)$ (*long dash*, right ordinate) vs θ_0. After Thomas [76].

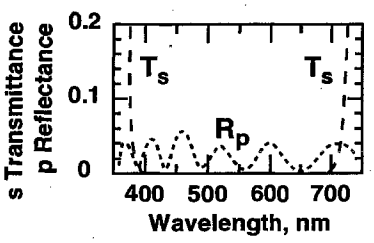


Fig. 1-174 $R_p(\lambda)$ (*short dash*) and $T_s(\lambda)$ (*long dash*) of "Type I polarizer." After Meadowlark [91].

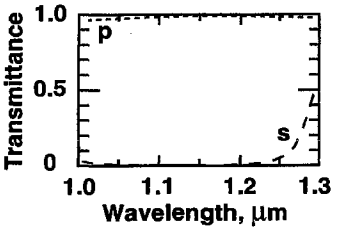


Fig. 1-175 $T_p(\lambda)$ (*short dash*) and $T_s(\lambda)$ (*long dash*) of "narrow-band polarizing cube." After Lambrecht [90].

One disadvantage of the slab is that it functions at angles of incidence between 50° and 60°, thus producing a total angle between incident and reflected beams of at least 100°. The 90° beam separation of a cemented cube is often more convenient.

As depicted in Fig. 1-172, a slab polarizer is an edge filter with polarization splitting. This coating functions in a relatively narrow spectral region where T_p zooms upwards to high values and T_s is still relatively low.

T_p of a slab polarizer can be adjusted by "angle tuning." Figure 1-173 shows p and s transmittances vs angle at a wavelength of 1.06 μm [Thomas 76]. T_p is in excess of 99% and T_s is below 0.1%. It is also evident that $T(\lambda)$ of this coating could not be measured accurately on a conventional spectrophotometer, where the semi-angle of the converging cone of flux on a sample is typically ±4°. A spectrophotometer would indicate a lower T_p and poorer rejection than is actually achieved.

1.8.3.3 MacNeille polarizer

Figure 1-168 shows an immersed coating that transmits p and reflects s polarization [MacNeille 46]. It is quite efficient in reducing unwanted s in its transmitted beam — T_s is usually below 10^{-3}. It is more difficult, however, to reduce its R_p so that the reflected beam is pure s polarized. In many polarizers 2% to 4% of the reflected flux is the (unwanted) p polarized component. Figure 1-174 graphs $T(\lambda)$ of a coating that functions over most of the visible spectrum. The polarizer, whose $T(\lambda)$ appears in Fig. 1-175, polarizes in the near IR. Similar to a slab polarizer, T and R of a MacNeille can be adjusted by "angle tuning," although this destroys the directional orthogonality of transmitted and reflected beams.

The $T(\lambda)$ cited above occurs when its "tuned" incidence angle is 45°. What happens when this is varied by a few degrees? At the risk of making a generalization, T_s usually remains quite low. Inevitably, its p transmittance degrades — T_p may decrease from nearly 100% at its "tuned" angle to 90% or lower.

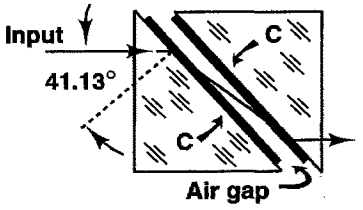


Fig. 1-176 The Ward polarizer consists of two coated prisms separated by a relatively thin air gap. C is the optical coating

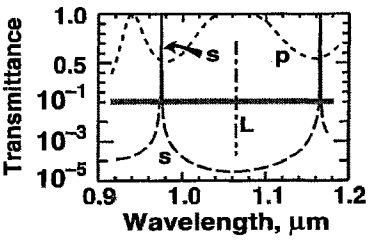


Fig. 1-177 At θ_0 of 41.16° in glass, $T_s(\lambda)$ (*long dash*) and $T_p(\lambda)$ (*short dash*) of a Ward polarizer. The design is in §1.10.2.5. The scale of the ordinate changes from linear to log at 0.10.

1.8.3.4 Ward polarizer for lasers

This unique coating (invented by Ron Ward [96]) is designed to function only in a beam produced by a well-collimated single wavelength laser. The coating manifests a quite substantial laser damage threshold. As depicted in Fig. 1-176, the hypotenuses of two prisms are each coated with an all-dielectric multilayer and then separated by an air gap. The coatings are an antireflection for p-polarized light but highly reflect the s. The system functions like an "air-gap single-cavity filter" (see §1.6.2.1) for s polarization. The s transmission "spikes" in Fig. 1-177 are the quite narrow passbands of this filter.

Figure 1-177 shows that there is a spectral region where T_p approaches 100% — that is where each of the coatings has a quite low p reflectance. T_s is at least 3 or 4 absorbance units at all wavelengths except where narrow transmission "spikes" occur. By adjusting the spacing of the air gap during assembly, those spikes are well separated

from the laser wavelength at which the polarizer functions. Caveat: This coating is exceedingly sensitive to angle of incidence. For this reason, its $T(\lambda)$ could never be measured accurately on a conventional spectrophotometer.

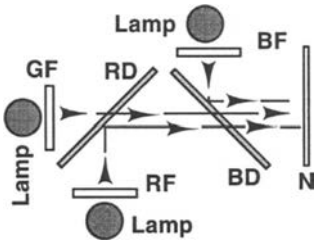


Fig. 1-178 In a color printer, red-reflecting *RD* and blue-reflecting *BD* dichroics combine light converging on the negative *N*. *BF*, *GF* and *RF* are trim filters. After White [58].

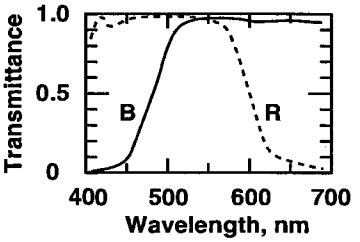


Fig. 1-179 At 45° angle of incidence in air, $T_a(\lambda)$ of blue-reflecting (*solid*, *B*) and red-reflecting dichroic mirror (*dashed*, *R*). After Bausch [60].

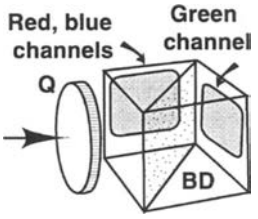


Fig. 1-180 Color balance of an immersed beamdivider in a cinematographic camera is improved by a quarterwave plate *Q* that converts linearly polarized light into circularly polarized. After Ball [39].

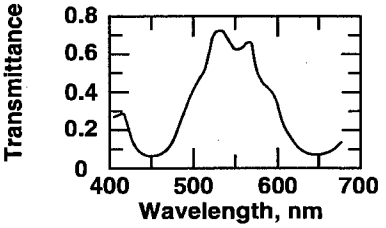


Fig. 1-181 At 45° incidence angle, measured average transmittance of a green-transmitting and red-blue reflecting coating in a cemented cube, as in Fig. 1-180. After Kelly [56]

1.8.3.5 Specification of a polarizer

This topic is discussed in §1.8.6.3.

1.8.4 Dichroic reflectors

Mineralogists use *dichroism* to describe materials that transmit various hues of light, depending upon the polarization of its illumination. Its usage has broadened to include optical coatings that reflect and transmit different colors. In essence, *dichroic beamdividers* — also called *dichroic mirrors* — may be viewed as edge filters (see §1.5.2) that function at nonnormal incidence.

1.8.4.1 Slab dichroic reflectors

Dichroic coatings are widely used for color separation [H. H. Schroeder 53]. Reflected and transmitted colors from a dichroic are complementary because it is nonabsorbing. Dichroics in a color photographic printer (shown in Fig. 1-178) combine the light from filtered lamps that illuminate a color negative with red, green and blue primary colors [White 58]. $T(\lambda)$ curves in Fig. 1-179 are typical of such coatings [Bausch 60].

Blue-reflecting and red-reflecting dichroics reflect the corresponding colors from the lamps. Both dichroics transmit green.

1.8.4.2 Dichroic reflectors for a color camera

Figure 1-180 shows the beam divider used in the Technicolor camera that transmits the green, yet reflects the blue and red. Its transmittance, as depicted in Fig. 1-181, is limited by the severe polarization effects that occur in such immersed systems — at this incidence angle of 45°, at least.

Although the color balance between the reflected and transmitted channels functions adequately when randomly polarized light is incident, the balance changes when partially polarized light is collected. Figure 1-180 shows that the solution is to insert a *quarterwave linear retarder* (also termed a *quarterwave plate*) into the optical train [Ball 39]. This plate converts any linearly polarized light into circularly polarized, which is reflected and transmitted by the dichroic in the same way as randomly polarized.

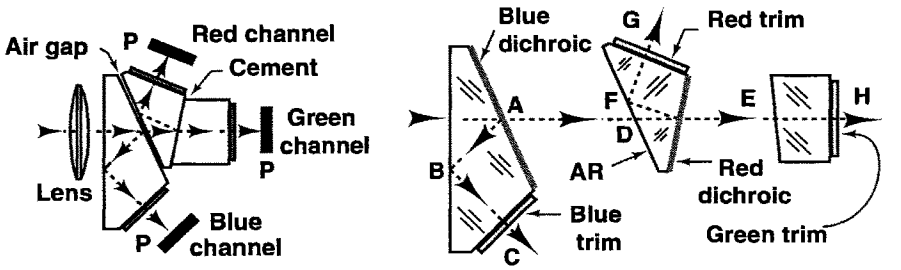


Fig. 1-182 (*Left panel*) The objective lens of a color TV camera forms images on Plumbicon tubes *P*. Blue-reflecting and red-reflecting dichroics direct light onto those respective tubes, while transmitting green. *(Right panel)* Exploded view. The heavy shaded lines represent the dichroic reflectors. After de Lang [65].

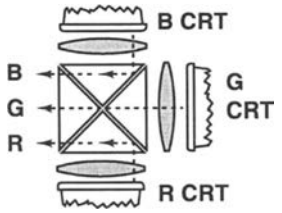


Fig. 1-183 *B CRT, G CRT* and *R CRT* are, respectively, the blue, green and red cathode ray tubes. Blue-reflecting and red-reflecting dichroics in the X-prism direct blue light *B* and red light *R* to the projection lens (*shaded*). The green light *G* is transmitted. After Schmidt [86].

1.8.4.2.1 Philips color TV camera

The designers of this system were aware of the severe polarization effects that occur at 45° incidence angle in the Technicolor system (§1.8.4.2). The solution was to use a smaller incidence angle of 30°, which reduces such effects. This design has become an industry standard for that reason [de Lang 65].

Figure 1-182 shows rays traveling from an objective lens to Plumbicon tubes in the blue, red and green channels. The objective is corrected for spherical aberration introduced by traversal of the prisms. The blue reflector is first in the chain because it is easier to fabricate a long-wavelength pass filter, as compared to a short pass. The right panel of Fig. 1-182 shows that blue light is reflected at *A* by the dichroic coating, internally reflects from a glass-air interface at *B* and finally traverses a *blue trim filter* before converging upon the Plumbicon *P*.

The surface *D* of the red-channel prism is antireflection coated, since it is noncemented. This antireflection coating need function only in the green and red, however. Design of the red reflector *E* is less complicated because it need transmit only green, since the blue has been removed. There are trim filters on all channels that reject any unwanted light that might have leaked through the dichroic.

1.8.4.3 Three-color projection systems for CRTs

The goal is to combine the light from three Cathode Ray Tubes (CRTs) so that a single projection lens may be used. Although dichroic slab beam dividers have been used in an "egg-crate" configuration [Dimmick 52b; Sachtleben 54; Albright 54; de Lang 59; Rowe 61], an immersed system is preferable for two reasons. First, the slabs may introduce astigmatism into the imaging system that would be removed by anamorphic elements or by other means. Second, it is advantageous to cement the beamdivider to the face of the CRT. This increases the effective luminance by a factor of n^2, where n is the refractive index of the glass. This cementing is not done in Fig. 1-183.

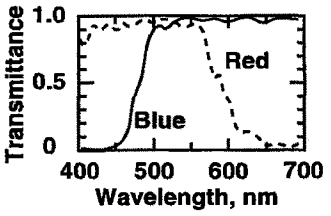


Fig. 1-184 At angle of incidence of 45°, $T(\lambda)$ of red-reflecting (*dash*) and blue-reflecting (*solid*) immersed dichroics in an X-prism similar to those in Fig. 1-183. After Nitto [98].

1.8.4.3.1 X-prism

The *X-prism* was used by Kelly [60] and Gretener [56] as a beam combined in a three-color cinematographic camera. Figure 1-183 shows dichroics are coated on the hypotenuses of at least two prisms. They are subsequently glued together. As used by Schmidt [85] in a color projection system, the light from the red cathode ray tube (R CRT) and blue cathode ray tube (B CRT) is reflected by the dichroics. The light from the green cathode tube (G CRT) is transmitted.

Figure 1-184 depicts $T(\lambda)$ of red-reflecting and blue-reflecting dichroics. Given the angle of incidence of 45°, it is by no means a trivial problem to minimize polarization splitting of these dichroics. Such splitting is reduced considerably by the 30° incidence angle in the Philips system (mentioned in §1.8.4.2.1).

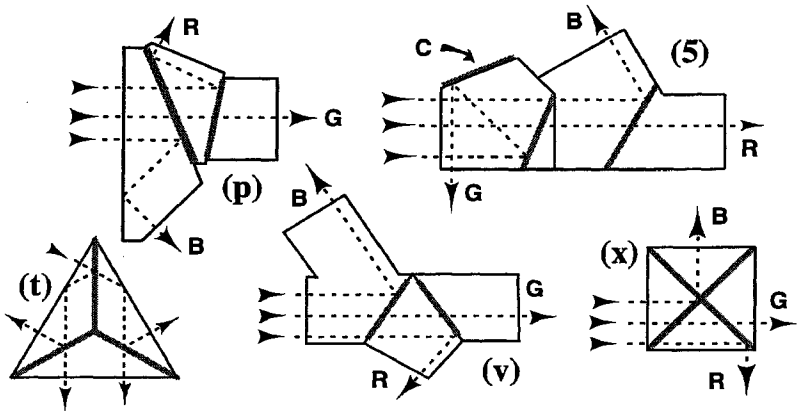


Fig. 1-185 Beam-combining prisms used in color projectors include: (*t*) triangle prism [Ball 39]; (*p*) Philips; (*v*) Vee prism; (5) Penta prism; and (*x*) X-prism. The heavy shaded lines are multilayer dichroic reflectors. The *C* coating (penta-prism) is for phase compensation. *B*, *G* and *R* refer to the blue, green and red channels, respectively.

1.8.4.4 Overview of projection systems using spatial light modulators

The cathode ray tubes in a color projection system supply their own light, whereas a system incorporating spatial modulators must be illuminated. That is, the modulator acts as a valve that varies the intensity level of the light that impinges upon it. It accomplishes this, pixel-by-pixel, in both transmissive and reflective systems. Examples of a modulator are a liquid crystal display (LCD) panel and a deformable mirror, such as the Digital Micromirror Device [Texas Instruments 99]. The LCD functions as a valve that, under control of a voltage, enables a selected pixel to alter the state of polarization of the light that reflects from it or is transmitted through it.

Figure 1-185 depicts prism assemblies that have been either used in or proposed for use in three-color projection systems. Attributes of these systems are:

(a) Back focal length.
(b) Polarization splitting of the "edge" transmission curve of a dichroic.
(c) Angle shift of the "edge" transmission curve of a dichroic.
(d) Preservation of the state of polarization.

Each is discussed below.

1.8.4.4.1 Back focal length

The projection lens forms a magnified image of the LCD panel upon a viewing screen. The lens back focal length is the total optical path from its principal plane to the panel. The smaller this distance, the closer the projector may be positioned to the screen — for a given magnification. Hence, a minimal back focal length is usually desirable.

As an example, the penta prism and the V-prism in Fig. 1-185 contain more glass than other systems, such as the Philips. They are used infrequently because of this longer back focal length.

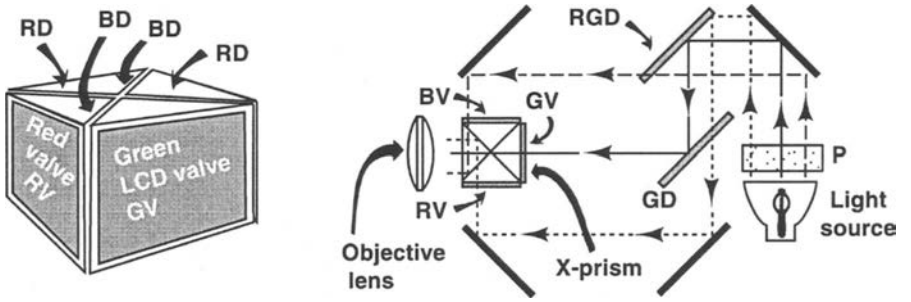


Fig. 1-186 (*Right panel*) Path of axial rays for the blue (*long dash*), red (*short dash*) and green (*solid*) that propagate from the light source, through a polarization converter, slab dichroics and thence to the X-prism. (*Left panel*) Perspective view of an X-prism. Please consult the text for explanations of the symbols. After [Epson 99].

1.8.4.4.2 "Edge" transmission of a dichroic — its polarization splitting

The section title refers to item (b) in §1.8.4.4 and is illustrated by a color projector system [Epson 99]. The conclusion is that it is possible to design a system in which the polarization splitting of the edge transmission is not a problem. This is done by controlling the linear polarization state of the light that is reflected from and transmitted by the X-prism.

Referring to Fig. 1-186, the light emerges from the source and traverses a polarization converter (see §1.8.3.1). The emergent light is mostly *s* linearly polarized — the electric vector is perpendicular to the paper. After reflecting from a turning mirror, the blue light (*short dash*) is transmitted by the red-green beamdivider (*RGD*). That blue light then impinges upon the blue LCD valve *BV*. As depicted in Fig. 1-187, the blue dichroic (*BD* in left panel of Fig. 1-186) has a low transmittance in the spectral region 420 nm to 490 nm and hence that nonabsorbing coating reflects in that spectral region. This light then enters the objective lens.

The red light reflects from the red-green beamdivider *RGD* and is transmitted by the green slab beamdivider *GD*. After reflecting from several turning mirrors (*solid black*), the red light hits the red LCD valve *RV*. The red light finally reflects from the red-reflecting dichroics in the X-prism and is directed to the objective lens.

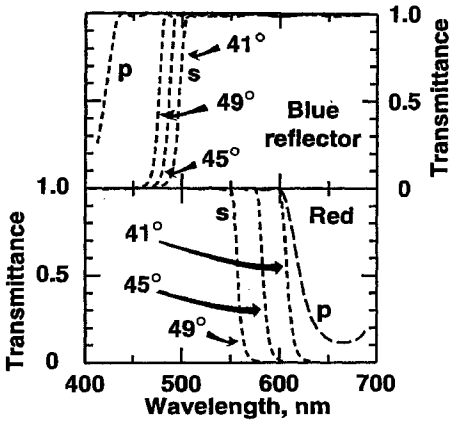


Fig. 1-187 At three angles of incidence measured in the glass of index 1.52, $T_s(\lambda)$ (*short dash*) of the blue-reflecting (*upper*) and red-reflecting (*lower*) dichroic reflectors in the X-prism of Fig. 1-186. $T_p(\lambda)$ (*long dash*) is shown at 49°. The designs appear in §1.10.2.14.

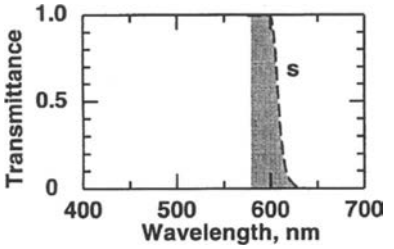


Fig. 1-188 At an angle of 41°, $T_s(\lambda)$ from Fig. 1-187. The shaded area is the spectral region encompassing the nominal edge of the red band (580 nm) and the transmittance curve.

The green light is reflected by the green beamdivider *G D* and propagates towards the green LCD valve. Before it arrives at this valve, however, there is a *halfwave linear retarder* (also termed a *halfwave plate*) that converts the *s* linearly polarized light into *p*. Figure 1-187 shows that the X-prism transmits *p*-polarized green light with high efficiency.

Finally, a question about the X-prism arises: Does the polarization splitting (between *s* and *p*) of the transmittance curves degrade the system performance? In this configuration, the answer is essentially no. The reason is that the red and blue dichroics need only reflect *s* polarized light and only the green *p* polarization should be transmitted. Thus the only requirement for *p* polarization is a relatively high transmittance in the green spectral region — 490 nm to 580 nm.

A final comment is that the system in Fig. 1-186 is simplified for purposes of illustration. Numerous condenser lenses are omitted. There are dye absorption type linear polarizers in each channel. The horizontal physical dimensions are expanded for the purposes of illustration.

1.8.4.4.3 Angle shift of the edge transmission curve of a dichroic

Figure 1-187 shows the transmission curves shifting progressively to shorter wavelengths as the angle of incidence (measured in the glass) shifts from 41° to 49°. Does this angle shift cause the system to degrade? The answer is yes, to a limited extent. For example, the goal of the red dichroic in the X-prism is to reflect *s* polarized light at wavelengths in the spectral region 580 nm to 680 nm. The shaded area under the $T_s(\lambda)$ curve in Fig. 1-188 shows that at the 41° incidence angle, the light in the spectral band 580 nm to 610 nm is not reflected by the dichroic.

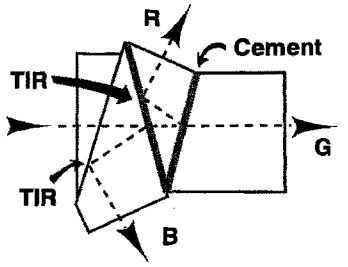


Fig. 1-189 In the Rank modification of the Philips color TV camera prism, path of axial rays (*dashed*) in the blue channel *B*, red *R* and green *G*. After [Cook 75].

1.8.4.4.4 Control of the differential phase shift — first example

The aforementioned Philips system (see §1.8.4.2.1) has been used in a projection system by placing a LCD valve in the focal plane of each channel. In a modification [Cook 75], Fig. 1-189 shows that an additional wedge prism is added to the entrance face, which permits the angle of incidence upon the blue dichroic (that follows) to be 16° for an axial ray. This is compared with the 30° angle for the system in §1.8.4.2.1. It follows from the discussion in §1.8.5.1 that this relatively small angle reduces the polarization splitting.

Figure 1-189 shows that there are air gaps between the prisms that produce total internal reflection (TIR). For meridional rays, the light illuminating the LCD in both the blue and red channels is linearly polarized, with the plane of polarization either parallel or perpendicular to the plane of the paper. The problem is that for skew rays there is a differential phase shift upon reflection for the TIR light. This alters the polarization state of the light to elliptical. A consequence is that it is no longer possible to obtain total extinction of the light reflected from the LCD in the "dark" state.

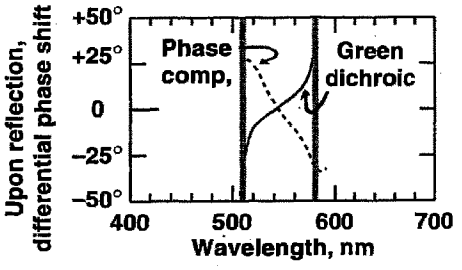


Fig. 1-190 At an angle of incidence of 22.5° in glass, differential phase shift upon reflection (using Abelès sign convention) for the modified dichroic (*solid curve*) and phase compensation coating (*dashed*). The vertical lines delineate 510 nm and 580 nm. Coating designs appear in §1.10.2.15.

The aforementioned differential phase shift upon reflection is graphed in Fig. 2-29 for an uncoated glass-air interface. However, the surface at which TIR occurs is not uncoated. There are antireflection coatings antireflection coating on the air-gap side of the prism that reduce the reflectance of the air-glass interfaces. The light that reflects TIR bounces off of the opposite sides of those antireflection coatings. TIR (i.e., 100% reflection) still occurs as long as the angle of the ray in the glass is in excess of the critical angle (see §2.2.2.2) and the layers are nonabsorbing.

The solution is to modify the antireflection coatings so that they not only function as an antireflection coating but, in addition, produce as close to zero (as possible) differential phase shift upon reflection. It has been the experience of the author that this additional requirement easily doubles or even triples the number of layers in the antireflection coating.

1.8.4.4.5 Control of the differential phase shift — second example

The light in the green channel of the penta prism in Fig. 1-185 reflects from the green dichroic reflector and thence from the *C* (phase compensation) reflector coating on the

other interface. The goal is to minimize the net differential phase shift upon reflection occurring as the light reflects successively from the green dichroic and the *C* coating. This is accomplished by requiring that the differential phase shift upon reflection of the green dichroic and the *C* reflector be approximately equal in magnitude, but opposite in sign — in the spectral region 510 nm to 580 nm, at least.

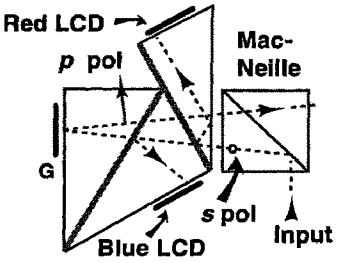


Fig. 1-191 Color projector with MacNeille polarizer, red, blue and green *G* LCD panels. The heavy shaded lines are dichroic reflectors. *s pol* and *p pol* refer to the linear polarization of the light. After Doany [97].

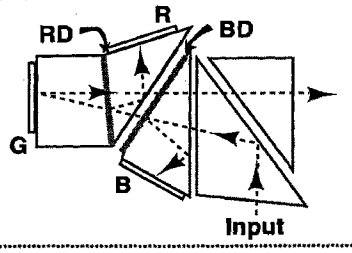


Fig. 1-192 Input light from an illumination source reflects TIR and traverses the red dichroic *RD* and blue dichroic *BD*. *B*, *G* and *R* are the DMD valves in the blue, green and red channel, respectively. After [Hardin 99].

The solid curve in Fig. 1-190 graphs the differential phase shift upon reflection of the green dichroic. The differential phase is approximately –25° at the lower wavelength end of the green channel. It increases to +25° degrees at the upper wavelength end of that channel.

The differential phase shift upon reflection of the *C* reflector is the dashed line in Fig. 1-190. It manifests a differential phase shift upon reflection that is the negative of the differential phase shift of the green dichroic. For example, the differential phase of the *C* coating is approximately +25° degrees at the lower wavelength end of the green channel. The combination of the light reflecting successively from these two coatings achieves the desired phase compensation.

1.8.4.4.6 Liquid crystal display (LCD) projection system

A feature of the Philips beamdivider mentioned in the previous section is that the light in each channel reflects twice, as in the blue and red channels, or not at all, as in the green channel. The pixel order in the image must be reversed electronically if the light were to reflect once. Advances in electronics allowed this reversal to happen in the decade of the 1990s and beyond.

Figure 1-191 shows a projector in which each of the LCD valves functions in the reflection mode. The randomly polarized input beam travels upwards where it reflects from the coating of a MacNeille polarizer (see Fig. 1-168).

The reflected light is *s* polarized — that is, the electric vector is perpendicular to the paper. When the proper voltage is imposed on a pixel in the LCD valve in the green channel, it functions as a "polarization mode converter." The light reflected from the LCD is *p* polarized, which in turn is transmitted by the MacNeille. In other words, this "voltage on" is the "bright" state. If no voltage were imposed, the reflected light remains *s* polarized, little of which is transmitted by the MacNeille. This is the "dark" state.

Similar to the Philips system (see §1.8.4.2.1), the reflection of the light in the blue and red channels from either dichroics and/or internal reflections converts linearly

polarized light into elliptically polarized light, which is partially transmitted by the MacNeille.

1.8.4.4.7 DMD projection system

This system [Hardin 99] uses the DMD light deflectors developed by Texas Instruments. Figure 1-192 shows that the input illumination enters the prism. The angle of the ray exceeds the critical angle and is reflected TIR. The light in the blue spectral region reflects from the blue dichroic and thence travels to the blue DMD *B*.

The remaining incident light penetrates another air gap and impinges upon the red dichroic *RD*. The light in the red spectral region is directed towards the red DMD R. The transmitted light illuminates the green DMD G. As the reflected light from each of the DMD devices travels towards the objective lens, it must traverse two air gaps. The angle of incidence (in air) is as high as 75° and the design of antireflection coatings that function effectively at these large angles is by no means an easy task.

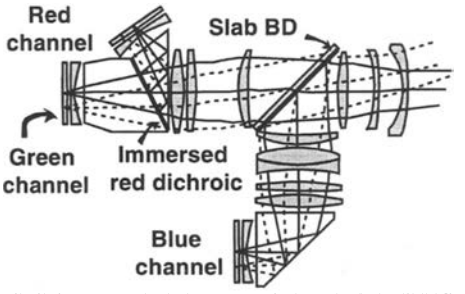


Fig. 1-193 A three-color projector for high definition TV incorporates both an immersed dichroic and a dichroic coated upon a slab [Jeffs 94].

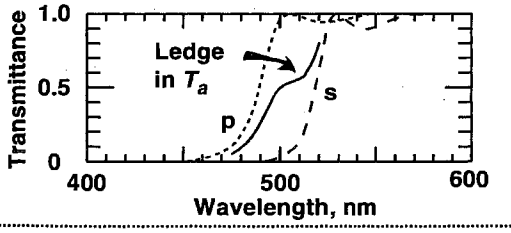


Fig. 1-194 For the blue channel dichroic in Fig. 1-183 and at 30° angle of incidence, $T_a(\lambda)$ (*solid*), $T_p(\lambda)$ (*short dash*) and $T_s(\lambda)$ (*long dash*). Its design appears in §1.10.2.17.

1.8.4.4.8 Projection system with slab beamdivider

Figure 1-193 shows a high definition color television projection system in which the blue dichroic reflector is in quasi-collimated light. This optical arrangement permitted a slab, rather than an immersed, beamdivider to be used. The beamdivider for the red channel is immersed, similar to the MacNeille design.

1.8.5 Miscellaneous topics

1.8.5.1 Polarization splitting of a dichroic

Figure 1-194 graphs $T(\lambda)$ of a coating that could function as a blue-reflecting dichroic in Fig. 1-183. The cuton of T_p is at 470 nm and at 505 nm for T_s. This polarization splitting is welcomed if a linear polarizer (see Fig. 1-172) is constructed. Figure 1-194 shows a manifestation of this splitting — a ledge *L* in the T_a curve. §1.8.5.2 discusses the difficulty of eliminating this ledge.

The ledge in the $T(\lambda)$ curve is less of a problem as long as light impinging upon a camera is randomly polarized. At least the color balance — the relative amount of light in each channel — remains constant. Suppose a camera collects light reflected from a wet surface. That light is highly polarized — the effect is that its color balance is perturbed. Inspection of Fig. 1-194 shows that the 50% *T* point could move 30 nm due to a change in the polarization of the illumination.

1.8.5.2 Nonpolarizing dichroic

Figure 1-195 depicts a dichroic blue reflector filter designed by computer optimization. Notwithstanding its virtue of minimal polarization splitting at the transmission edge, its detractions are: (a) a layer count of 101, which is quite large; and (b) three coating materials are used. At least twenty additional layers must be added, if the blue dichroic reflector is to reflect in the spectral region 410 nm to 480 nm. An entire chapter of Thelen [89] elaborates on the design of edge filters that manifest minimal polarization splitting.

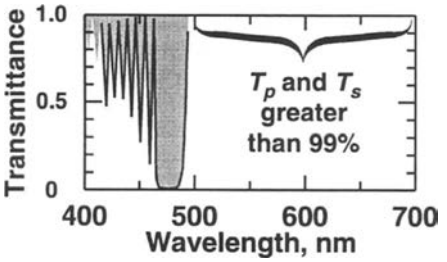


Fig. 1-195 At angle of incidence of 45° in air, $T_s(\lambda)$ (*solid*) and $T_p(\lambda)$ (*shaded*) of an 101 layer dichroic edge filter coated on a slab, whose design appears in §1.10.3.

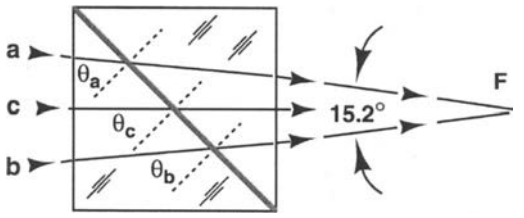


Fig. 1-196 In an *f*/3.8 beam, meridional rays *a* and *c*, as well as the axial ray *b*, converge on the focus *F* through a cemented cube. The "local" incidence angles on the coating (*shaded*) are θ_a, θ_c and θ_b, respectively.

1.8.5.3 Optimal performance over range of angles

Performance of coatings mentioned previously is determined in collimated light. What happens when a cone of light (see Fig. 1-17) impinges on a coating? This is best illustrated by an example.

Figure 1-196 depicts a cross section of a cone (*f*/3.8 beam). Meridional rays impinge upon a cemented cube linear polarizer of the edge filter type (see Fig. 1-172) that transmits *p* polarization and reflects *s*. The coating is optimized at 633 nm and 45°. The center panel of Fig. 1-197 shows that this dichroic functions well at 45°.

Consider rays at other angles. Ray *a* in Fig. 1-196 is inclined at 7.6° to the horizontal axis. After that ray refracts at the external air-glass interface and enters the glass, its angle to the horizontal is 5.0°. The angle of incidence θ_a on the coating is 45° + 5° = 50°. By similar reasoning, the incidence angle θ_b of ray *b* is 45° − 5° = 40°. Figure 1-197 shows that T_p at 633 nm drops to 35% at 40° and that T_s is getting quite close to its long wavelength transmittance edge.

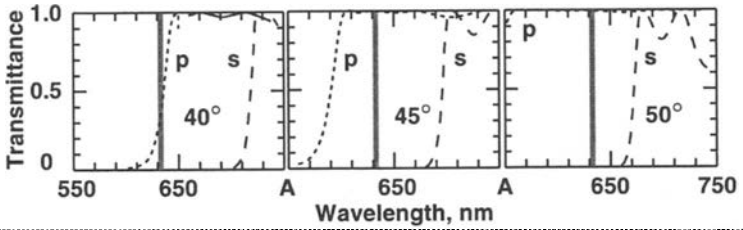


Fig. 1-197 At θ_0 of 40°, 45° and 50°, $T_p(\lambda)$ (*short dash*) and $T_s(\lambda)$ (*long dash*) of a linear polarizer in a cemented cube. The design is listed in §1.10.2.13.

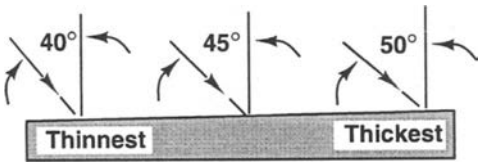


Fig. 1-198 At θ_0 of 40° (*left*), 45° (*middle*) and 50° (*right*), relative thickness of a layer of refractive index 1.50 "tuned" to operate optimally at those angles. The incident medium index is 1.517.

1.8.5.3.1 Wedged thickness layers

Wedged describes a coating in which the thickness of each layer is deliberately adjusted to compensate for varying angle of incidence [Dimmick 52b; Kelly 56]. In a sense, this thickness adjustment is a "tuning" so that each layer performs optimally at a particular angle. Films are thickest where the "local" angle of incidence is largest and thinnest where it is smallest.

Figure 1-198 shows, for example, a layer of index 1.50 "tuned" at three angles. Wedging its thickness makes it 20% thicker at the 50° angle of incidence portion, as compared to the 40° section. Figure 1-199 shows that the performance of the polarizer with wedged thickness layers is dramatically improved, as compared to the transmittances shown in Fig. 1-197.

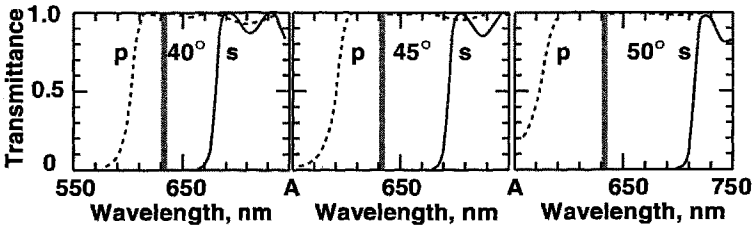


Fig. 1-199 Caption to Fig. 1-197 obtains, with the following exceptions: At 40°, **L** and **H** layers are 8% and 2% thinner, respectively. At 50° **L** and **H** layers are 11% and 2.6% thicker, respectively.

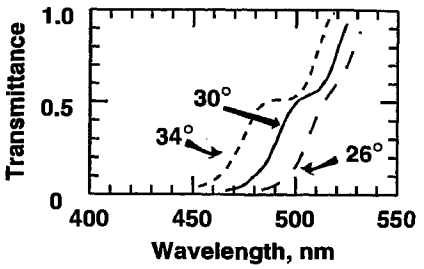


Fig. 1-200 At θ_0 (*inside* the glass) of 34° (*short dash*), 30° (*solid*) and 26° (*long dash*), $T_a(\lambda)$ of coating captioned in Fig. 1-194.

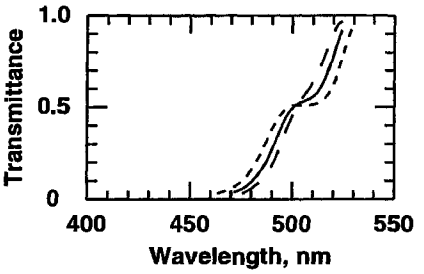


Fig. 1-201 Caption to Fig. 1-200 obtains, with these exceptions: At 26° the **L** and **H** layers are 3.7% and 1.3% thinner, respectively. At 34° the **L** and **H** layers are 4.6% and 1.6% thicker, respectively.

A blue-reflecting dichroic (see Fig. 1-194) serves as another example. Figure 1-200 depicts T_a at incidence angles (inside the glass) of 26°, 30° and 34°. There is a 40 nm spread between the edges at the two extreme angles of incidence. Figure 1-201 shows the same design, with the exception that layer thicknesses are wedged to compensate for the "local" angle of incidence. Although there is an unavoidable "ledge" in $T_a(\lambda)$ due to polarization splitting, the curves at the three angles bunch together much more closely.

Wedging the layers usually increases significantly the cost of coating a slab or a prism face. The reason is that each coated part must be rotated behind a mask to obtain the gradient thickness. This usually reduces the number of parts that may be coated in a given "coating run." An ultimate system uses a different mask for each coating material.

1.8.6 Coatings: costs and specifications

This section considers problems unique to beamdividers, dichroics and polarizers.

1.8.6.1 Costs of a coating

1.8.6.1.1 Nonrecurring costs

In addition to the cost of depositing layers on a prism face or slab, there are expenses unique to the aforementioned coatings:

- Tooling. This is mechanical fixturing that holds prisms or planar parts while they are being coated. Prisms are usually held in plastic pads of polyamide. There is the expense of design and fabrication of the pads and metal fixtures that hold them. The clear aperture should be considered in tooling design (see §1.9.1).
- If required, masking for the production of wedged thickness layers must be designed and installed (see §1.8.5.3.1).
- Fixturing for QA. Suppose that a trapezoidal prism (see Fig. 1-157) is to be coated and that its length is 100 mm. It is too large to fit into the compartment of a spectrophotometer. A prism with the same angles but a 25 mm face is fabricated as a witness for QA. After it is coated with an actual part, it is attached to an identical prism with immersion oil whose refractive index matches that of the optical cement. These "oiled together" prisms are held in a mechanical fixture that positions them in the sample beam of a

spectrophotometer — at the proper angle. Such fixtures and "witness prisms" add to the cost.

1.8.6.1.2 Additional costs

There are additional costs particular to these coatings:

- QA — transmittance in polarized light. Suppose that $T(\lambda)$ of a coated part is to be measured in polarized light. This coated part is a cube polarizer whose $T(\lambda)$ is similar to that in Fig. 1-175.

  A crystal linear polarizer (i.e., Glan-Thompson) is inserted into reference and sample beams of the spectrophotometer — a "100% T line" is established. If there is evidence that physical bulk of the coated part causes a beam to defocus and not converge properly on the detector in the spectrophotometer, a "100% T line" is run with a "dummy cube" of glass (of the same size as the coated part) in each beam. The "dummy cube" is then removed and a spectral scan is now made of the coated part.
- Thermal cycling. In many instances coatings are deposited at temperatures in excess of 250 °C. The larger the bulk of the prism, the more slowly the material is raised from ambient to coating temperature. This also applies to the cool down cycle after the coating has been deposited.
- Cementing and alignment. Exceedingly tight tolerances on parallelism or angles add significantly to the cost of the final assembly. A challenging alignment problem is the X-prism in Fig. 1-183. Four prisms must be orthogonal and/or parallel to each other and pairs of their faces coplanar. The coating extends to the 90° corner of each prism, which is not beveled, nor can it be chipped.

1.8.6.2 Cost reduction methods

A reduction in manufacturing cost is possible if the following points are considered:

- Angle of incidence — is it possible to adjust or alter it? See §1.8.6.2.1.
- In an immersed coating, is it feasible to coat a thin slab that is subsequently cemented between prisms? See §1.8.6.2.2.
- Always bevel edges of prisms. Prisms with nonbeveled edges chip easily.

1.8.6.2.1 Angle of incidence

Consider an immersed linear polarizer that should manifest T_p close to 100% over the visible spectral region and T_s below 0.5%. The s-polarized reflected light is collected in a beam dump and thus is not used. The glass must be BK-7 because of cost considerations.

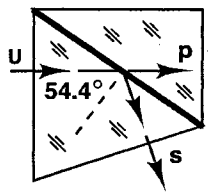


Fig. 1-202 Ray of randomly polarized light *U* impinges upon MacNeille polarizer coating (*black*) at θ_0 of 54.4°. Transmitted light is *p* polarized.

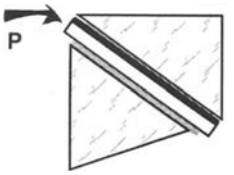


Fig. 1-203 The polarizer in Fig. 1-202 is constructed by depositing coating (*solid black*) on plane parallel plate *P* and encapsulating it between prisms with optical cement (*shaded*). Thicknesses are exaggerated for purposes of illustration.

There are two approaches to this problem: (a) A customer mandates an angle of incidence of 45°, as in Fig. 1-168. (b) A customer is willing to modify the internal angle of the prism such that optimum coating performance is obtained.

Assume that the first approach is taken. A problem is that this MacNeille polarizer must be constructed using magnesium fluoride and a material of index 1.705. The latter is difficult to deposit with a reliable refractive index. A result is that the coating is expensive because many coating houses would choose to no-bid it.

Is it possible that a more enlightened customer is willing to adjust the internal angle inside the prism to 54.4°, as shown in Fig. 1-202? Why this angle? Vendors advise that if this angle is used, the multilayer will be composed of alternate layers of titania and silica — coating materials that are normally used and well understood. A coating with superior performance is a result.

A final example relates to the production of a slab linear polarizer for a Nd:YAG laser, as shown in Fig. 1-172. There are two approaches: (a) The customer mandates that at incidence angle of 56.4° and wavelength of 1.056 μm, T_p must exceed 99.5% and T_s not exceed 0.1%. (b) The foregoing transmittance requirements are to be met, with an exception that the incidence angle can range from 55° to 58°. In the latter approach, the customer is willing to "angle tune" each slab polarizer when it is installed in an instrument.

An advantage to the second approach is that coating yields can be higher and consequently be produced at lower cost. At one time the author's employer — Coherent Medical Division — designed instruments with a multilayer linear polarizer that included this flexibility in its angle of incidence.

1.8.6.2.2 Prism — coating of a slab

Suppose the linear polarizer in Fig. 1-202 has a face of dimension 150 mm x 150 mm. That prism is by no means easy to handle and install in a coating chamber — it weighs nearly 6 kg. An alternative to coating a prism is to coat a slab with plane parallel surfaces. This slab is subsequently cemented into the prism, as shown in Fig. 1-203.

1.8.6.3 Specification of beamdividers, dichroics and polarizers

Items in the following list might appear in a specification:

- Type of glass or material for incident medium and/or substrate. This includes optical cements.
- Range of incidence angles. Are there tolerances on those angles?
- Should the vendor assume that the incident light is randomly polarized? If not, can the customer specify its state of polarization?

- Is it necessary to wedge layer thicknesses to improve performance?
- What is a lower limit for the *RT* product? [See §1.8.1.2(g).]
- What is the type of illumination? Is it "coherent" (as from a laser) or "white light" (as seen by the human eye)?
- Clear aperture (see §1.9.1).
- Are there laser damage specifications?
- Does the opposite side of a slab or faces of a prism require an antireflection coating?

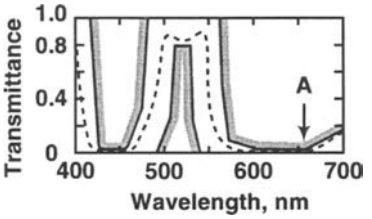


Fig. 1-204 Upper and lower limits of $T(\lambda)$ of a dichroic beamdivider are indicated by a stick diagram. At *A* the coating is out of specification.

1.8.6.3.1 Specification of a dichroic

If a $T(\lambda)$ curve is relatively complex, its upper and lower bounds are displayed, as in Fig. 1-204. This is a green channel trim filter for a color printer. $T(\lambda)$ (*dashed*) should not trespass upon shaded regions, as it does at point *A* in the Fig.

1.8.6.3.2 General observations

Comments below are the result of attempts (and failures) by the author to design the "supercoating."

- The author has yet to cast his eyes upon an all-dielectric beamdivider centered at λ_0 that is nonpolarizing over a spectral bandwidth that exceeds 10% of λ_0.
- If an immersed beamdivider is required that functions over a wide range of angles and broad spectral bandwidth in the visible, a coating with silver (see §1.8.2.2.1) is the best choice. Yes, its *RT* product suffers.

1.8.6.3.3 Rules for the procurement

Some general rules are:

- If the polarization effects of the beamdivider are to be minimized, avoid a 45° angle of incidence. Use the smallest angle possible. Although this rule applies to both slab and immersed beamdividers, it is especially true of the latter.
- Remember that it quite difficult to reflect *p* polarized light in an immersed system, especially when the angle of incidence exceeds 40°.
- When BK-7 glass is used, a MacNeille polarizer functions much more efficiently when the angle of incidence is in the range 50° to 55°.
- If the layer thicknesses are not to be wedged, then place the immersed dichroic and/or MacNeille polarizer in telecentric image space.

1.9 Miscellaneous topics

1.9.1 Clear aperture, jig marks and bevels

An assumption is that physical vapor deposition is used to apply a coating to a substrate. Although it is possible to evaporate downwards so that an entire surface is coated, this is expensive and rarely done. It is easier to sputter downwards — one problem is that dust particles settle more easily on upward-facing substrates.

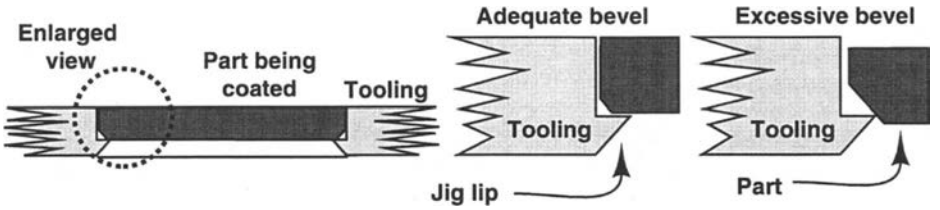


Fig. 1-205 Beveled glass part (*darker shading*) is held in metal tooling (*lighter shading*) during deposition. (*Right*) Thermal expansion permits part with wide bevel to ride on jig lip.

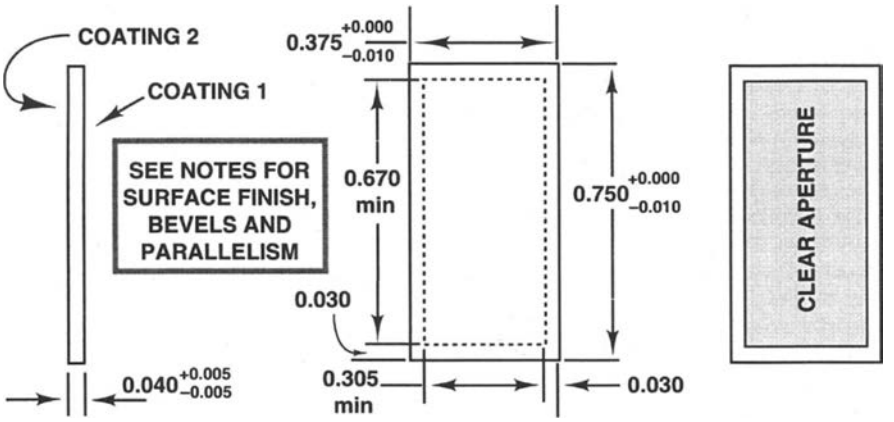


Fig. 1-206 Drawing of part coated *only* in its clear aperture. A shop note, which is a separate document, specifies the substrate material, bevels, surface finish, parallelism, specifications of coatings and QA procedures.

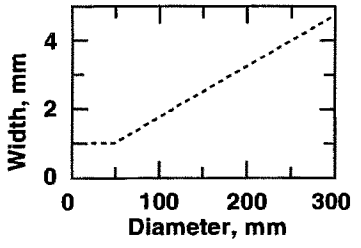


Fig. 1-207 Suggested minimum width of a jig mark of coated part vs diameter (or diagonal). After MIL-F-48616.

Figure 1-205 shows a cross section of tooling that holds a part (i.e., substrate) as it is coated. A jig lip supports the part and also defines the limits of the *clear aperture* — the coated area. This is usually defined in a drawing, as depicted in Fig. 1-206. The uncoated area is called the *jig mark*.

Consider the effects of heating the parts and tooling during deposition. In most instances thermal expansion of metal tooling exceeds that of glass. The gap between the tooling and edge of the part increases with temperature. Figure 1-205 shows a part with an excessively wide bevel that is physically supported on a relatively slim jig lip. Thermal expansion permits a part to ride on the protrusion of the lip. It is certain to chip when the tooling is cooled and contracts.

1.9.1.2 Jig mark width

How wide (or narrow) should the jig mark be? Figure 1-207 presents a conservative width for a jig mark vs diameter of the part. For larger sizes, the jig mark width is approximately 2% of the part diameter.

Negotiations between customer and vendor can reduce the jig lip width cited above. An example is that the jig lip dimension in Fig. 1-206 is 0.030 inch, whereas Fig. 1-207 mandates 1 mm, which is 0.040 inch.

A jig mark should be sufficiently wide that a lens or optical element mounts properly. Figure 1-208 shows, for example, a bezel in a lens mount that abuts the coating, rather than the uncoated lens. This displaces the lens element from its proper position.

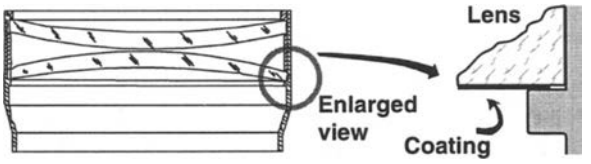


Fig. 1-208 Lens in a mount abuts a bezel. The thickness of a coating displaces the lens from its proper position. After Matteucci [80].

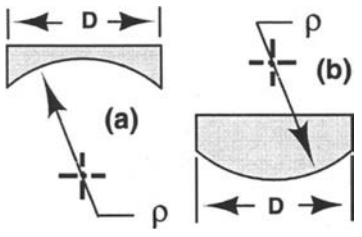


Fig. 1-209 Radius ρ and diameter *D* of concave (*left*) and convex (*right*) parts, undersides of which are to be coated.

1.9.1.3 Tooling

The cost of coating a part depends upon the number of parts that can be loaded in the coating tank at one time. As an illustration, a coating tank might be able to coat 95 *flat* parts of 13 mm ϕ in a coating run. If surfaces are highly curved, however, this number may be greatly reduced because each part deserves "special treatment." A manifestation of this "special treatment" is that during coating each *individual* part must be rotated in front of a mask to achieve thickness uniformity. Under these circumstances, it might be possible to coat only 11 parts at one time, rather than 95. This increases the cost.

Whether a part deserves this "special treatment" depends upon the ratio of the radius ρ of the curved surface to its diameter *D* (see Fig. 1-209). Rules of thumb are:

> If $\rho/D > 1$ for concave surfaces or
> if $\rho/D > 1.2$ for convex surfaces,
> no "special treatment" is needed.

In other words, if either condition above is met, parts with curved surfaces are treated the same way as plano parts — for the purposes of coating.

1.9.2 Cementing of coatings

One reason a coating is cemented is to protect it from the environment, such as high humidity or salt spray. Cements also join prisms used with immersed coatings.

A vast number of cements are available. It is possible that silver chloride could be used as a "cement" in the IR. Most cements are organic, however. Pellicori [64, 66] examines UV transmitting cements. A cement eventually absorbs in the UV due to electronic transitions, as illustrated by the $T(\lambda)$ of RTV in Fig. 1-210 [Williams 73]. Stretching and bending of molecular bonds produce absorption in the IR.

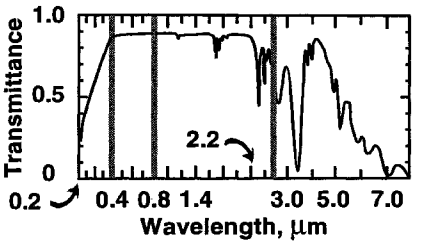


Fig. 1-210 $T(\lambda)$ of RTV cement. Abscissa changes scale at 0.4 μm, 0.8 μm and 2.5 μm. After Williams [73].

A cement should be compatible with the coating. A UV setting cement is often used in a coating containing silver, because its chemicals do not degrade the metal. There are often chemical constituents in epoxy cements that attack silver. Penetration of the optical cement into a coating invariably alters its refractive index which, in turn, causes its $T(\lambda)$ to shift.

The way a coating is cemented is important. For example, its edge should not be exposed to humidity, which would permit water to penetrate the edge of each layer, thus causing the multilayer eventually to degrade. This moisture edge penetration is lessened by removing a strip of coating on its periphery and filling that space with optical cement, as shown in Fig. 1-211. Virtually all organic cements permit a modicum of water vapor to diffuse through them, however. Metals are much more impervious to water vapor. The right panel of Fig. 1-211 shows that the coating is sealed by a layer of metal deposited on its periphery. A similar process is described by Miles [73].

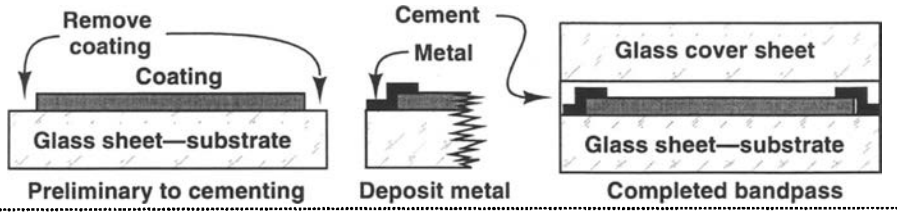


Fig. 1-211 Prior to cementing, a strip is removed from periphery of a coating. That outer border is overcoated with metal. The coating is then cemented.

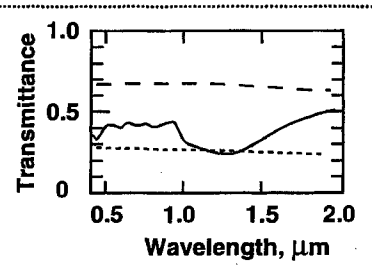


Fig. 1-212 (*Dashed*) $T(\lambda)$ of thin film neutral density coatings. After Hampton [90]. (*Solid*) $T(\lambda)$ of "ND-40" glass, after Hoya [88].

1.9.2.1 Optical contacting

Optical contacting is a method of joining two surfaces by the molecular forces that occur when they are in intimate contact [Karow 93]. Austin [73a] optically contacted a bandpass and Lees [79b] contacted prisms to form an IR polarizer.

1.9.3 Neutral density

Neutral implies that $T(\lambda)$ of a neutral density (ND) is relatively constant with wavelength. A neutral density is used to adjust the illuminance of a light beam, for example, in an automated photographic printer. Suppose a laser illuminates a surface and scattered flux is collected with a photomultiplier. The level of scattered flux is 10^{-8} below input beam and since the detector is not linear over that range, the only recourse is to insert neutral density coatings. Most neutral densities are absorbing because they contain one or more metal layers.

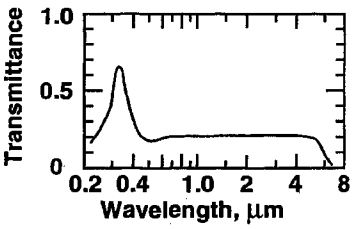


Fig. 1-213 $T(\lambda)$ of *air* **D M D M D M D** *sapphire* where **D** and **M** represent layers of BaF_2 (20 nm thick) and silver (5 nm thick), respectively. After Flint [74].

1.9.3.1 Cascading of neutral density coatings

If the neutral density coatings manifest no appreciable reflectance, the net transmittance T_0 of three NDs in tandem is

$$T_0 = T_1\, T_2\, T_3\,, \qquad (1\text{-}21)$$

where T_1, T_2 and T_3 are transmittances of individual filters. If the net absorbance is D_0 [see Eq. (1-6a)], the foregoing reduces to

$$D_0 = D_1 + D_2 + D_3\,. \qquad (1\text{-}22)$$

Thus the absorbances add. Equation (1-22) may be generalized to an array of m neutral density coatings.

1.9.3.2 Advantages of neutral density coatings

An alternative to a neutral density with metal layers is absorbing glass. Figure 1-212 shows that although its $T(\lambda)$ is relatively flat in the visible part of the spectrum, it dips appreciably in the near IR. In contrast, a neutral density prepared by vacuum evaporation is usually quite flat. Figure 1-213 graphs $T(\lambda)$ of a neutral density for the IR. A drop in T at 7 μm is due to absorption of the substrate. If deposited on a proper substrate, this neutral density should function well at wavelengths as long as 15 μm.

1.9.4 Absorbers — both selective and broadband

Absorbers are coated onto baffles, mirror mounts and other hardware in an optical system to reduce reflectance and, hence, scattered light. In the visible part of the spectrum, the tube interior and support elements of an inexpensive Gregorian telescope are coated with black lacquer. Why is a more expensive evaporated coating used in lieu of lacquer? As shown in Fig. 1-214, it is not unusual that its reflectance is well below that of any paint.

Another factor is that evaporated coatings are inorganic, which excludes any possibility that outgasing of the coating could contaminate an optical system, as could occur with a lacquer.

There are many ways of constructing absorbers. Dupoisot [74] deposited a two-layer metal-dielectric coating that absorbed 633 nm. Dimmick [53] used a mixture of magnesium fluoride and copper as an absorber.

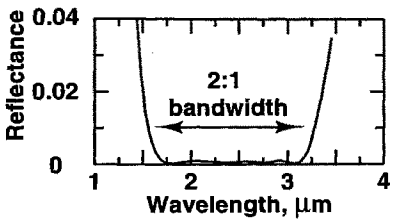


Fig. 1-214 $R(\lambda)$ of a dark mirror containing ten layers. Its design appears in §1.10.2.12.

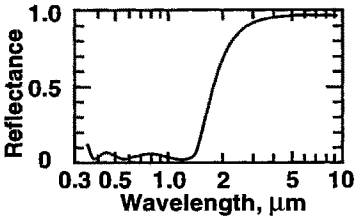


Fig. 1-215 $R(\lambda)$ of dark mirror. Its design appears in Fig. 1-216. After Hass [55].

1.9.4.1 Selective absorbers

Figure 1-215 shows $R(\lambda)$ of a coating often called a *dark mirror* [Hass 55]. Its reflectance is below 5% in the visible and near IR, yet it reflects in excess of 90% at 4 μm. Figure 1-216 shows that this coating consists of opaque aluminum overcoated with a germanium film and two silicon monoxide layers. The absorption of the germanium masks the reflective aluminum film in the visible spectrum. A low reflectance is produced. The transparency of the germanium in the IR, however, permits the IR to reflect from the aluminum.



Fig. 1-216 Design of a dark mirror. The shaded layer is germanium and the other layers are silicon monoxide. After Hass [55].

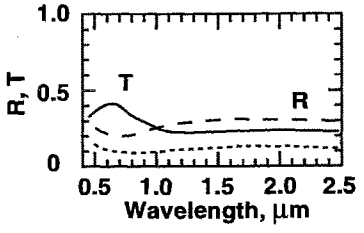


Fig. 1-217 $T(\lambda)$ (*solid*), $R(\lambda)$ from air (*long dash*) and $R'(\lambda)$ from glass side (*short dash*) of an architectural coating for glass. After BOC-Airco [93].

The aforementioned coating has been used as a selective solar absorber. A selective coating that withstands high temperature is used in a solar energy collector [Agnihotri 81]. Similar coatings control the thermal balance of a space vehicle [Drummeter 64]. They are also utilized for visible or IR camouflage.

1.9.5 Coatings for glass on buildings

There are photographs of the World Trade Center building in New York City before its demise in 2001. The exterior was sheathed with glass coated with thin layers. As Berning [83] discusses in his review, there are several types of coatings with diverse purposes classed as "summer films" or "winter films."

1.9.5.1 Summer film

A function of this coating is to reduce transmittance so that intruders outside the building cannot peer into it. The most important mission, however, is to minimize the amount of solar energy leaking into the building. Such a "summer film" is applied to glass on buildings

in tropical and subtropical regions. Figure 1-217 shows $T(\lambda)$ and $R(\lambda)$ of such a coating. This 35% T in the visible is merely an example. There are other coatings that transmit 10%.

1.9.5.2 Winter film

This coating has two functions: (1) It permits the maximum amount of solar energy to enter a building through the glazing; and (2) it minimizes radiant heat flow from the room interior to the cold glass at the exterior of the window.

Figure 1-218 depicts a cross section of a double-glazed window. The temperature drop across an exterior sheet of glass is only a few degrees. As an illustration, if the exterior air temperature is 3 °C, the inner surface of an exterior sheet of glass is at 5 °C. There is heat transfer to that relatively cold exterior sheet by thermal convection and also by radiation. This radiant heat transfer is quite efficient because of the high emissivity ($\varepsilon = 83\%$) of glass. The heat transfer by radiation is reduced by covering the outer surface of the inner glass sheet with a low emissivity coating, as shown in Fig. 1-218.

Figure 1-219 depicts $T(\lambda)$ and $R(\lambda)$ of a winter film. The high T in the visible part of the spectrum is important because of the concentration of the sun's energy in that region. Its reflectance climbs from low values in the visible to 85% at 2 μm. From Kirchhoff's Law one can conclude that its emissivity ε must be lower than 20% at 2 μm. ε is even lower at longer wavelengths, where the peak of a Planckian curve occurs for a 300 K black body.

A final question is, how should the low-emissivity coated glass in Fig. 1-218 be oriented? The answer is that if it is truly a winter film, the coated glass should be contiguous to the inside of the building. This minimizes the heat loss and is desirable in cold climates, such as northern Europe or the upper regions of Canada.

In many more southerly climates, however, it is more important that the coating reject the nonvisible solar radiation, which heats the building in the summer. In this instance, the glass with the coating is oriented so that it is the outside light.

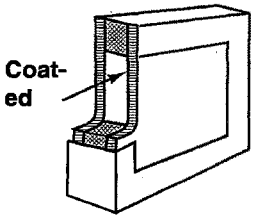


Fig. 1-218 Cross section of a double glazed window. The glass sheet adjacent to building interior is coated to lower its emissivity. After Franz [74].

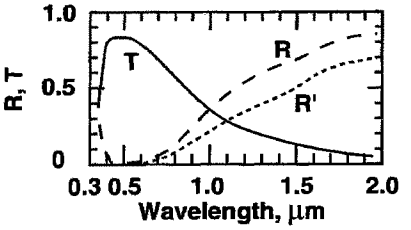


Fig. 1-219 $T(\lambda)$ (*solid*), $R(\lambda)$ from air side (*long dash*) and $R'(\lambda)$ from glass side (*short dash*) of a "winter film" coated on glass. After Airco [93].

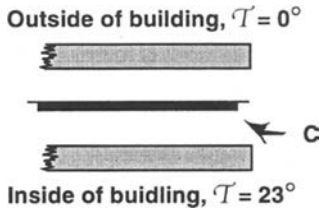


Fig. 1-220 Cross section of double glazed window in which there is a thin plastic sheet with a coating *C* on one side. Temperatures $\mathcal{T}$ are shown.

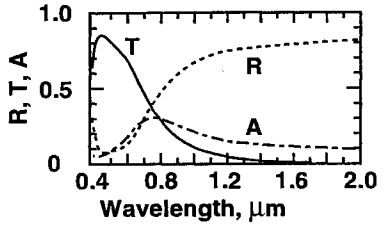


Fig. 1-221 $T(\lambda)$ (*dashed*), $R(\lambda)$ and absorption $A(\lambda)$ of a "winter film" coated on plastic sheet. After Southwall [90].

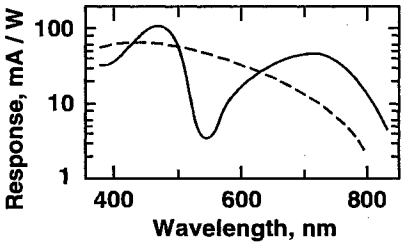


Fig. 1-222 (*Solid*) Responsivity (A/W) of a photocathode of the design
vacuum NaKCsSb SiO *aluminum*
where the *aluminum* is opaque. (*Dashed*) SiO layer is omitted. After Novice [66].

Figure 1-220 depicts another approach to energy conservation, in which a thin plastic film is stretched between glass sheets. Its coated side (i.e., winter film) faces inward. $T(\lambda)$ and $R(\lambda)$ of such a coating appear in Fig. 1-221.

1.9.6 Interference photocathode

As mentioned in a review of Kossel [69], this coating enhances absorption of light in a photocathode and thus increases its quantum efficiency. Delineated in Fig. 1-222 is responsivity vs wavelength (*dashed*) of a conventional opaque photocathode in which a cesium-antimonide layer is deposited directly upon opaque aluminum. In an enhanced design (*solid*), a layer of evaporated SiO is sandwiched between the aluminum and photocathode. This increases its quantum efficiency five-fold at 700 nm.

In an "end on" photocathode configuration appearing in Fig. 1-223, light impinges on one side, is absorbed and creates photo-electrons that emerge from the opposite side. The configuration in Fig. 1-223 is immersed (see §1.8.1.1) — a prism is cemented onto the glass of the tube [Hirschfeld 68; Livingston 66]. A coating (see Fig. 1-224) enhances the absorption at the outer surface of the photocathode, where photo-electrons escape more easily into the vacuum [Love 68; Novice 67].

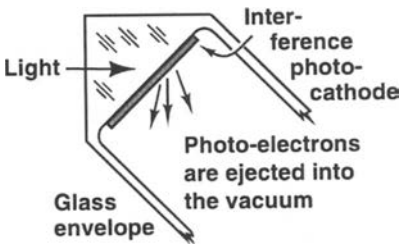


Fig. 1-223 Semitransparent photocathode in which attenutated total reflection enhances its responsivity.

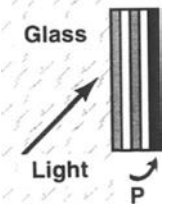


Fig. 1-224 Details of the interference photocathode in Fig. 1-223. *P (solid black)* is the photocathode and other layers are dielectric.

1.9.7 Electrically conducting coatings

Coatings that are both optically transparent and electrically conducting are widely used in liquid crystal displays and for electromagnetic (EMI) shielding.

Electrical properties of a coating are usually measured in ohm/square, which is determined as follows. Electrical resistance is measured by selecting a square of material of width w and attaching ohmic contacts to both ends, as shown in Fig. 1-225. Resistance $\mathcal{R}$ of a film of thickness h is

$$\mathcal{R} \equiv \rho \frac{w}{w\,h} = \rho / h \quad (1\text{-}23)$$

and is thus independent of w, provided its specific resistivity ρ remains constant with thickness. $\mathcal{R}$ is in units of ohm per square.

Figure 1-226 shows that semitransparent gold layers on polyester attain a $\mathcal{R}$ of 11 Ω/square and transmit over 60% in the visible spectrum. It follows from Eq. (1-23) that sheet resistance is inversely proportional to layer thickness.

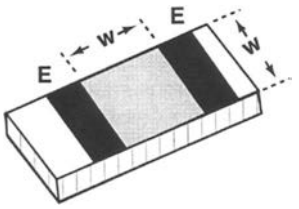


Fig. 1-225 In order to measure its sheet resistance, conducting electrodes *E (solid black)* are deposited so that a square area remains of an electrically conducting film *(shaded)*, whose width and breadth are *w*.

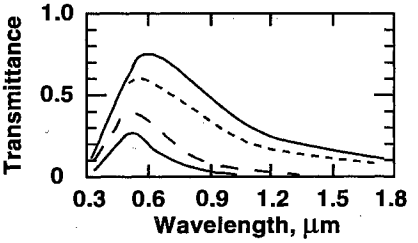


Fig. 1-226 $T(\lambda)$ of gold films on polyester. Sheet resistance 11 Ω/square with *(upper, solid)* and without *(short dash)* "antireflection layers." Sheet resistance 2.5 Ω/square with *(long dash)* and without *(lower, solid)* "antireflection layers." After Andus [90].

A thicker layer, however, also has a lower T. Figure 1-226 shows a 2.5 Ω/square coating with a T of 40%. Coatings with higher T in Fig. 1-226 include "antireflection layers" that enhance T. Figure 1-227 shows a relationship between luminous T and sheet resistance. Figure 1-228 depicts $T(\lambda)$ of a gold layer in the visible spectral region.

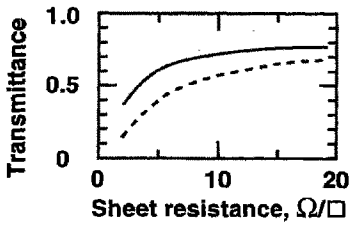


Fig. 1-227 "Visible T" vs sheet resistance of gold deposited on polyester with (*solid*) and without (*dashed*) "antireflection layers." After Andus [90].

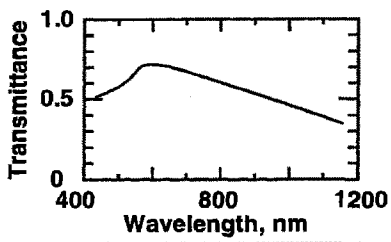


Fig. 1-228 $T(\lambda)$ of gold film on soda-lime glass. Sheet resistance is 15 Ω/square. After Denton [90].

Table 1-14
Attributes of electrically conductive coatings

	Type of coating		
Attribute	Gold	Tin oxide	Indium oxide
Luminous T	75%	82%	85%
ohm/square	20	100	75

Luminous transmittance and sheet resistance (ohm/square) of electrically conductive coatings deposited on soda-lime glass. After Practical [78].

Semitransparent electrically conductive coatings are also made from oxide semiconductors, such as cadmium stannate and zinc oxide. Coatings consisting of a mixture of tin oxide and indium oxide are widely used. Part of the black magic in their production depends upon whether the coating can be baked after deposition, which lowers its sheet resistance. As compared to a glass substrate, sheet resistivity of a coating on plastic is higher because it cannot be baked. Vossen [77] reviews physical and chemical properties of electrically conducting films.

Figure 1-229 depicts $T(\lambda)$ of gold and indium-tin oxide coatings [Practical 78]. Table 1-14 compares their attributes. Figure 1-230 shows $R(\lambda)$ and $T(\lambda)$ of an indium-tin oxide coating with antireflection layers that enhance its visible transmittance.

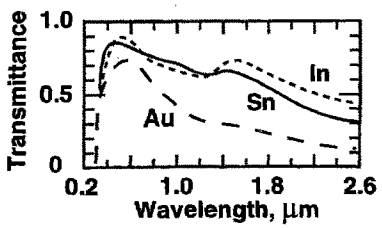


Fig. 1-229 $T(\lambda)$ of soda-lime glass coated with: (*long dash*) gold with overcoat of silicon monoxide; (*solid*) 300 nm of tin oxide doped with antimony; (*short dash*) 200 nm indium oxide doped with tin oxide. After Practical [78].

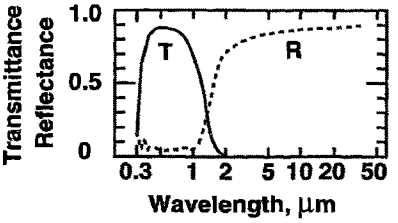


Fig. 1-230 $T(\lambda)$ and $R(\lambda)$ of soda-lime glass coated with indium tin oxide and overcoated with magnesium fluoride. After Granqvist [81].

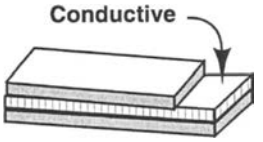


Fig. 1-231 Exposed ledge allows electrical contact with a conductive layer sandwiched between nonconductive dielectric layers (*shaded*).

An electrically conductive coating used for EMI shielding is ineffective unless it is connected electrically to ground. This means that if the outer layer of a stack is not electrically conductive, the producer of the coating must mask a small area after the conductive layer has been deposited. Figure 1-231 depicts an electrically conductive layer sandwiched between two dielectric films. Masking leaves bare a strip on the electrically conductive layer to which a grounding bus bar may be connected.

1.9.8 Lateral variation of *R* or *T*.

In most coatings, every effort is made to ensure that their R or T is relatively uniform at various lateral positions across their surfaces. There are coatings in which a lateral variation of R or T serves a useful purpose. Etched and patterned coatings have unique applications.

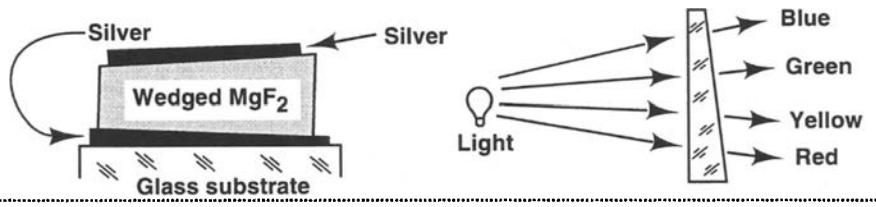


Fig. 1-232 (*Left*) An MDM bandpass with wedged thickness layers (thicknesses are exaggerated for purposes of illustration). (*Right*) Used as monochromator.

1.9.8.1 Bandpass with wedged thickness layers

Figure 1-232 shows a MDM bandpass with a wedged thickness spacer layer. Its passband T peaks at λ_0 whenever the optical thickness of the spacer is approximately $\lambda_0/2$. *Approximately* is used because its thickness is modified by the phase shift upon reflection of silver. At the thin end, its thickness is a halfwave for violet light and hence the peak of the passband occurs at 400 nm. At other locations, the spacer is thicker; these sections transmit progressively green, yellow and red. MDM filters often use a full wave thickness spacer in order to achieve a narrower passband width. A width of 10 nm is typical. These coatings have been used in conjunction with Charge-Coupled Detector (CCD) arrays.

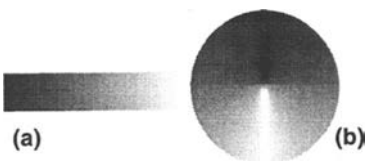


Fig. 1-233 In gradient transmission coatings, T varies linearly (*a*) and as the polar angle (*b*).

The MDM is used as a monochromator by positioning it behind a 1 mm slit. Either the bandpass or slit is moved to scan wavelength. This wedge has also been produced on an annular ring so that thickness varies as a function of the polar coordinate ϕ. These are termed *circular variable filters* [Apfel 65, Thelen 65b], *universal filters* [Holm 68] and *variable wavelength filters* [Ward 71]. Rotation of a filter past a stationary radial slit scans wavelength. IR wedge filters

transmitting at wavelengths as long as 14 μm have been incorporated into monochromators [Linsteadt 65, 67].

1.9.8.2 Gradient attenuators

These are used in a variety of applications. The maximum allowed power output from a He-Ne 633 nm laser is 0.5 mW. The 1.5 mW a newly manufactured tube outputs is attenuated by a gradient attenuator (see *a* in Fig. 1-233) driven into the beam with a servomechanism. Its transmittance is increased as the tube ages and power decreases.

Figure 1-233 also depicts an attenuator whose T is adjusted by rotation. Other attenuators have a "bulls-eye" or "Gaussian" radial change of T from the center and are used to apodize a laser beam.

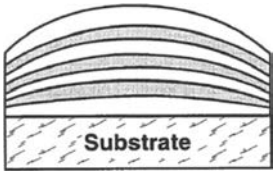


Fig. 1-234 Metric thicknesses of layers of magnesium fluoride (*shaded*) and yttria (*nonshaded*) in a gradient reflectance mirror. After Piegari [93].

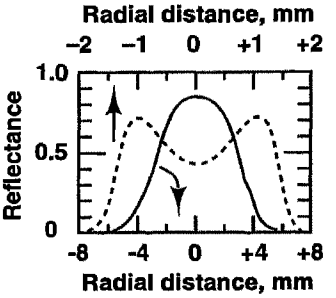


Fig. 1-235 R vs radial distance from center of gradient reflectance mirrors (*dashed* — upper scale and *solid* — lower scale). After Institut [92].

1.9.8.3 Gradient reflectors

R varies in a lateral direction in this coating. It is typically used as an internal mirror in a laser cavity. Its variable reflectance transforms an output beam whose spatial distribution is "ugly" into the Gaussian of the TE_{00} mode. Figure 1-234 depicts, for example, relative thicknesses of layers (which are magnesium fluoride and yttria) of a reflector for an excimer laser [Piegari 93]. Phase shift upon reflection, as well as radiant reflectance, of this coating is controlled. Figure 1-235 depicts R vs radius of several gradient reflectors [Institut 92].

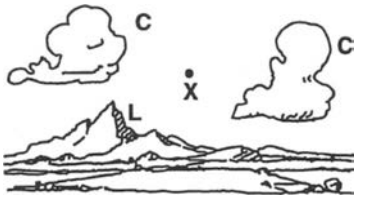


Fig. 1-236 Showing landscape L with sunlit clouds C and target X for heat-seeking missile. Blow them up better! After Whitney [64].

1.9.8.4 Patterned coatings

A patterned coating that many of us look at daily is the liquid crystal display of a wristwatch or computer screen. A reticle in a magnifier, microscope or fire control system is another example.

There are two methods of patterning: (1) Coating is etched after it is deposited; or (2) coating is deposited on top of patterned photoresist. The first method works well if only

one layer (such as a metal) is deposited. If a multilayer contains two or more materials, however, it is not always easy to develop an etching process that erodes away all coating materials with equal facility.

Figure 1-236 depicts a landscape into which is launched a heat-seeking missile that is intended to lock onto water vapor emission of target *X*. Sunlight glinting from clouds or from snow (on a mountain) are false targets. Target discrimination is provided by a "two-color" patterned reticle. The scanning reticle appearing on the left of Fig. 1-237 contains two bandpasses that are alternately swept in front of a detector. Wavelength centers of bandpasses and electrical gain are adjusted to produce a zero response from sunlight.

The right of Fig. 1-237 illustrates construction of a reticle. Opaque strips (often dark mirrors) are laid down between bandpasses to prevent any leakage of radiation. By successive applications of photoresist, coating deposition and lift-off, dark mirrors and bandpasses are sequentially deposited.

1.9.9 Unusual coatings and systems

This section includes "unusual" coatings and systems that use such coatings. "Unusual" merely means that they will probably not appear in a catalog as an off-the-shelf item.

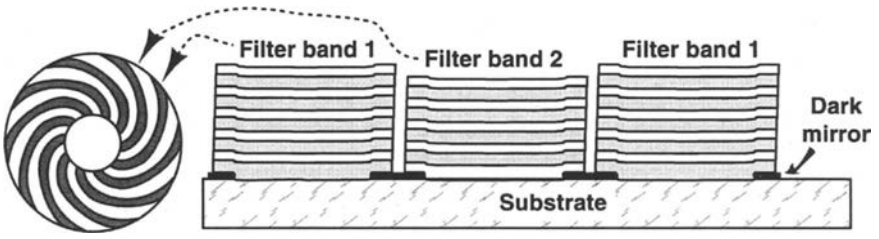


Fig. 1-237 (*Left*) "Two-color reticle" that spins about its center. (*Right*) Cross section of the reticle, showing bandpasses and dark mirrors *D*. The vertical dimensions are exaggerated for purposes of illustration. After Whitney [64].

1.9.9.1 Reststrahlen coatings

Reststrahlen describes the strong reflectance that occurs when IR flux (typically at λ = 5 μm and longer wavelengths) excites lattice vibrations in a crystalline solid. IR filters have been designed that utilize both optical interference effects and intrinsic properties of such materials [Turner 65].

1.9.9.2 Buried diffraction grating

Figure 1-238 shows that this device is a diffraction grating "substrate" overcoated with an all-dielectric multilayer that reflects strongly at λ_0 [Bostick 78]. §1.7.6 emphasizes that this type of reflector has a limited bandwidth over which it manifests high reflectance. Thus at longer wavelengths (such as 1.25 λ_0) or shorter wavelengths (such as 0.80 λ_0) the multilayer is no longer reflecting — the incident radiation penetrates the multilayer and hits the grating.

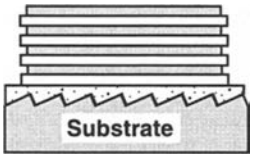


Fig. 1-238 Diffraction grating (saw teeth next to substrate) is overcoated with a smooth layer (*dotted*) and a multilayer reflector. After Bostick [78].

The grating diffracts that radiation. In other words, this system acts like a reflective dichroic beamdivider in the offband spectral region and a high reflector (see §1.7.1.2) at λ_0.

The first step in the manufacture of a "buried grating" is to deposit the first layer on the grooves of the grating. That layer is subsequently polished smooth. The remainder of the multilayer is then deposited on top of that polished layer.

1.9.9.3 Reflector and beamdivider

It is an understatement that many coatings manufactured in the 1990s were considered impossible to produce two decades earlier. Coating houses have developed both their design capabilities and control of their processes. As an example of a more difficult task, Fig. 1-239 depicts $T_p(\lambda)$ of a coating that, at 45°, reflects the lines of an argon ion laser and splits the *p*-polarized light into two beams in the shaded spectral region. According to the producer [Lys 93], theoretical is dashed and measured is solid. It is not difficult to produce the reflector, but the beamdivider part of the coating contains many relatively thin layers whose thicknesses are more difficult to control.

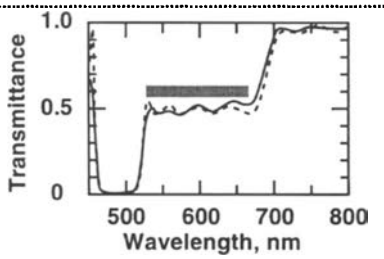


Fig. 1-239 At 45° incidence angle, $T_p(\lambda)$ theoretical (*dashed*) and measured (*solid*) of coating that reflects emission lines of an argon ion laser and divides the light into two beams in the shaded spectral region. After Lys [93].

1.9.10 Coatings for fiber optic communication systems

1.9.10.1 A fiber communication system with an amplifier

Regardless of whether a communication signal is carried by a copper wire or a glass fiber (in many instances at an infrared wavelength near 1550 nm), the signal is attenuated and eventually must be amplified. An earlier method was to convert the infrared from the fiber into an electrical signal, which was then amplified. An injection laser subsequently converted that electrical signal into infrared, which was fed into the output glass fiber.

With the advent of the Erbium Doped Fiber Amplifier (EDFA) in the late 1980s, it was possible to amplify the 1550 nm wavelength directly. Figure 1-240 depicts the rudiments of such a system. In practice, there are components such as gradient index lenses (GRIN) that couple the signal from fiber to fiber. These latter components are not shown in order to avoid cluttering the diagram.

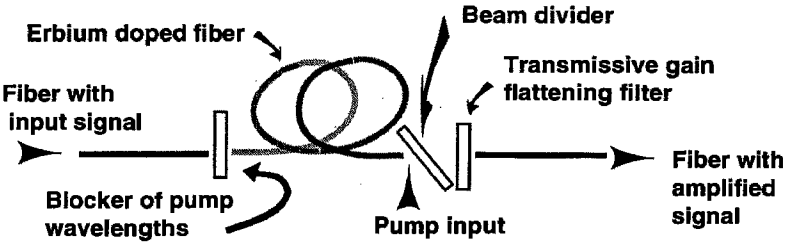


Fig. 1-240 Components of an erbium doped fiber amplifier. The details are given in the text.

The pump (at shorter wavelengths) is fed into the erbium doped fiber. Although Fig. 1-240 shows that this is done with a beam divider, "Y" fiber couplers are also widely used. The input signal is fed into the EDFA. The amplified signal emerges into the output fiber. The optical coatings often used in the system are listed below:

Blocker of pump radiation. This is an edge filter whose performance is discussed in §1.5.4.1.

Gain equalizer. The erbium does not have a uniform gain in the wavelength span over which the amplifier is used. An optical coating is used to correct this. This is discussed in §1.9.10.1.1.

Wavelength division multiplexing. This technology broadens the system bandwidth by permitting multiple wavelengths to propagate down the fiber. This is discussed in §1.6.3.2.

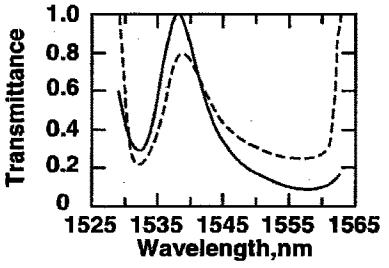


Fig. 1-241 $T(\lambda)$ of gain equalization filters for erbium doped fibers of varying chemical composition.

1.9.10.1.1 Gain equalization coatings

Both transmissive and reflective coatings are used to equalize the net gain versus wavelength. It is difficult to make a general comment about such systems. One of the problems is that the gain depends upon the composition of the core of the EDFA. Figure 1-241 depicts (for fibers of different composition) the transmittance of gain equalization filters. An advantage of the reflective gain equalizer is that in many cases, its loss (in the range of 0.2 to 0.3 dB) is smaller than that of a transmissive device (loss in the range of 0.4 to 0.5 dB). At least one paper (see Tilsch [99]) has been published on their design.

1.10 Appendices

1.10.1 The creation of an environment for an optical coating when the impinging light is noncollimated.

Suppose that light over a wide range of angles is converging upon an optical coating. The performance of the coating would be poor, indeed, if the coating were inserted "raw" into such an environment. This section illustrates that it is possible to create an optical system that reduces the range of angles and, thereby, improves the performance of the coating. In this section, "flux" (i.e., radiant flux) used rather than "light" so that the results may apply to the ultraviolet, visible, or infrared spectral regions.

1.10.1.1 Beam expanse and range of angles

§1.6.6.2 shows that the maximum cone angle is 15° for the bandpass analyzed in that section. This angle would be smaller if the bandwidth were narrower. What can be done? The answer: Expand the diameter of the beam — this reduces the range of angles.

An example is the flux collector in Fig. 1-242. There are three positions for a plano bandpass: at *B* in front of the objective lens, at *C* in the center of a Galilean telescope or in a convergent beam — in front of the detector *D*. The intent is to reduce the angular spread of rays that converges upon the bandpass, thus minimizing its angle shift and increasing the monochromatic flux flowing through it. The *Lagrange-Helmholtz invariant* can be used to calculate trade-offs involved in reducing the diameter of the bandpass. This is discussed in §1.10.1.2.

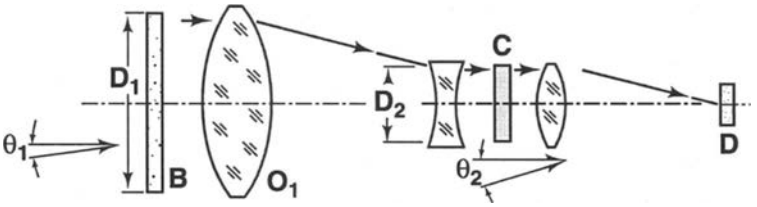


Fig. 1-242 Optical system collects flux from distant target of angular subtense θ_1, focuses it onto Galilean telescope and thence onto detector *D*.

A special case of the *Lagrange-Helmholtz invariant* obtains for paraxial optics ($\sin\theta \sim \theta$) in which a ray at an angle θ_1 travels through an aperture in an optical system of diameter D_1 ; it emerges through an aperture D_2 at an angle θ_1. This is

$$\theta_1 D_1 = \theta_2 D_2 \ . \qquad (1\text{-}24)$$

This equation is applied to the optical system in Fig. 1-242. The angular subtense (i.e., semi-angle) of the flux collected from a distant object is θ_1. If a bandpass of diameter D_1 is positioned in front of the objective, then the flux is relatively well collimated — the total cone angle is $2\theta_1$. It may be beyond present state of the art to fabricate such a large bandpass, especially if its bandwidth is quite narrow (see §1.6.6.3). A smaller bandpass of diameter D_2 can be inserted at *C* in the Galilean telescope. This decrease in diameter is obtained at the price of a wider angular divergence. Its semi-angle is θ_2

$$\theta_2 = \theta_1 D_1 / D_2 \qquad (1\text{-}25)$$

which is obtained from Eq. (1-24). As an example, suppose the objective is 120 mm ϕ and a 20 mm ϕ bandpass is located at *C* in Fig. 1-242. If the total angular target size is 3°, the divergence of the flux at *B* is 3° X (120/20) = 18° or ± 9°. §1.6.6.3 emphasizes that the layer thicknesses must be uniform across the lateral extent of the bandpass.

1.10.1.2 Lagrange-Helmholtz invariant

A general form of this invariant is

$$\mathcal{A}_1 n_1^2 \sin^2\theta_1 \geq \mathcal{A}_2 n_2^2 \sin^2\theta_2 \qquad (1\text{-}26)$$

where the symbols are illustrated in Fig. 1-243 [Winston 70; Rabl 76; Welford 78]. The incident flux (in a medium of index n_1) impinges upon an aperture of area $\mathcal{A}_1$ at a maximum angle θ_1. It passes through either a beam expander or concentrator and

emerges into a medium of index n_2; $\mathcal{A}_2$ is the area of the emergent beam and θ_2 is the maximum angle of a ray in that medium. Regardless of whether the system is imaging or is merely a nonimaging "flux bucket," Eq. (1-26) is valid. The equality holds only if there are no transmission losses.

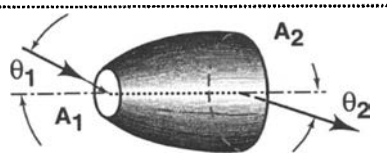


Fig. 1-243 Ray enters optical system (whose entrance pupil area is $\mathcal{A}_1$) at angle θ_1 and emerges from an exit pupil (area $\mathcal{A}_2$) at angle θ_2.

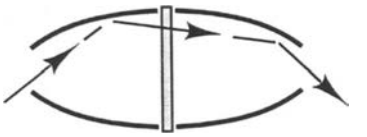


Fig. 1-244 Path of a meridional ray in cross section of a bandpass (*shaded*) sandwiched between two "back-to-back" compound parabolic concentrators. After Miles [81].

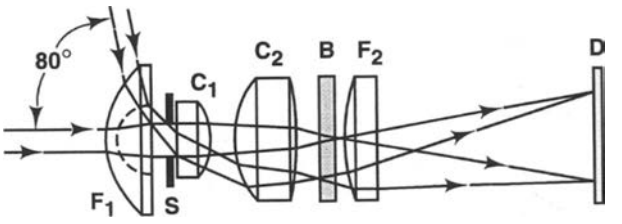


Fig. 1-245 Wide angle flux collector with the bandpass filter B in an internal telecentric image plane. Shown are the field compressor F_1, aperture S, condenser lenses C_1 and C_2, field lens F_2 and detector D. After Brixner [66].

The internally reflecting surface in Fig. 1-243 is a compound parabolic concentrator, which is an inexpensive way of expanding beam size in a non-imaging system [Winston 70]. Figure 1-244 shows two such concentrators "back-to-back," with a bandpass in the center [Miles 81]. Due to the expanded area, the divergence of the rays hitting the bandpass is reduced — as shown by Eq. (1-26).

1.10.1.3 Bandpass in a telecentric space

This optical system is advantageous when flux is collected over a vast range of angles [Brixner 66; Best 73]. As an illustration, suppose a naked filter and detector were substituted for the lens shown in Fig. 1-245. The angle of incidence upon the filter would range from zero to 80° — the angle shift (see §1.6.4.1) of its passband would be prohibitively large. A lens system with a telecentric space makes it possible to collect flux at such angles, while yet maintaining a relatively constant cone of convergent flux upon the filter.

With reference to Fig. 1-245, the lens combination C_1 and C_2 forms an image of the physical aperture S at infinity. Thus the bandpass filter B is in a telecentric space. Any ray that passes through the center of S impinges normally upon the surface of the filter, regardless of the field angle of that ray. The maximum angle of incidence on the filter is determined only by the f/number of the lens, which usually decreases with increasing field angle. This decrease, however, is mitigated somewhat by the field compression lens F_1, whose magnification distorts the pupil at larger field angles.

The field compression lens F_1 (see Fig. 1-245) is an added plus, although it is not essential to the system. It serves several purposes. First, it reduces the angle of incidence — from 80° on its incident side to nearly 60° on its emergent side. It also serves to distort the entrance pupil at high angles of incidence. Figure 1-245 shows that the entrance pupil size is nearly the same at normal incidence and at 80°. This enhances its flux collecting power at high angles and improves upon the "cosine theta" decrease in the entrance pupil area that would be unavoidable if this lens were omitted.

It is possible to insert a photomultiplier detector D immediately behind the filter B. But this has the disadvantage that the flux at each angle focuses at a different position on the photocathode. Variations in responsivity across the face of the photocathode would alter the signal, depending upon the angle. This is avoided by inserting the field lens F_2 that images the aperture onto the photocathode. In effect, it scrambles the rays and spreads the collected flux across the entire photocathode.

1.10.2 Coating designs

1.10.2.1 Reflector

The coating whose $R(\lambda)$ appears in Fig. 1-111 has the design:

air 92.85**H** 103.03**L** 68.08**H** 125.61**L** 66.67**H** 103.39**L** 58.187**H** 96.365**L** 53.527**H** 74.46**L** 38.74**H** 85.13**L** 54.18**H** 68.05**L** 24.72**H** *glass*

where the refractive indices of **H**, **L** and *glass* are 2.30, 1.35 and 1.52, respectively. Layer thicknesses are metric in nm.

1.10.2.2 Reflector

The design of a reflector (see Fig. 1-133) is

air 1.993**L** 0.646**H** 0.481**L** 0.676**H** 0.962**L** 1.158**H** 1.087**L** 0.555**H** 0.878**L** 0.72**H** 0.746**L** 0.767**H** 0.899**L** 0.993**H** 1.06**L** 1.111**H** 1.1**L** 1.13**H** 1.107**L** 1.133**H** 1.11**L** 1.128**H** 1.09**L** 1.072**H** 1.018**L** 0.97**H** 0.932**L** 0.887**H** 0.902**L** 0.921**H** *glass*

where the refractive indices of **H**, *glass* and **L** are 2.30, 1.52 and 1.47, respectively. Optical thickness of **H** and **L** is $\lambda_0/4$ at λ_0 of 665 nm and 725 nm, respectively.

1.10.2.3 Edge filter

The coating whose $T(\lambda)$ appears in Fig. 1-195 has the design:

air 95.01**L** 54.82**M** 22.24**H** 66.71**M** 140.56**L** 79.77**M** 38.17**H** 60.02**M** 132.82**L** 83.84**M** 32.90**H** 71.56**M** 139.60**L** 97.46**M** 22.92**H** 74.53**M** 138.54**L** 94.40**M** 37.57**H** 79.48**M** 136.50**L** 80.00**M** 35.75**H** 80.00**M** 134.52**L** 80.00**M** 35.75**H** 80.00**M** 134.52**L** 80.00**M** 35.75**H** 80.00**M** 134.52**L** 80.00**M** 35.75**H** 80.00**M** 134.52**L** 80.00**M** 35.75**H** 80.00**M** 134.52**L** 80.00**M** 35.75**H** 80.00**M** 134.52**L** 80.00**M** 35.75**H** 80.00**M** 134.52**L** 80.00**M** 35.75**H** 80.00**M** 134.52**L** 80.00**M** 35.75**H** 80.00**M** 134.52**L** 80.00**M** 35.75**H** 80.00**M** 134.52**L** 80.00**M** 35.75**H** 80.00**M** 134.52**L** 80.00**M** 35.75**H** 80.00**M** 134.52**L** 80.00**M** 35.75**H** 80.00**M** 134.52**L** 80.00**M** 35.75**H** 80.00**M** 134.52**L** 80.00**M** 35.75**H** 80.00**M** 134.52**L** 80.00**M** 35.75**H** 80.00**M** 134.52**L** 80.00**M** 35.75**H** 80.00**M** 134.52**L** 73.62**M** 40.27**H** 82.13**M** 132.08**L** 72.25**M** 32.96**H** 75.20**M** 127.63**L** 76.95**M** 40.17**H** 70.71**M** 127.78**L** 81.08**M** 26.17**H** 65.98**M** 134.46**L** 91.36**M** 30.10**H** 120.32**M** 61.54**L** *glass*

where the refractive indices of *glass*, **L**, **M** and **H** are 1.52, 1.38, 1.65 and 2.35, respectively. Layer thicknesses are metric, in nm.

1.10.2.4 Edge filter

The coating whose $T(\lambda)$ appears in Fig. 1-76 has the design:

qtz 52.5H 64L 51.5H 54.6L 38.4H 71L 49H 64L 55.1H 79.4L 45.5H 59.5L 49.1H 58.7L 48.1H 61.9L 52H 73.2L 54.1H 61.1L 49.5H 57.9L 47.6H 60.2L 42H 68.5L 54.2H 70.3L 58.2H 62.4L 46.1H 54.3L 43.7H 58.8L 44.3H 54.8L 43.6H 53.7L 43H 54.1L 44H 55.3L 44.5H 55.7L 45H 56.1L 44.8H 55.1L 43.4H 53.5L 42.7H 53.2L 42.8H 54.1L 58.6H 48.2L 111.6H 45.7L 71H *air*

where the refractive indices of *qtz*, H and L are 1.48, 1.85 and 1.45, respectively. Layer thicknesses are metric, in nm.

1.10.2.5 Ward polarizer

The arrangement of the two prisms in the Ward polarizer appears in Fig. 1-176. The design of the system is:

glass 1.114H 1.37L 1.114H 1.37L 1.114H S 1.114H 1.37L 1.13H 1.37L 1.114H *glass*

where the indices of *glass*, H, S and L are 1.508, 2.25, 1.00 and 1.45, respectively. The optical thickness of H and L is $\lambda_0/4$ at λ_0 of 1.065 μm. The metric thickness of the air-gap S is 23 μm.

1.10.2.6 Reflector with −90° differential phase shift upon reflection

The coating whose $R(\lambda)$ and $\Delta(\lambda)$ appear in Fig. 1-133 has the design:

air 0.06147L 0.10377H 0.15412L 0.09102H 0.15903L 0.09260H 0.15611L 0.08978H 0.15339L 0.09101H 0.16009L 0.09515H 0.16388L 0.09480H 0.15721L 0.08542H 0.13404L 0.07412H 0.12458L 0.07233H 0.12377L 0.07225H 0.12408L 0.21582H 0.12403L 0.07224H 0.12396L 0.21550H *glass*

where the refractive indices of *glass*, H and L are 1.52, 2.30 and 1.46, respectively. Layer thicknesses are metric, in μm.

1.10.2.7 Reduction of the differential phase shift upon reflection

The design (see Fig. 1-136) is

glass 0.147H 0.299L 1.357H 0.619L 0.526H *air*

where indices of H, *glass* and L are 2.30, 1.52 and 1.45, respectively. The optical thicknesses are multiples of $\lambda_0/4$ at λ_0 of 1.0 μm.

1.10.2.8 YAG reflector (1064 nm)

The design (see Fig. 1-131) is

air H $(L\ H)^6$ *glass*

where indices of H, *glass* and L are 2.25, 1.51 and 1.45, respectively. The optical thicknesses are $\lambda_0/4$ at λ_0 of 1.064 μm.

1.10.2.9 Cemented reflector for *p* polarization

The design (see Fig. 1-140) is

glass $(H\ L)^{28}$ *glass*

where the refractive indices of *glass*, L and H are 1.52, 1.7168 and 2.30, respectively. The metric thicknesses of H and L are 76.8 nm and 123.5 nm, respectively.

1.10.2.10 Nonpolarizing slab beamdivider

The design (see Fig. 1-161) is

air $(M\ H\ M\ L)^5$ M *glass* .

Refractive indices of *glass*, L, M and H are 1.52, 1.442, 1.61 and 2.26, respectively. Optical thicknesses of L, M and H are $\lambda_0/4$ at λ_0 of 726 nm, 705 nm and 667 nm, respectively.

1.10.2.11 Immersed absorbing beamdivider

The design (see Fig. 1-163) is

glass 1.52**L** 0.67**H** **M** 0.511**H** 0.367**L** 0.247**H** 0.385**L** 0.246**H** *glass* .

Refractive indices of *glass*, **L** and **H** are 1.52, 1.35 and 2.25, respectively. **M** is silver of 24.8 nm thickness. Optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 1 μm.

1.10.2.12 Dark mirror

The design (see Fig. 1-214) of dark mirror is

air 1.758**L** 0.633**H** 0.0123**M** 1.484**H** 0.0456**M** 1.918**H** 0.032**M** 0.969**H** 0.0145**M** 1.125**H** *metal*

Optical constants of **M** (its metric thickness is 1.0 μm) and *metal* are those of bulk titanium [Weaver 81]. The refractive indices of **L** and **H** are 1.45 and 2.30, respectively. Optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 1 μm.

1.10.2.13 Cemented linear polarizer

The design (see Fig. 1-197) is

glass 0.71**H** 1.436**L** 1.084**H** 1.351**L** 1.052**H** 1.371**L** 1.11**H** 1.42**L** 1.154**H** 1.44**L** 1.156**H** 1.43**L** 1.135**H** 1.42**L** 1.124**H** 1.42**L** 1.135**H** 1.43**L** 1.156**H** 1.44**L** 1.154**H** 1.42**L** 1.11**H** 1.37**L** 1.05**H** 1.35**L** 1.084**H** 1.436**L** 0.71**H** *glass* .

The refractive indices of **L**, *glass* and **H** are 1.50, 1.517 and 2.30, respectively. Optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 550 nm.

1.10.2.14 Red-reflecting and blue-reflecting dichroics

The red-reflecting design (see Fig. 1-187) is

glass 88.19**S** 99.23**H** 143.10**S** 89.38**H** 135.42**S** 85.31**H** 135.61**S** 83.26**H** 134.54**S** 84.03**H** 131.13**S** 85.21**H** 130.47**S** 85.21**H** 131.12**S** 84.02**H** 134.53**S** 83.26**H** 135.61**S** 85.31**H** 135.41**S** 89.38**H** 143.09**S** 99.22**H** 88.19**S** *glass*

and the blue-reflecting design is

glass 32.45**H** 58.47**S** 61.18**H** 73.04**S** 56.75**H** 55.01**S** 60.46**H** 72.99**S** 61.83**H** 58.87**S** 58.51**H** 67.02**S** 62.94**H** 67.00**S** 58.52**H** 58.88**S** 61.82**H** 72.99**S** 60.45**H** 55.02**S** 56.75**H** 73.01**S** 61.20**H** 58.47**S** 32.45**H** *glass* .

The refractive indices of **S**, *glass* and **H** are 1.65, 1.52 and 2.32, respectively. Layer thicknesses are metric, in nm.

1.10.2.15 Green-reflecting dichroic and phase compensation coating

The differential phase shift curves appear in Fig. 1-190. The all-dielectric green-reflecting design is

glass 0.0652**H** 0.02589**M** 0.019**H** 0.08287**M** 0.02014**H** 0.0605**M** 0.06145**H** 0.091395**M** 0.085236**H** 0.065035**M** 0.04887**H** 0.065944**M** 0.02163**H** 0.00997**M** 0.0525**H** 0.077245**M** 0.061563**H** 0.07956**M** 0.062975**H** 0.080535**M** 0.06283**H** 0.0797**M** 0.06222**H** 0.079333**M** 0.062276**H** 0.079687**M** 0.062614**H** 0.079976**M** 0.062614**H** 0.079687**M** 0.062276**H** 0.079333**M** 0.06222**H** 0.07969**M** 0.06283**H** 0.080535**M** 0.062975**H** 0.079559**M** 0.061563**H** 0.077245**M** 0.0525**H** 0.00997**M** 0.02163**H** 0.065944**M** 0.04887**H** 0.065035**M** 0.085236**H** 0.091395**M** 0.061445**H** 0.0605**M** 0.020136**H** 0.082865**M** 0.019**H** 0.025888**M** 0.0652**H** *glass*

and the design of the phase compensation coating is

glass 0.16641**H** 0.016307**L** 0.068331**H** 0.102587**L** 0.060998**H** 0.065014**L** 0.017043**H** 0.090510**L** 0.065746**H** 0.114461**L** 0.077697**H** 0.381169**L** 0.078212**H** 0.113362**L** 0.066914**H** 0.106520**L** 0.065015**H** 0.089444**L** *aluminum*

where **H**, **M** and **L** represent layers of refractive indices 2.25, 180 and 1.46, respectively. Layer thicknesses are metric, in micrometer. The *aluminum* is opaque to visible light.

1.10.2.16 Three cavity bandpass design

Its design is

air 0.23**L** 0.5**H** 1.23**L** **H L H** 4**L** (**H L**)^6 **H** 4**L** (**H L**)^6 **H** 4**L** (**H L**)^3 *glass*

where *glass*, **L** and **H** represent media of refractive indices 1.52, 1.45 and 2.30, respectively. Optical thicknesses of **L** and **H** are multiples of $\lambda_0/4$ at λ_0 of 633 nm.

1.10.2.17 Dichroic blue reflector

Its design (see Fig. 1-194) is

glass (**H L H**)^{10} *glass*

where refractive indices of *glass*, **L** and **H** are 1.517, 1.50 and 2.30, respectively. The optical thicknesses of **L** and **H** are 84 nm and 77 nm, respectively.

The loss $\mathcal{L}$ is the fraction of the flux that is not reflected

$$\mathcal{L} \equiv 1 - R = A + T \qquad \text{(1-4b)}$$

regardless of whether the loss is due to transmittance or absorption. Loss is important, for example, in the design of a coating for a laser cavity.

Chapter 2
Fundamentals

2.0 Overview of chapter

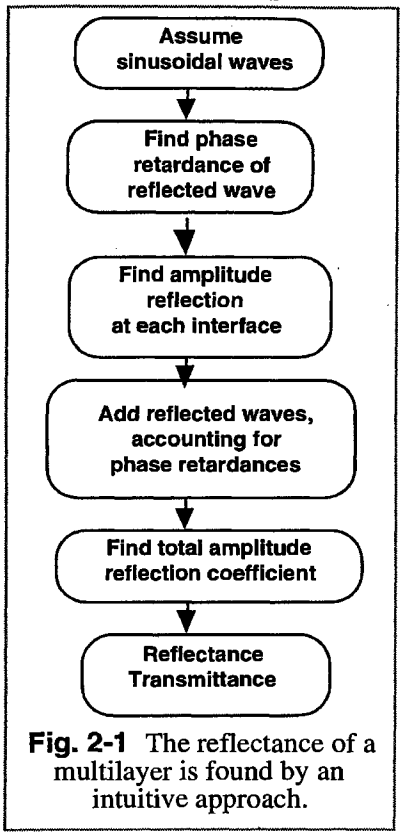


Fig. 2-1 The reflectance of a multilayer is found by an intuitive approach.

The ultimate goal is to find the radiant reflectance, transmittance, phase shift upon reflection and so on, of one or more thin layers — the multilayers whose performances are described in Chap. 1.

There is an introductory section that computes the reflectance by an intuitive method. On the left (Fig. 2-1) is a summary of that section.

A more rigorous derivation is provided in an appendix (§2.10.1) and further sections of this book — as summarized in Fig. 2-2.

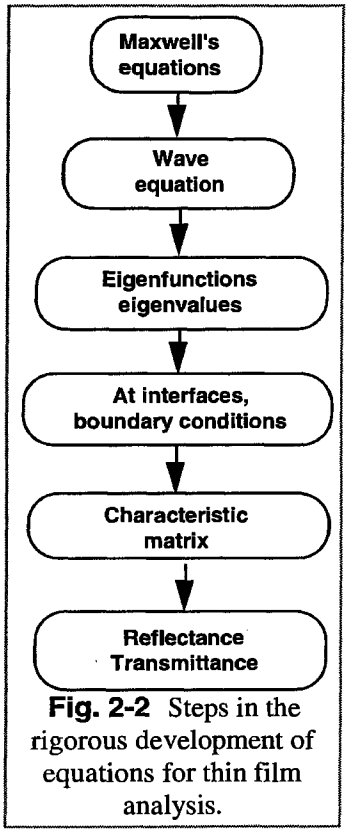


Fig. 2-2 Steps in the rigorous development of equations for thin film analysis.

2.1 Overview of intuitive approach

The initial goal in §2.1 is to pursue the intuitive approach — the "addition of waves" method of finding the reflectance of one or more layers cited in Fig. 2-1 above. §2.1 discusses the propagation of waves — waves that penetrate a multilayer stack. These light waves bounce off of the interfaces between the layers of a stack. The reflection at an interface is treated in §2.2.3. The addition of waves is the subject of §2.2.6 and §2.2.7.

2.1.1 Waves propagating in a nonabsorbing medium

2.1.1.1 Traveling waves

A light source illuminates a stack of films. Examples of the source are a filtered sodium lamp or laser. Because the waves *travel* from the light source to the stack, this is termed a *traveling wave.* An ocean wave serves as an example (albeit rather crude) of such a traveling wave. Surfboard riders are transported by this wave, as shown in Fig. 2-3.

Figure 2-4 shows a cross section of the wave at an instance of time. Each surfboard rider must adjust the velocity of the board so that it remains about 20° in phase

in front of the crest of the wave. If this is not done, the board will not always travel "down hill" towards the beach and (alas!) the rider will "dump."

The shape of the wave in Fig. 2-4 is approximated by a sine curve. The height above mean sea level is the *amplitude* of the wave. The frequency ν of the wave is found by measuring the time it takes for successive crests to pass a given point—and taking its reciprocal. The distance from "crest to crest" is the wavelength λ'—also depicted in Fig. 2-4. A coordinate system is added—the "plus zee" direction z is towards the beach.

An equation may be written for the amplitude E of the wave (varying with both time t and distance z)

$$E^{(+)} = E_0^{(+)} \sin(2\pi\nu\ t - 2\pi z / \lambda') \qquad (2\text{-}1a)$$

where $E_0^{(+)}$ is the maximum amplitude. The superscript "(+)" indicates that this wave travels in the "+z" direction. The foregoing equation also represents the propagation of a light wave. In fact, the word *amplitude* is used in older books on physical optics. The *amplitude* is, in fact, a scalar component of the electric field of the light wave.

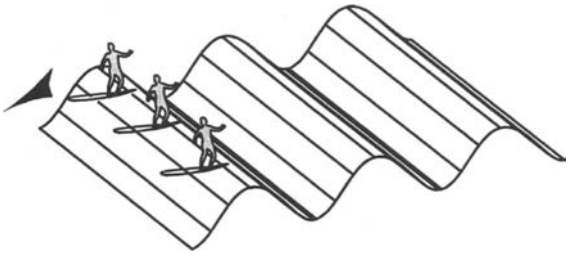


Fig. 2-3 Positioned slightly in advance of a crest of a wave, surfboard riders travel towards the shore.

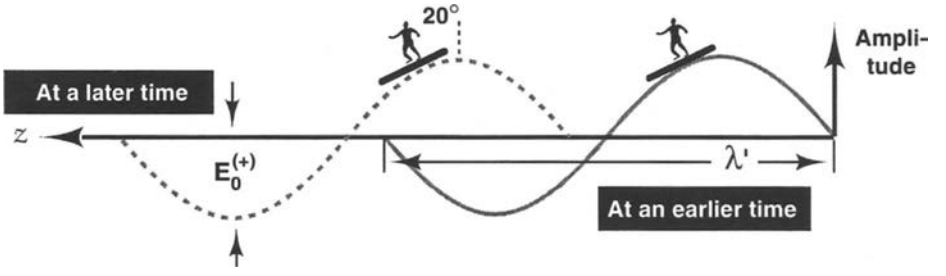


Fig. 2-4 A surfboard rider is transported in the +z direction by riding slightly in advance of the crest of a traveling wave. $E_0^{(+)}$ is the amplitude of the wave.

2.1.1.2 Retardance of a wave

An analogy is made between the velocity of a runner and the phase velocity of a wave. The time t that it takes a runner to traverse a distance h is related to the velocity v

$$t = h / v \ . \qquad (2\text{-}1b)$$

For example, the author accompanied a group that recently backpacked along a 13 km stretch of beach on the California coast. Under normal circumstances the trip would have been completed in four hours. But hiking in loose sand and on wave-swept rock slowed

us down—we arrived in six hours. This is evident in Eq. (2-1b)—the lower the velocity, the longer the time.

Applying the foregoing equation to a light wave, the phase velocity v of a wave is related to the refractive index n

$$v = c/n \tag{2-1c}$$

where c is the velocity of light. This shows that a light wave slows down when it penetrates a medium of higher refractive index, such as glass. Combining the previous equations provides for the time delay

$$t = \frac{h}{v} = \frac{h}{c/n} = \frac{n\,h}{c} \ . \tag{2-1d}$$

It depends upon the product of the physical distance h and refractive index n. The product $n\,h$ is termed the *optical thickness* or *optical path* and appears frequently in this book.

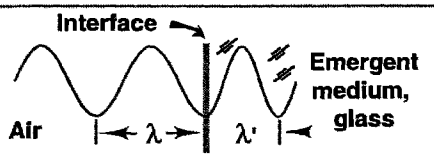


Fig. 2-5 At an instance of time, the electric field of a light wave as it impinges upon a glass surface. λ and λ' are the wavelengths in air and in the emergent medium of index 1.52, respectively.

A consequence of Eq. (2-1c) is that the wavelength λ' in a medium of index n is related to the wavelength λ in a vacuum by

$$\lambda' = \lambda / n \tag{2-1e}$$

and Eq. (2-1a) is written

$$E^{(+)} = E_0^{(+)} \sin(2\,\pi\,\nu\,t - 2\,\pi\,z\,n/\lambda). \tag{2-1f}$$

Figure 2-5 shows the total amplitude of a light wave (at a given instance of time) as it travels from air and into glass of index 1.52. Inside the glass, the wavelength is foreshortened by a factor of 1/1.52 = 0.66. In subsequent sections of the book, the wavelength λ, even when appended by subscripts such as λ_0 or λ_1, always refers to the vacuum wavelength.

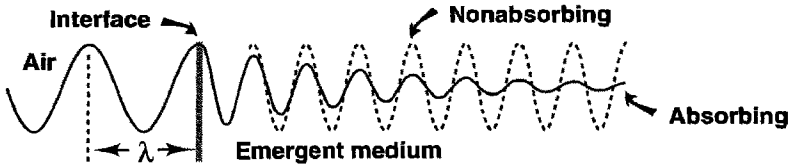


Fig. 2-6 The caption to Fig. 2-5 obtains, with the exception that the emergent medium (into which the wave penetrates) has an index of 2.00 and is nonabsorbing (*dashed*) or absorbing (*solid*), with $\hat{n} = 2.00 - j\,0.10$.

For a wave propagating a distance $z = h$ into a medium of index n, its phase retardance β is

$$\beta \equiv 2\,\pi\,n\,h\,/\,\lambda \tag{2-1g}$$

and is expressed as either radians or degrees.

2.1.2 Properties of an absorbing medium

The optical properties of a medium are characterized by two parameters:

- refractive index n and
- extinction coefficient k.

The refractive index (encountered in §2.1.1.2) quantifies the phase velocity of the wave in that medium. The extinction coefficient k establishes the level of absorption. That is, the larger k value, the greater the rate at which the energy of the plane wave is converted into joule heat in the medium in which it is propagating. k is zero if the medium is truly nonabsorbing—such a medium is often called a *dielectric*. The n and k are combined into the (complex) optical constant $\hat{n}$ as

$$\hat{n} = n - j\,k \ . \qquad (2\text{-}1h)$$

Figure 2-6 depicts a "snapshot" of the total electric field of a wave as it penetrates an absorbing medium with an n of 2.00 and a k of 0.1. With the origin at the air-medium interface, the equation that describes this wave is

$$E^{(+)}(z,t) = E_0^{(+)} \cos(2\pi\nu\ t - 2\pi z n/\lambda) \exp(-2\pi k z/\lambda) \ . \qquad (2\text{-}2a)$$

As it travels a distance z, phase is retarded by $2\pi n z/\lambda$ radian and its amplitude is attenuated by "$\exp(-2\pi k z/\lambda)$".

An example is to find the decrease of its amplitude when the optical thickness increases by a wavelength. This is the distance between $z = 0$ and $z\ n\ /\ \lambda = 1$. The exponential in Eq. (2-2a) reduces to

$$\exp(-2\pi k/n) = \exp(-2\pi\ 0.1/2.00) = \exp(-0.31) = 0.73 \ . \qquad (2\text{-}2b)$$

In other words, each time the optical thickness increases by a wave, the amplitude is attenuated by 73%, as shown in Fig. 2-6. The electric field is quite attenuated after the wave penetrates eight wavelengths into the medium.

Table 2-1 *n* and *k* of thin films

Material	n	k	λ	Comment
cryolite	1.35	10^{-4}	450 nm	Low index material for "cemented" BP construction.
zinc sulfide	2.35	10^{-3}	450 nm	High index material for "cemented" BP construction.
germanium	4.2	2×10^{-3}	3 μm	This is transparent in much of the IR.
germanium	5	1	500 nm	Material is an absorber in the visible spectrum.
silver	0.2	3.0	600 nm	Base metal for a reflector (see §1.7.4).
gold	32	70	12 μm	Base metal for a reflector (see §1.7.4).
chromium	2.5	3.5	500 nm	Used in neutral density filters (see §8.7.3).

Representative n and k values of optical coating materials at the specified wavelength λ. BP refers to "bandpass filter."

2.1.2.1 Range of *n* and *k* values

Table 2-1 shows that there is a large range of n and k values for thin films. n of dielectric materials in the visible part of the spectrum ranges from 1.35 to 2.35—the k can be as low as 10^{-5} for a titania layer. The index of an IR coating material such as germanium is as large as 4.2. In the visible part of the spectrum, however, that material is absorbing—a k of 1.00 is typical. The n and k of a transition metal, such as nickel or chromium, are approximately equal. The n of a silver layer is at least ten times smaller than its k at wavelengths near 600 nm. Both the n and k of a noble metal are large in the middle IR.

2.1.3 Dispersion of the medium in which the wave propagates

Equation (2-1g) shows that the phase retardation is proportional to frequency — i.e., inversely proportional to wavelength. This is also true of the attenuation of the wave — the argument of the exponential in Eq. (2-2a). In addition, n and k themselves in those equations vary with wavelength; this is termed *dispersion*. It is important to account for this, especially when a multilayer functions over a wide range of wavelengths. In an appendix (§2.11.1) it is shown that in the shorter wavelength region where a material is reasonably transparent, the refractive index may be approximated by

$$n^2 \approx A + B\lambda^{-2} \quad . \tag{2-2c}$$

§2.11.1 discusses a method of determining the A and B constants.

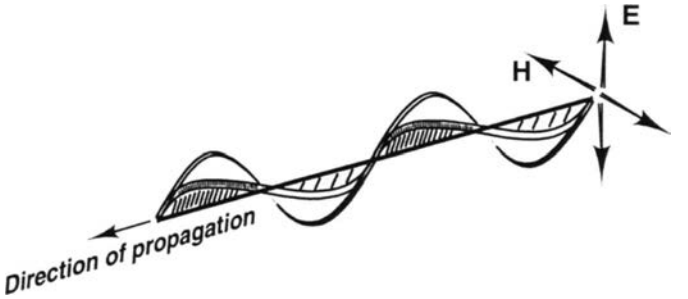


Fig. 2-7 At a given instance of time, a snapshot of the electric field E (vertical) and magnetic field H (horizontal) of an electromagnetic wave.

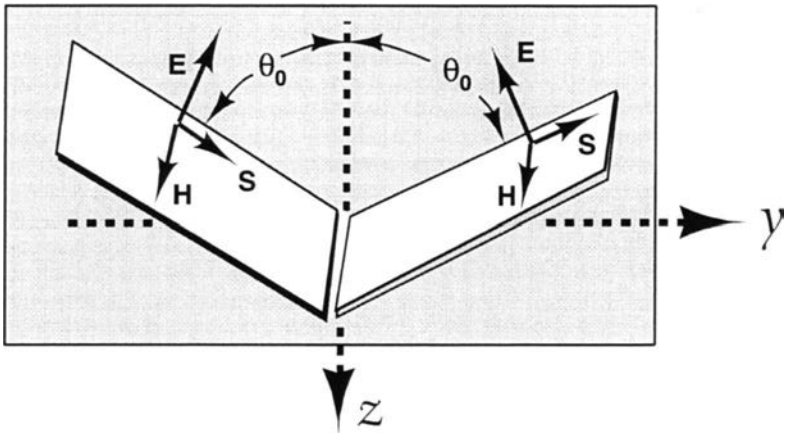


Fig. 2-8 A linearly polarized plane electromagnetic wave impinges upon an interface (in the x-y plane) between two media — the x axis is into the paper. The vector $\vec{s}$ is a normal to the wave front that propagates at an angle θ_0 to the z axis. The vectors representing the electric field E and magnetic field H are for TM polarization.

2.2 Reflection and transmission at an interface

2.2.1 Overview and establishment of terminology

§2.2.1.1 cites a mathematical expression for a wave propagating at nonnormal incidence in a medium that may be absorbing. Figure 2-7 in §2.1.2 depicts the electric field E and magnetic field H of that wave—as they appear in the incident medium. §2.2.1.2 considers the linear polarization states of such a wave. §2.2.2 applies the boundary conditions for E and H—the reflection coefficients are found. Earlier parts of §2.2 refer to §2.2.5, in which the electric field and the magnetic fields are related by their admittance.

A light wave is a manifestation of an electromagnetic (EM) wave. Figure 2-7 depicts the E and H fields at a given instance of time for a linearly polarized EM wave. The E field vibrates in a plane perpendicular to the plane of the H field. The directions of both E and H are orthogonal to the direction of propagation. In Fig. 2-8, the E and H field directions are shown as vectors—the sinusoidal variation is omitted.

2.2.1.1 Wave propagation at nonnormal incidence

An appendix (§2.10.2) relies upon Maxwell's equations to derive an equation for a plane wave. For that reason, references to §2.10.2 appear below. The medium in which the wave propagates before it impinges upon a multilayer is termed the *incident medium.* Its refractive index is n_0. Figure 2-8 shows the unit vector $\vec{s}$ ($\perp$ to the wave front) that lies in the y-z plane at an angle of incidence θ_0 to the z axis. §1.2.2.5 (see Fig. 1-15) mentions the linear polarization of such a wave and also that the plane of incidence contains the incident ray (normal to the wave front) and the normal to the surface. By definition, the wave does not propagate in the x direction; k_x [in Eq. (2-70a)] is zero. These substitutions are made into Eq. (2-70b)

$$k_y = 2\pi n_0 \sin\theta_0 / \lambda \qquad (2\text{-}2d)$$

and

$$k_z = 2\pi n_0 \cos\theta_0 / \lambda \ . \qquad (2\text{-}2e)$$

where λ is the vacuum wavelength. Taking the square root of Eq. (2-70b) results in

$$2\pi n_0 \lambda^{-1} \sqrt{\sin^2\theta_0 + \cos^2\theta_0} \ = 2\pi n_0/\lambda \ . \qquad (2\text{-}2f)$$

When the medium is absorbing, the general form of the wave function is

$$E(y,z,t) = \hat{E}_0 \exp(j\omega t) \exp[-j\, 2\pi \ (y\hat{\eta} + z\hat{\gamma}) / \lambda'] \qquad (2\text{-}2g)$$

where λ' is the wavelength in that medium. $\hat{\eta}$ and $\hat{\gamma}$ are yet to be determined. These are derived in §2.2.2 by application of the boundary conditions at an interface.

2.2.1.2 Waves impinging upon the interface

The interface consists of the plane boundary between the ith medium of index n_i and a medium of optical constant $\hat{n}_{i+1}$ on its emergent side. A coordinate system is established in which the interface lies in the x - y plane, with the z axis $\perp$ to the interface. The vector $\vec{s}$ (normal to the wave front) defines the direction of the incident wave; θ_0 is the angle between $\vec{s}$ and the z axis. The plane containing $\vec{s}$ and the z axis defines the plane of incidence. If the emergent medium is not anisotropic, then the vector representing the normal to the wave front of the reflected wave lies in the plane of incidence, as shown in Fig. 2-8.

2.2.2 Boundary conditions at an interface

The propagation constant of the wave in the emergent medium is determined by the boundary conditions at the interface. At this interface it is assumed that: (*a*) there is no permanent charge contributing to the electromagnetic field at optical frequencies; and (*b*) no currents exist that can do likewise. Consider the electric field E and magnetic field H at the boundary between the ith and $(i+1)^{th}$ media. It is shown in texts on electrodynamics [Jackson 75] that

$$H_i\,]_{tan} = H_{i+1}]_{tan} \tag{2-2h}$$

and

$$E_i\,]_{tan} = E_{i+1}]_{tan} \tag{2-2i}$$

where *tan* refers to the components tangential to the interface. The incident and reflected components of the wave in the incident medium (medium "zero") are described by

$$\hat{E}^{(+)}(y,z,t) = \hat{E}_0^{(+)} \exp(j\omega t) \exp[-j k_0 (y \sin\theta_0 + z \cos\theta_0)] \tag{2-3a}$$

$$\hat{E}^{(-)}(y,z,t) = \hat{E}_0^{(-)} \exp(j\omega t) \exp[j k_0 (-y \sin\theta_0 + z \cos\theta_0)] \ , \tag{2-3b}$$

respectively, where

$$k_0 \equiv 2\pi n_0 / \lambda \tag{2-3c}$$

and λ is the vacuum wavelength.

The wave $\hat{E}_i$ in the ith medium of optical constant $\hat{n}_i$ may be represented in several ways. Either Eq. (2-2g) or

$$\hat{E}_i^{(+)}(y,z,t) = \hat{E}_{i,0}^{(+)} \exp(j\omega t) \exp[-j (y k_y' + z k_z')] \tag{2-3d}$$

is used—$\hat{E}_{i,0}^{(+)}$ is the (complex) amplitude of the wave in that medium.

In either case, in order for Eq. (2-2i) to be satisfied *at all values* of y, the y components of the propagation constant [in Eqs. (2-3b) and (2-3d)] must match

$$k_0 \sin\theta_0 = 2\pi \hat{\eta}\, n / \lambda \tag{2-3e}$$

and hence $2\pi\, n\, \hat{\eta}/\lambda$ in Eq. (2-2g) must be pure real. If $\hat{\eta}$ [see Eq. (2-2g)] were not real, then the beam would be attenuated in the y direction, in contradiction to observations. Equation (2-1e) has been used above.

Substituting Eq. (2-70c) for the appropriate medium in Eq. (2-3e) results in

$$n_i \sin\theta_i = n_i\ \hat{\eta} = n_0 \sin\theta_0 \ . \tag{2-3f}$$

θ_i is the angle of incidence in the ith medium. The foregoing is reiterated in Eq. (2-4a). k_y' and k_z', or, alternatively, $\hat{\eta}$ and $\hat{\gamma}$ are not independent but are related. A combination of Eqs. (2-70b) and (2-70c) provides

$$k_z' = 2\pi \lambda^{-1} \sqrt{\hat{n}_i^2 - n_0^2 \sin^2\theta_0} \ . \tag{2-3g}$$

In the ith medium, the (complex) propagation constant in the z direction is

$$\hat{\gamma}_i \equiv \sqrt{1 - \hat{n}_i^{-2}\, n_0^2 \sin^2\theta_0} \quad = \hat{n}_i^{-1} \sqrt{\hat{n}_i^2 - n_0^2 \sin^2\theta_0} \tag{2-3h}$$

where n_0 is the index of the incident medium. $\hat{\gamma}$ is pure real, pure imaginary or complex, depending upon the angle of incidence and relative refractive indices of the layers.

2.2.2.1 Nonnormal incidence, angle of incidence less than critical angle

In this instance, it is assumed that the ith medium is nonabsorbing and that if a critical angle (see §2.2.2.2) exists, the incidence angle θ_0 is less than that angle. Eq. (2-3h) simplifies to

$$\hat{\gamma}_i = \cos\theta_i + j0 \ . \tag{2-3i}$$

As a result of the analogies between Eqs. (2-3h) and (2-3i), $\hat{\gamma}$ is often called the *complex cosine*. The only physical significance of that terminology is that it reduces to Eq. (2-3i) when the medium is nonabsorbing.

§1.8.1.1 mentions that the incidence angle may be defined in a reference medium (of refractive index n_{ref}) or as θ_0 in the incident medium of index n_0. Equation (2-3f) generalizes to Snell's Law

$$L \equiv n_0 \sin\theta_0 = n_{ref} \sin\theta_{ref} = n_i \sin\theta_i \tag{2-4a}$$

in the ith medium. L [defined also in Eq. (1-19)]is termed the *numerical aperture* or the *immersion constant* (see §1.8.1.1). The incidence angle θ_i in the ith layer is

$$\theta_i = \arcsin(L/n_i) \ . \tag{2-4b}$$

Finally, the *phase of retardation* β is introduced by writing for the ith layer a recasting of Eq. (2-3a)

$$\hat{E}_i = \hat{E}_i^{(+)} \exp(-j\, 2\pi\, y\, n_0 \sin\theta_0 \lambda^{-1}) \exp(-j\beta_i) \exp(j\omega t) \tag{2-4c}$$

where

$$\beta_i \equiv 2\pi n_i h_i \cos\theta_i / \lambda \ . \tag{2-4d}$$

h_i is the distance the wave penetrates along the z axis.

2.2.2.2 Angle of incidence exceeds the critical angle

Consider the case where the refractive index of the incident medium exceeds the index of the substrate or a layer in a stack. Mathematically, $n_i < n_0$ and there exists a critical angle θ_c in the incident medium

$$\theta_c \equiv \arcsin(n_i/n_0) \ . \tag{2-4e}$$

A consequence is that when $\theta_0 > \theta_c$ for that medium, θ_i in Eq. (2-4b) is no longer real. In this case Eq. (2-3g) becomes

$$k_z' = -j\, 2\pi \lambda^{-1} \sqrt{L^2 - n_i^2} \tag{2-4f}$$

and is pure imaginary. Why is the negative root chosen? With this sign, the retardance β_i appearing in Eq. (2-4d) is

$$\beta_i = -j\, \beta_i' \tag{2-4g}$$

where β_i' is pure real and positive. It is written as

$$\beta_i' \equiv 2\pi h_i \lambda^{-1} \sqrt{L^2 - n_i^2} \ . \tag{2-4h}$$

Omitting the time dependence in Eq. (2-4c), the wave propagation is

$$\hat{E}_i = \hat{E}_i^{(+)} \exp(-j\, 2\pi\, y\, L\, \lambda^{-1}) \exp(-\beta_i') \tag{2-4i}$$

which shows that $\hat{E}_i$ is exponentially attenuated as it penetrates a distance h in the +z direction. Furthermore, the electric field becomes vanishingly small as the layer becomes thick compared to a wavelength. This would not be the case if the sign of Eq. (2-3g) were reversed.

Alternatively, $\hat{\gamma}$ [defined in Eq. (2-3h)] is pure imaginary when θ_0 is large enough so that it exceeds the critical angle in the ith layer. Again, there is an ambiguity in the sign for the square root in Eq. (2-4f). This may be visualized if

$$Arg \equiv n^2 - L^2 \tag{2-4j}$$

is plotted on the Argand plane, as shown in Fig. 2-9. *Arg* lies at B on the negative real axis when it is negative and pure real. Its square root could lie at point C' on the positive y axis or at C on the negative y axis. As stated above, the latter root is correct. Often the routine in a computer code that finds the root of a complex number may betray the

programmer and return the root at C'. The remedy is to add a modicum of negative imaginary part (such as $-j10^{-7}$) to the optical constant in Eq. (2-4j). The correct root at C is then found by the computer code. The optical properties are not affected by the small imaginary part unless the layer is thick compared to a wavelength.

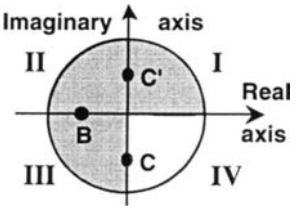


Fig. 2-9 The complex dielectric constant [see Eq. (2-69j)] is located in quadrants III or IV of an Argand diagram. The optical constant $\hat{n}$ [(see Eq. (2-1h)] lies in quadrant IV (unshaded). If *Arg* in Eq. (2-4j) lies at point B, the square roots lie at either C' or at C.

2.2.2.3 The medium in which the wave propagates is absorbing

In this case the phase retardance is complex and is

$$\hat{\beta}_i \equiv \beta_r + j\beta_j = 2\pi\sigma h_i \sqrt{n_i^2 - k_i^2 - L^2 - j2n_i|k_i|} \qquad (2\text{-}5a)$$

where σ is the vacuum wave number. h_i, n_i and k_i are the metric thickness, refractive index and extinction coefficient, respectively, of the ith layer. L appears in Eq. (2-4a). Also, $k = |k|$ in the foregoing equation and the root should lie in quadrants III or IV in the complex plane. $L = 0$ at normal incidence and Eq. (2-5a) simplifies to

$$\hat{\beta}_i = \beta_r + j\beta_j = 2\pi\sigma h_i n_i - j2\pi\sigma h_i|k_i| \ . \qquad (2\text{-}5b)$$

The imaginary part of the phase retardance is

$$\beta_j \equiv -2\pi h_i|k_i|/\lambda = -2\pi\sigma h_i|k_i| \ . \qquad (2\text{-}5c)$$

2.2.3 Fresnel coefficients at an interface

2.2.3.1 Definitions

As noted in §2.2.1, there is a wave propagating in the ith medium. This is resolved into its components, which consist of the incident component (propagating in the $+z$ direction) and the reflected component (propagating in the $-z$ direction). The reflection and transmission coefficients of these waves are found by assuming that there is no reflected wave in the $(i+1)^{th}$ medium. In classical physical optics, the constants $\hat{E}_0^{(+)}$, $\hat{E}_0^{(-)}$ and $\hat{E}_{i,0}^{(+)}$ in Eqs. (2-3a), (2-3b) and (2-3d) are usually termed the *amplitudes* of these waves. The *assumed* directions of the E, H fields appear in the left panel of Fig. 2-10. The sign convention for the amplitude is:

> The electric field is positive if a component is directed in the direction along the $+y$ axis or $+x$ axis of the coordinate system.

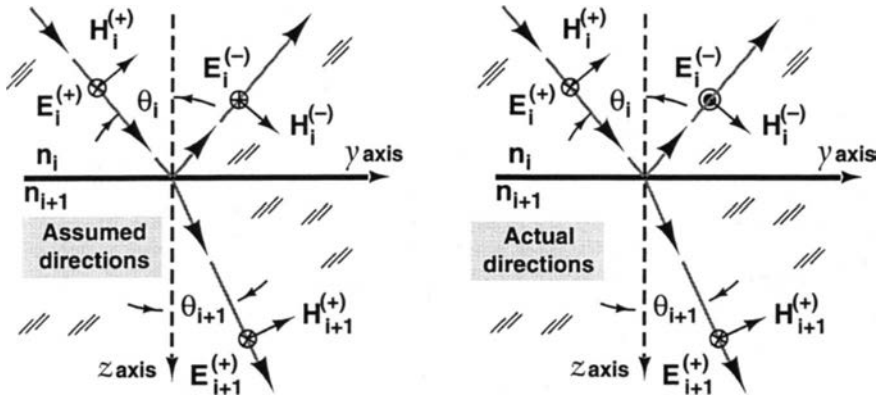


Fig. 2-10 A TE (i.e., *s*) polarized EM wave propagates at an incidence angle θ_i in a medium of index n_i and is reflected at the interface at the medium of index n_{i+1}. (*Left*) *Assumed* directions of the electric field E and magnetic field H. (*Right*) *Actual* field directions, $n_i < n_{i+1}$. The Cartesian coordinates y and z are shown. x is perpendicular to and into the paper. The symbols ⊙ and ⊗ represent a vector *out of* the paper and *into* the paper, respectively.

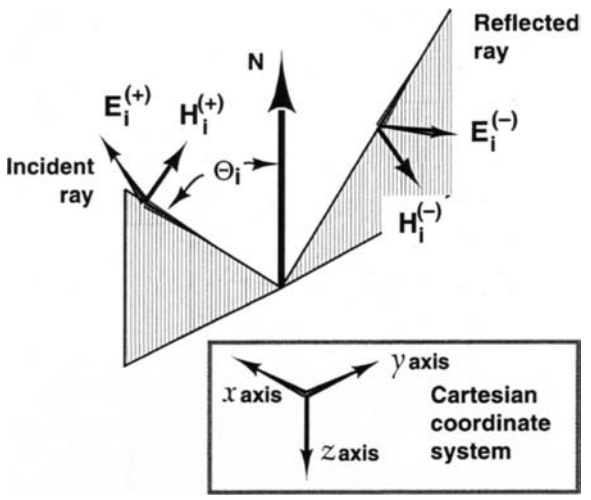


Fig. 2-11 A drawing of the right panel of Fig. 2-10, showing the *actual* directions of the electric field E and magnetic field H as the ray is reflected at a dielectric interface. N represents a normal to the surface at the interface.

The convention above is important when the differential phase shift upon reflection is measured on an ellipsometer, as discussed further in §2.3.6. At an interface the reflection coefficient r_i for the E field is

$$r_i \equiv \hat{E}_i^{(-)} / \hat{E}_i^{(+)} \tag{2-6a}$$

and the corresponding transmission coefficient $t_{i,i+1}$ is

$$t_{i,i+1} \equiv \hat{E}_{i+1}^{(+)} / \hat{E}_i^{(+)} \tag{2-6b}$$

The terms *amplitude reflection* and *amplitude transmission*, or simply *Fresnel coefficients*, also are used to identify r and t.

The foregoing pertain to a wave propagating from the ith to the $(i+1)^{th}$ medium. When the propagation is from the $(i+1)^{th}$ medium to the ith, the corresponding coefficients from the $(i+1)^{th}$ to the ith medium are

$$r_i' \equiv \hat{E}_{i+1}^{(+)} / \hat{E}_{i+1}^{(-)} \tag{2-6c}$$

and

$$t_{i+1,i} \equiv \hat{E}_i^{(+)} / \hat{E}_{i+1}^{(+)} \ . \tag{2-6d}$$

It is curious that four variables [appearing in Eqs. (2-6a), (2-6b), (2-6c) and (2-6d)] describe the reflection and transmission at an interface, whereas only two parameters, n_i and n_{i+1}, characterize the adjacent media. In fact, r_i, r_i', $t_{i,i+1}$ and $t_{i+1,i}$ are not independent. The continuity of the E field at the interface [see Eq. (2-2i)] requires that

$$1 + r_i = t_{i,i+1} \tag{2-6e}$$

and

$$1 + r_i' = t_{i+1,i} \ . \tag{2-6f}$$

At nonnormal incidence, the foregoing equations apply also for the TE wave. They do not apply for the TM wave if the definitions of r_p and t_p in Eqs. (2-9c) and (2-10a) are used, but are valid if Eqs. (2-10e) and (2-10f) are used. Other equations that follow are

$$r_i = -r_i' \tag{2-6g}$$

and

$$1 - r_i^2 = t_{i,i+1}\, t_{i+1,i} \ . \tag{2-6h}$$

2.2.3.2 Fresnel coefficients for the TE polarized wave

This is the *s* polarization — the E vectors are perpendicular to the plane of incidence. The continuity of the E, H fields described in Eqs. (2-2h) and (2-2i) requires that

$$\hat{E}_i^{(+)} + \hat{E}_i^{(-)} = \hat{E}_{i+1}^{(+)} \tag{2-7a}$$

$$\hat{H}_{y,i}^{(+)} + \hat{H}_{y,i}^{(-)} = \hat{H}_{y,i+1}^{(+)} \tag{2-7b}$$

where the subscript refers to the y component of the H field in the ith medium.

A *caveat* is that the actual directions of $\hat{E}_i^{(-)}$ or $\hat{H}_i^{(-)}$ must be different than those in the left panel of Fig. 2-10. Either E or H should reverse direction in the reflected wave to maintain the outward flow of radiant flux. The *actual* directions are shown in the right panel of Figs. 2-10 and 2-11 for a dielectric interface. In the incident medium there are both incident and reflected components. There is only a traveling wave component in the emergent medium. This boundary condition is discussed further in §2.4.2.4.

Due to the fact that the E in Eq. (2-7a) is the x component, H_y is found using Eq. (2-15b). Equation (2-7b) becomes

$$\hat{n}_i\, \hat{\gamma}_i\, (\hat{E}_i^{(+)} - \hat{E}_i^{(-)}) = \hat{n}_{i+1}\, \hat{\gamma}_{i+1}\, \hat{E}_{i+1}^{(+)} \tag{2-7c}$$

where Eqs. (2-3a), (2-3b) and (2-3d) have been differentiated with respect to z. The foregoing is quite general and is valid when the media on both sides of the interface are absorbing. When both media are nonabsorbing, this simplifies to

$$n_i \cos\theta_i (\hat{E}_i^{(+)} - \hat{E}_i^{(-)}) = n_{i+1} \cos\theta_{i+1} \hat{E}_{i+1}^{(+)} \ . \tag{2-7d}$$

At normal incidence each of the $\cos\theta$ terms is unity and this reduces to

$$n_i\, (\hat{E}_i^{(+)} - \hat{E}_i^{(-)}) = n_{i+1} \hat{E}_{i+1}^{(+)} \ . \tag{2-8a}$$

Equation (2-7d) is identical to Eq. (2-8a) if these substitutions are made

$$n_i \rightarrow n_i \cos\theta_i \ . \tag{2-8b}$$

This leads to the introduction of the effective index in §2.5.3. The reflection coefficient for the E field is found by solving for $\hat{E}_i^{(+)}$ and $\hat{E}_i^{(-)}$

$$r_s = \frac{n_i \cos\theta_i - n_{i+1} \cos\theta_{i+1}}{n_i \cos\theta_i + n_{i+1} \cos\theta_{i+1}} \tag{2-8c}$$

where the subscript s refers to the TE wave. t_s is obtained by eliminating $\hat{E}_i^{(-)}$ from Eqs. (2-7a) and (2-7d)

$$t_s = \frac{2\, n_i \cos\theta_i}{n_i \cos\theta_i + n_{i+1} \cos\theta_{i+1}} . \tag{2-8d}$$

The appropriate $\hat{\gamma}$ replaces $\cos\theta$ in the foregoing equation when either medium is absorbing. The result is

$$r_{s,i} = \frac{\hat{n}_i \hat{\gamma}_i - \hat{n}_{i+1} \hat{\gamma}_{i+1}}{\hat{n}_i \hat{\gamma}_i + \hat{n}_{i+1} \hat{\gamma}_{i+1}} \tag{2-8e}$$

where $\hat{\gamma}$ appears in Eq. (2-3h). At normal incidence, the amplitude reflection and transmission coefficients simplify to

$$r_i = \frac{n_i - n_{i+1}}{n_i + n_{i+1}} \tag{2-8f}$$

$$r_i' = \frac{n_{i+1} - n_i}{n_{i+1} + n_i} \tag{2-8g}$$

$$t_{i,i+1} = \frac{2\, n_i}{n_i + n_{i+1}} \tag{2-8h}$$

$$t_{i+1,i} = \frac{2\, n_{i+1}}{n_i + n_{i+1}} \tag{2-8i}$$

2.2.3.2.1 Actual directions of the electric vectors

The right panel of Fig. 2-10 and Fig. 2-11 show the *actual* directions of the electric and magnetic vectors in the reflected wave. It is assumed that the refractive index of the incident medium is less than that of the emergent medium, as would be the case if the incident medium were air and the emergent medium glass. Using the language of classical physical optics, this is termed a *rare-to-dense* reflection. The electric vector for the incident wave is along the positive x axis. For the reflected wave, its direction has "flipped" and is directed along the negative x axis,

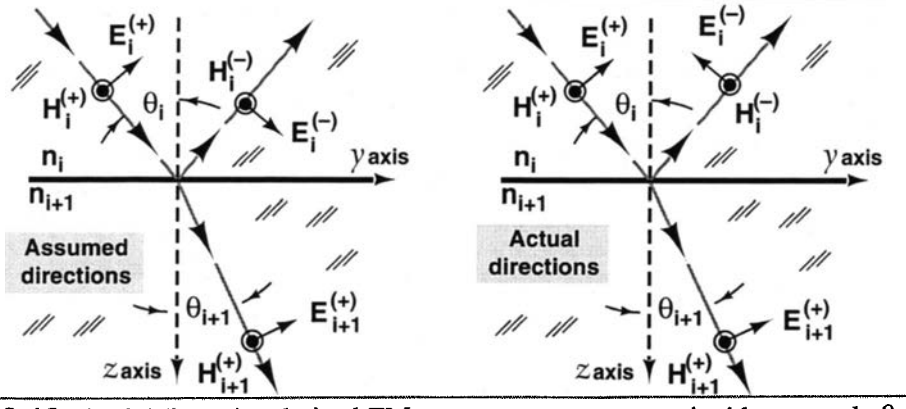


Fig. 2-12 A TM (i.e., p) polarized EM wave propagates at an incidence angle θ_i in a medium of index n_i and is reflected at the interface at the medium of index n_{i+1}. (*Left*) *Assumed* directions of the electric field E and magnetic field H. (*Right*) *Actual* field directions, $n_i < n_{i+1}$ and θ_i is less than the Brewster angle. Figure 2-13 provides another view of the E and H fields in the right panel.

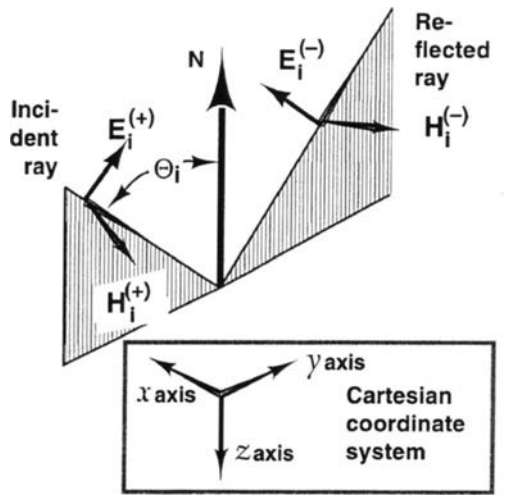


Fig. 2-13 A drawing of the right panel of Fig. 2-12, showing the *actual* directions of the electric field E and magnetic field H as the ray is reflected at a dielectric interface. N represents a normal to the surface at the interface.

2.2.3.3 Fresnel coefficients for the TM polarized wave

The left panel of Fig. 2-12 depicts the *assumed* directions of the E and H fields for the p polarized wave. Consider first the case where both media are nonabsorbing. It is then relatively simple to apply the boundary conditions. Figure 2-12 (*left panel*) shows that H is along the negative x axis; the tangential components of E are found from its projections upon the y axis.

If Eqs. (2-2h) and (2-2i) are to be satisfied, this requires that

$$\cos\theta_i(\hat{E}_i^{(+)} + \hat{E}_i^{(-)}) = \cos\theta_{i+1}\hat{E}_{i+1}^{(+)} \qquad (2\text{-}9a)$$

$$-\hat{H}_i^{(+)} - \hat{H}_i^{(-)} = -\hat{H}_{i+1}^{(+)} \qquad (2\text{-}9b)$$

An equation similar to Eq. (2-15c) that relates Eq. (2-9b) to the electric field is used. This results in

$$r_p = \frac{\dfrac{n_i}{\cos\theta_i} - \dfrac{n_{i+1}}{\cos\theta_{i+1}}}{\dfrac{n_i}{\cos\theta_i} + \dfrac{n_{i+1}}{\cos\theta_{i+1}}} \qquad (2\text{-}9c)$$

and

$$t_p = \frac{\dfrac{2\,n_i}{\cos\theta_{i+1}}}{\dfrac{n_i}{\cos\theta_i} + \dfrac{n_{i+1}}{\cos\theta_{i+1}}} . \qquad (2\text{-}10a)$$

The foregoing is valid at nonnormal incidence for nonabsorbing media. When the incident medium is absorbing, the equations for r and t are more complex and not derived.

Mahan [56] points out that his expression for t at a dielectric-metal interface differs from that of König [29]. To reproduce either of these two equations here would probably confuse rather than enlighten. Mahan, in particular, uses an "angle-dependent optical constant" for the emergent medium — a convention avoided here.

§2.2.3.2 cites r and t for the TE wave and demonstrates that the appropriate *effective indices* may be substituted into Eqs. (2-8c) and (2-8d) to convert relationships valid at normal incidence into the appropriate forms for nonnormal incidence. *Effective indices* may also be defined for the TM polarization as

$$n_i \rightarrow n_i / \cos\theta_i \ . \tag{2-10b}$$

The substitution of the foregoing into Eqs. (2-9c) and (2-10a) provides the correction expression for r but *not* for t.

An alternate approach is to express the r and t in terms of the tangential components of E, i.e., E_y. If Eqs. (2-2h) and (2-2i) are to be satisfied, this requires that

$$\hat{E}_{y,i}^{(+)} + \hat{E}_{y,i}^{(-)} = \hat{E}_{y,i+1}^{(+)} \ . \tag{2-10c}$$

The foregoing equation is substituted into one of Maxwell's equations — Eq. (2-69c) — which appears in an appendix. Equating the y component of Eq. (2-10c) results in

$$\hat{H}_{x,i}^{(\pm)} = \mp \Upsilon_0 \hat{n}_i (\hat{\gamma}_i)^{-1} \hat{E}_{y,i}^{(\pm)} \tag{2-10d}$$

where $\hat{\gamma}$ and Υ_0 appear in Eqs. (2-3h) and (2-15e), respectively. Writing a similar equation for the $(i+1)^{th}$ medium results in

$$r_p = \frac{\hat{n}_i / \hat{\gamma}_i - \hat{n}_{i+1} / \hat{\gamma}_{i+1}}{\hat{n}_i / \hat{\gamma}_i + \hat{n}_{i+1} / \hat{\gamma}_{i+1}} \tag{2-10e}$$

and

$$t_p'' \equiv \frac{2\,\hat{n}_i / \hat{\gamma}_i}{\hat{n}_i / \hat{\gamma}_i + \hat{n}_{i+1} / \hat{\gamma}_{i+1}} = \frac{\hat{E}_{y,i+1}^{(+)}}{\hat{E}_{y,i}^{(+)}} \ . \tag{2-10f}$$

Effective indices may be substituted into the normal incidence equations for r and t to provide the correct relationships at nonnormal incidence for the TM wave, provided Eq. (2-10f) defines the amplitude transmission coefficient t_p''.

2.2.3.3.1 Actual directions of the components of the electric vectors

The right panel of Fig. 2-12 and Fig. 2-13 show the *actual* directions of the electric and magnetic vectors for the reflected wave. It is assumed that the refractive index of the incident medium is less than that of the emergent medium, as would be the case if the incident medium were air and the emergent medium glass. Using the language of classical physical optics, this is termed a *rare-to-dense* reflection. The component of the electric field is along the positive y axis for the incident wave. For the reflected wave, the component of the electric vector along the negative y axis.

2.2.3.3.2 Definition of the Brewster angle

The numerator of Eq. (2-9c) is zero when the angle is

$$\theta_B = \theta_0 = \arctan(n_s / n_0) \tag{2-10g}$$

which is termed Brewster's angle. r_p is zero. Another consequence of Brewster's angle is that the direction of the y component of the E vector reverses when $\theta_0 > \theta_B$. This is evident when the right panel of Figure 2-12 is compared with Fig. 2-14. §2.3.6 discusses the effect of the Brewster angle on the differential phase shift upon reflection.

2.2.4 Radiant reflectance and transmittance

The radiant reflectance R and transmittance T are proportional to the radiant flux Φ (i.e., power in Watt) reflected and transmitted by an interface or a multilayer. This is sometimes called the *intensity* transmittance, or *energy* transmittance. It is a measure of the response of a thermal detector such as a thermopile.

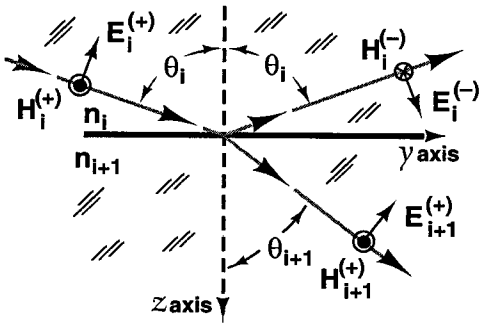


Fig. 2-14 The caption to Fig. 2-12 obtains, with the exception that the *actual* field directions are shown for the following conditions: $n_i < n_{i+1}$, θ_i exceeds the Brewster angle θ_B.

R and T are defined with reference to a specific example in which a flux detector is inserted into the incident, reflected and transmitted beams. As shown in Fig. 2-8, a beam of radiant flux is incident at an angle θ_i from a nonabsorbing medium. This is identified by the subscript i. The flux incident is

$$\Phi_i^{(+)} = \mathcal{E}_i^{(+)} \mathcal{A} \cos\theta_i \qquad (2\text{-}10h)$$

where $\mathcal{E}_i^{(+)}$ is the irradiance and $\mathcal{A}$ is the area of the irradiated surface.

The emergent medium is identified by the subscript j. In the example shown in Fig. 2-15, its index n_j is higher than that of the incident medium. It is also possible to bury a detector in that medium.

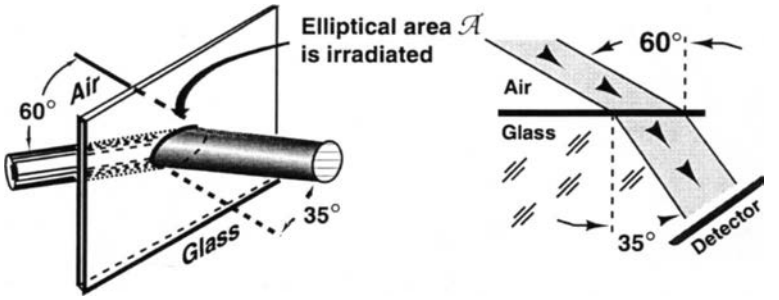


Fig. 2-15 A beam of radiant flux of circular cross section is incident at an angle θ_0 upon an area $\mathcal{A}$ at an interface between air and a medium of index 1.52. The cross sectional area of the beam changes. (*Left*) System in perspective. (*Right*) Cross section — a detector immersed in the emergent medium collects all of the radiant flux.

Even if the multilayer were nonabsorbing, the irradiances (i.e., flux per unit area) of the incident and emergent beams would be different because the beam width changes (see Fig. 2-15). The transmittance is the ratio

$$T \equiv \frac{\text{Response of detector inserted into emergent beam}}{\text{Response of detector inserted into incident beam}} \qquad (2\text{-}11a)$$

There are two definitions for T, depending upon the area of the detector relative to the beam area — which, in turn, determines the amount of flux the detector is allowed to collect. If the detector collects the total flux in both the incident and reflected beams,

$$T_\Phi \equiv \Phi_j^{(+)} / \Phi_i^{(+)} \ . \qquad (2\text{-}11b)$$

An alternative definition is

$$T_e \equiv \mathcal{E}_j^{(+)} / \mathcal{E}_i^{(+)} \ , \qquad (2\text{-}11c)$$

i.e., the ratio of the *irradiances*.

The transmittance T_Φ is valid if the detector area always exceeds the beam area. Thus, the total *flux* in each beam is collected, as shown in the right panel of Fig. 2-15. The transmittance T_e must be used if the areas of *both* beams exceed the area of the detector—the ratio of the irradiances determines the transmittance. If $n_i = n_j$, the two definitions are identical $T_\Phi = T_e$.

T_Φ is the conventional definition and is adhered to in this book for several reasons. First, most spectrophotometers are designed to collect total flux and thus measure T_Φ. Also, T_Φ provides a convenient radiant flux balance. Conservation of energy requires that for a nonabsorbing system

$$R_\Phi + T_\Phi = 1 \tag{2-11d}$$

where R_Φ is the radiant reflectance. The "Φ" subscript is dropped in the remainder of this book and Eq. (2-11d) is

$$R + T = 1 \ . \tag{2-11e}$$

R and T are expressed in terms of the Poynting vector $\vec{S}$ (see §2.10.2.4). Its scalar components in the incident, reflected and transmitted beams are $S_i^{(+)}$, $S_i^{(-)}$ and $S_j^{(+)}$, respectively. A unit vector $\vec{s}$ defines the direction of propagation (see Fig. 2-8). The angles between $\vec{s}$ and the z axis are θ_i, $-\theta_i$ and θ_j, respectively, for these components. The incident flux impinging upon the area $\mathcal{A}$ is $\Phi_i^{(+)} = \mathcal{A} S_i^{(+)} \cos\theta_i$. Similar relationships exist for $\Phi_i^{(-)}$ and $\Phi_j^{(+)}$. The radiant reflectance R and radiant transmittance T are

$$R \equiv \frac{S_i^{(-)} \mathcal{A} \cos\theta_i}{S_i^{(+)} \mathcal{A} \cos\theta_i} = \frac{S_i^{(-)}}{S_i^{(+)}} \tag{2-11f}$$

and

$$T \equiv \frac{S_j^{(+)} \mathcal{A} \cos\theta_j}{S_i^{(+)} \mathcal{A} \cos\theta_i} \ . \tag{2-12a}$$

The physical interpretation of the cosine factors in the two equations above is that R and T depend on the components of $\vec{S}$ normal to the interface. A consequence is that if the components of E and H tangential to the interface are known, their product provides directly the normal components of $\vec{S}$ and, hence, the R and T are readily obtained.

The absorption A is discussed in §1.2.2.3. Given the flux incident in the ith medium and the flux emergent into the jth medium, Eq. (1-3a) is cast into the form

$$A_{ij} = \frac{\Phi_i^{(+)} - \Phi_i^{(-)} - \Phi_j^{(+)}}{\Phi_i^{(+)}} = 1 - R - T \ . \tag{2-12b}$$

In the numerator, the reflected flux $\Phi_i^{(-)}$ and transmitted flux $\Phi_j^{(+)}$ are subtracted from the incident flux $\Phi_i^{(+)}$. The absorption results when that numerator is normalized by dividing by the incident flux.

There is no absorption at an interface. The foregoing statement can be misleading, however. It is true that Eq. (2-11e) is valid for the mathematically defined interface used as the model to compute the Fresnel coefficients (see §2.2.3.1). In practice, however, the interface between layers may contain a thin but highly absorbing

transition layer. This interfacial absorption was studied extensively in the development of coatings that withstand high irradiance from a laser beam without being damaged [Harrington 79].

Given the amplitude reflection coefficient $\hat{\rho}$ in a nonabsorbing medium [appearing in Eq. (2-16g)], the reflectance R [Eqs. (2-12a) and (2-71i)] is

$$R = \hat{\rho}\,\hat{\rho}^* = |\hat{\rho}|^2 \ . \tag{2-12c}$$

The R and T at the interface between the substrate (index n_s) and air (index n_0) are

$$R = r\,r^* = |r|^2 \qquad \text{and} \qquad T = n_s\,n_0^{-1}\,|t|^2. \tag{2-12d}$$

2.2.4.1 *R* and *T* at an interface

Although R and T at an interface may be derived from the equations in the foregoing sections, it is nevertheless useful to list explicit expressions. There are several formulations. The decision on which to use may depend upon the availability of complex arithmetic on a calculator or in a computer language. The incidence angle is θ_0 in the incident medium of index n_0. The optical constant of the emergent medium is $\hat{n}_s = n_s - jk_s$. This is substituted into Eq. (2-16h) and from Eq. (2-12d) at normal incidence

$$R = 1 - T = \frac{(n_0 - n_s)^2 + k_s^2}{(n_0 + n_s)^2 + k_s^2} = 1 - \frac{4\,n_s\,n_0}{(n_0 + n_s)^2 + k_s^2} \ . \tag{2-12e}$$

For a nonabsorbing substrate ($k_s = 0$), this simplifies to

$$R = 1 - T = \frac{(n_0 - n_s)^2}{(n_0 + n_s)^2} = 1 - \frac{4\,n_s\,n_0}{(n_0 + n_s)^2} \ . \tag{2-12f}$$

If $k_s = 0$ and $\theta_0 \neq 0$, effective indices may be substituted into the foregoing using Eq. (2-8b) or Eq. (2-10b) to find R_s and R_p, respectively.

If $k_s \neq 0$ and $\theta_0 \neq 0$, an expression involving complex arithmetic is

$$R_s = \left| \frac{1 - \zeta}{1 + \zeta} \right|^2 \tag{2-12g}$$

where

$$\zeta \equiv \sqrt{n_s^2 - k_s^2 - L^2 - j\,2\,n_s\,|k_s|} \quad (n_0 \cos\theta_0)^{-1} \tag{2-12h}$$

and the vertical lines denote the absolute value. The expression for R_p is similar

$$R_p = \left| \frac{1 - \zeta'}{1 + \zeta'} \right|^2 \tag{2-12i}$$

where

$$\zeta' \equiv \zeta\,n_0^2\ \hat{n}_s^{-2} \ . \tag{2-12j}$$

Querry [69] shows that the foregoing equations can be inverted, provided $\theta_0 \neq 45°$. Thus given R_s, R_p and θ_0, n_s and k_s are found. Using his notation,

$$R_s = \frac{(Q - \cos\theta_0)^2 + P^2}{(Q + \cos\theta_0)^2 + P^2} \tag{2-13a}$$

and

$$R_p = R_s\,\frac{(Q - \sin\theta_0 \tan\theta_0)^2 + P^2}{(Q + \sin\theta_0 \tan\theta_0)^2 + P^2} \tag{2-13b}$$

where

$$Q^2 - P^2 \equiv n_s^2\, n_0^{-2} - k_s^2\, n_0^{-2} - \sin^2\theta_0 \tag{2-13c}$$

and

$$P\,Q \equiv n_s\, k_s\, n_0^{-2} \ . \tag{2-13d}$$

When $n_0 = 1.00$, P and Q represent, respectively, the imaginary and real parts of the normalized surface admittance for s polarization. One fact that follows immediately is that at 45° incidence angle, $R_p = R_s^2$. Querry shows that, given R_p and R_s, the foregoing equations can be solved for P and Q. n and k follow from solving a quadratic equation derived from Eqs. (2-13c) and (2-13d).

Table 2-2
Normalized surface admittances of absorbing materials

	Surface admittance		Surface admittance	
θ_0	p polarization	s polarization	p polarization	s polarization
0°	$5.7000 - j\,1.2500$	$5.7000 - j\,1.2500$	$0.6300 - j\,0.4700$	$0.6300 - j\,0.4700$
30°	$5.7210 - j\,1.2453$	$5.6790 - j\,1.2546$	$0.7254 - j\,0.3365$	$0.5113 - j\,0.5791$
60°	$5.7636 - j\,1.2356$	$5.6368 - j\,1.2640$	$0.6761 - j\,0.0758$	$0.3541 - j\,0.8363$
85°	$5.7800 - j\,1.2306$	$5.6164 - j\,1.2686$	$0.6150 - j\,0.0153$	$0.3100 - j\,0.9552$
0°	$0.1800 - j\,3.6400$	$0.1800 - j\,3.6400$	$1.7900 - j\,1.8600$	$1.7900 - j\,1.8600$
30°	$0.1816 - j\,3.6062$	$0.1783 - j\,3.6741$	$1.8225 - j\,1.8242$	$1.7568 - j\,1.8952$
60°	$0.1845 - j\,3.5413$	$0.1751 - j\,3.7414$	$1.8809 - j\,1.7481$	$1.6925 - j\,1.9672$
85°	$0.1857 - j\,3.5111$	$0.1736 - j\,3.7736$	$1.9059 - j\,1.7094$	$1.6624 - j\,2.0028$

Normalized surface admittance [see Eqs. (2-16c) and (2-16d)] of four media, the optical constants of which appear at 0°. The angle of incidence θ_0 is measured in air.

2.2.4.2 Drude approximation for the amplitude reflection coefficient *r*

At the beginning of the last century, Drude [1900] showed that approximate equations for r_p and r_s are obtained when L in Eq. (2-12h) can be neglected in comparison to the other terms. In this case, the surface admittance (§2.2.5.4) at non-normal incidence is assumed to be identical with the optical constant $\hat{n}_s$:

$$r_s \approx \frac{n_0 \cos\theta_0 - \hat{n}_s}{n_0 \cos\theta_0 + \hat{n}_s} \tag{2-13e}$$

and

$$r_p \approx \frac{n_0 \sec\theta_0 - \hat{n}_s}{n_0 \sec\theta_0 + \hat{n}_s} \ . \tag{2-13f}$$

The data in Table 2-2 illustrate the validity of the approximation. Listed also in the Table is the *actual* normalized surface admittance [see Eqs. (2-16c) and (2-16d)]. The medium with $\hat{n} = 5.7 - j\,1.25$ is typical of germanium in the visible spectral region. Its $|\hat{Y}|$ differs from $\hat{n}_s$ by only a few percent at 85° incidence. For silver ($\hat{n} = 0.18 - j\,3.64$), $|\hat{Y}|$ changes by only 4% as θ_0 ranges from 0° to 85°. Figure 2-16 shows that for nickel,

the change is greater—7%. Nevertheless, Eqs. (2-13e) and (2-13f) are reasonable approximations.

An exception is when $n < 1$ and $k < 1$, as occurs for metals in the vacuum UV. Figure 2-17 shows the surface admittance for such a case. At large angles of incidence, the imaginary part of the surface admittance doubles for s polarization and plunges to quite small values for p polarization. The Drude approximation unusable.

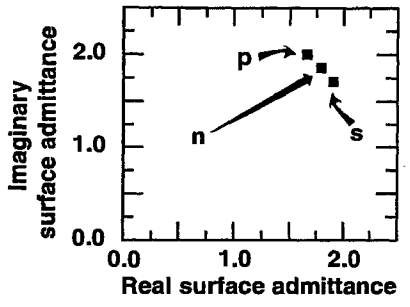


Fig. 2-16 For a metal of optical constant $\hat{n} = 1.79 - j\,1.86$, its surface admittance [see Eqs. (2-16c) and (2-16d)] is graphed at normal incidence n and for p and s polarization at 85° incidence.

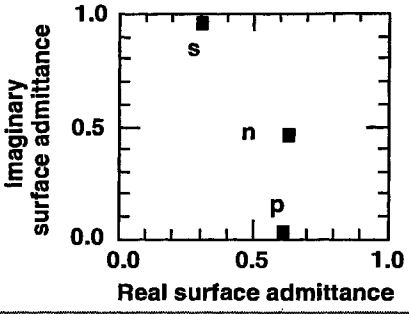


Fig. 2-17 The caption to Fig. 2-16 obtains, with the exception that the optical constant is $\hat{n} = 0.63 - j\,0.47$.

Equations (2-13e) and (2-13f) are a special case of Eqs. (2-16f) and (2-16g), with just an incident medium and substrate of optical constant $\hat{n}$, with $\hat{n}_i\,\hat{\gamma}_i = n_0 \cos\theta_0$, $\hat{n}_i / \hat{\gamma}_i = n_0 \sec\theta_0$ and $\hat{Y} = \hat{n}$.

2.2.4.3 Fluence and energy density

Multilayers are often irradiated with laser pulses of short duration — 10^{-9} seconds and less. Under these circumstances there is little opportunity for any substantial amount of heat generated by the pulse to exit by means of thermal conduction. A consequence is that the damage is proportional to areal energy density or *fluence*. Notwithstanding the preference for the SI unit of J/mm^2, the convention at the time of this writing is to measure fluence in J/cm^2, as plotted on the abscissa of Fig. 1-41.

2.2.4.4 Attenuation of the irradiance of a plane wave

The objective is to find the attenuation of the irradiance of a plane wave in a medium in which $k << 1$, i.e., a weakly absorbing material. In this case, $\hat{\gamma}_i$ is approximated well by $\cos\theta_i$ and the Poynting vector is nearly coaxial with the wave front normal. According to Eq. (2-71i), at $z = 0$ in the ith medium the irradiance is proportional to

$$S_i = \mathcal{E}_i = \tfrac{1}{2}\,\Re\left[\hat{E}_i^{(+)}\,(\hat{H}_i^{(+)})^*\right] \tag{2-13g}$$

where $\hat{E}_i$ appears in Eq. (2-3d). $\hat{H}_i$ is the amplitude of the component of the H vector at $z = 0$, as obtained from Eq. (2-15b) with the subscript i for "zero." At a distance z below the origin and lateral distance y from the origin, S is

$$S_i' = \mathcal{E}_i' = \tfrac{1}{2}\Re\left\{\hat{E}_i^{(+)} \exp[-j\xi\,(n_i - jk_i)]\left[H_i^{(+)}\right]^* \exp[j\xi\,(n_i + jk_i)]\right\} \tag{2-13h}$$

where

$$\xi \equiv (z\cos\theta_i + y\sin\theta_i)\,2\pi/\lambda\,. \tag{2-13i}$$

From geometry, $y \sin\theta_i = z \tan\theta_i \sin\theta_i$ and ξ reduces to

$$\xi = z \sec\theta_i \; 2\pi/\lambda \; . \tag{2-13j}$$

In other words, the distance z normal to the surface is increased by "$\sec\theta_i$" to obtain the slant distance. The absorption coefficient α_i is defined as

$$S_i'/S_i = \mathcal{T}_i'/\mathcal{T}_i = \exp(-\alpha_i h \sec\theta_i) \tag{2-14a}$$

where h is the distance along a normal to the surface the wave propagates and θ_i is the angle of refraction in that medium. Substituting S and S' from Eqs. (2-13g) and (2-13h) relates the bulk absorption coefficient to the extinction coefficient

$$\alpha_i \equiv 4\pi |k_i| / \lambda \; . \tag{2-14b}$$

As an example, suppose that an absorbing Schott glass is used as a substrate at a wavelength λ of 500 nm (i.e., 5×10^{-4} mm). Its thickness h is 2 mm and it has an internal transmittance of 10^{-4}. What is its k? Taking the natural log of Eq. (2-14a) and solving Eq. (2-14b) for k provides

$$|k| = -(4\pi h)^{-1} \lambda \ln(10^{-4}) = -(4\times\pi\times 2)^{-1} \times (5\times10^{-4}) \times (-9.21) = 1.8\times10^{-4} \tag{2-14c}$$

and hence is small compared to one. The conclusion is that it would be an acceptable approximation to consider the substrate to be "dielectric."

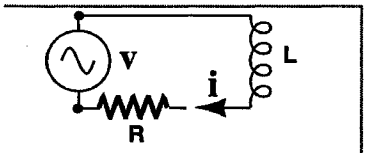


Fig. 2-18 An AC voltage generator is connected in series to an inductor L and resistor R. v_1 and i_1 are, respectively, the voltage and current.

At normal incidence, i.e., $\theta_i = 0$, and when h is the "standard thickness," the "$\exp(-\alpha_i h)$" in Eq. (2-14a) is the "internal transmittance" T_0 mentioned in the first paragraph of §10.3.4.2.

2.2.5 Analogy between admittance and an optical coating

This section borrows from a concept of electrical engineering — admittance. This includes: (a) Admittance of an electrical circuit; (b) admittance of a wave in free space; (c) admittance of an electrical transmission line; and (d) admittance of a surface and/or multilayer.

2.2.5.1 Admittance of an electrical circuit

Figure 2-18 depicts a steady-state ac generator (i.e., its output vs time is a sinusoid) connected to a circuit consisting of an inductor (of inductance L) in series with a resistance R. When driven at a steady-state angular frequency ω, the load admittance $\hat{Y}$ is the ratio of the current i to voltage v

$$\hat{Y} \equiv i/v = \frac{1}{R + j\,\omega L} = \frac{R - j\,\omega L}{R^2 + L^2\,\omega^2} \; . \tag{2-15a}$$

The fact that $\hat{Y}$ is complex implies that the current is not in phase with the voltage.

2.2.5.2 Admittance of a wave propagating in free space

Although Fig. 2-7 shows the E and H fields of an electromagnetic EM wave, previous sections have focused upon the electric field E—the magnetic intensity H has been largely ignored. Figure 2-19 shows the directions of the E and H fields of a wave propagating in the +z direction.

An appendix (§2.10.2) derives an expression for the E field of an EM wave. In the same way that the H field is eliminated from Eqs. (2-69b) and (2-69c) to provide E in Eq. (2-69g), the propagation equation for H could be obtained. It can be shown that it is similar to Eq. (2-70i), which in a nonabsorbing medium is

$$\hat{H}_x = \hat{H}_0 \exp[j(\omega t - kz)] \ . \tag{2-15b}$$

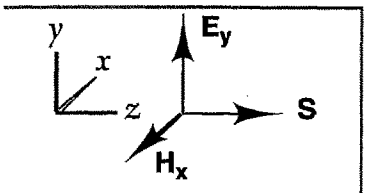


Fig. 2-19 In a Cartesian coordinate system, directions of the electric field E and magnetic field H of an electromagnetic wave propagating in the +z direction, which is collinear with S.

However, $\hat{H}_0$ in the foregoing equation and $\hat{E}_0$ in Eq. (2-70i) are related when a plane wave is propagating in free space. As shown in Fig. 2-19, the electric field has a component E_y and no others. Equation (2-69b) then reduces to

$$-\frac{\partial E_y}{\partial z} = -\mu\,\mu_0 \frac{\partial H_x}{\partial t} \ . \tag{2-15c}$$

Substituting Eqs. (2-70i) and (2-15b) into Eq. (2-15c) leads to

$$\hat{E}_0 = \mathcal{Z}_0 \sqrt{\frac{\mu}{\varepsilon}}\, \hat{H}_0 \tag{2-15d}$$

where

$$\mathcal{Z}_0 \equiv \sqrt{\mu_0/\varepsilon_0} \equiv (\Upsilon_0)^{-1} \tag{2-15e}$$

which is the impedance of free space—377 ohms in MKS units. Using the fact that $\mu = 1$ at optical frequencies, Eq. (2-15d) becomes

$$\hat{H}_0 = \hat{n}\,\Upsilon_0 \hat{E}_0 \tag{2-15f}$$

where Υ_0 is the reciprocal of $\mathcal{Z}_0$. The units of Υ_0 are $(377)^{-1}$ siemens.

As discussed in §2.10.2.4, complex notation is used for the amplitudes of electric field $\hat{E}$ and magnetic field $\hat{H}$. Circumflexes denote complex variables. A general form of Eq. (2-15f) is

$$\hat{Y}' \equiv \hat{H}_0 / \hat{E}_0 \tag{2-15g}$$

and defines the admittance (in units of siemens) at optical frequencies. The $\hat{H}_0$ and $\hat{E}_0$ are measured at an arbitrary reference point in space. If $\hat{Y}'$ is not pure real, the physical interpretation is that $\hat{H}$ and $\hat{E}$ are not in phase. When the actual field strength E (in units of volt/m) is to be found, it is the *real part of* (using the symbol $\Re$) Eq. (2-70i).

The symbol $\hat{Y}$ (*without* the prime) used in the remainder of this book is the *normalized* admittance, which is the admittance divided by the admittance of free space

$$\hat{Y} \equiv \hat{Y}' / \Upsilon_0 \equiv Y_r + jY_j \tag{2-16a}$$

where Y_r and Y_j are its real and imaginary parts, respectively. Normalized admittance is dimensionless. The advantage of its use (in lieu of actual admittance) is that many equations are simplified. A simple example is that an absorbing emergent medium of optical constant $\hat{n}$ has an admittance of $\hat{n}\Upsilon_0$ siemens and a *normalized* admittance of $\hat{n}$. Normalized admittance is used in Chaps. 4, 6, 7 and 8 to design multilayer reflection reducing coatings, beamdividers and bandpass filters.

2.2.5.3 Analogy—electrical transmission line and optical coating

2.2.5.3.1 Characteristic admittance of a transmission line

A simple form of an electrical transmission line is a pair of parallel wires. Suppose that this line is sufficiently long such that a reflected pulse is not returned. A short pulse of

voltage dumped into its input. This voltage pulse (and the accompanying current) propagates along the line. The voltage may be measured with a fast oscilloscope. The current could also be measured by measuring the voltage drop across a small resistor in the line.

Table 2-3
Analogies — optical coating and electrical transmission line

Optical coating	Transmission line	Waveguide
Electric field	Voltage	Electric field
Magnetic field	Current	Magnetic field
Coherent source of radiant flux	Driver — sinusoidal oscillator	Klystron — the generator
Refractive index of incident medium	Characteristic admittance of the generator	Characteristic admittance of the generator
Refractive index of substrate	Admittance of resistive load	Admittance of resistive load
Metallic substrate	Lossy inductive load	Lossy termination of guide
Refractive index of a layer	Characteristic admittance of a transmission line section	Characteristic admittance of a waveguide section
Optical thickness	Electrical length	Electrical length
**	Shunt section	Shunt section
Index of an equivalent layer	Image admittance — a symmetrical two port lossless network	Image admittance — same as transmission line
FTR layer (see §2.4.2.8)		Iris or waveguide section operating beyond cutoff
Absorbance of coating§	Insertion loss, dB	Insertion loss, dB
Beamdivider	Four-port lossless network	Magic tee

**Does not exist. § |dB| /10 = absorbance

Analogies between optical coatings, electrical transmission lines and waveguides.

The ratio of the current to the voltage (which has units of admittance) is constant. This characteristic admittance is determined by the geometry of the line and the dielectric constant of the material surrounding it.

2.2.5.3.2 Analogies

Analogies between an electrical transmission line and an optical thin film appear in the publications of Muchmore [48], Dorgelo [48], Leurgans [51], Altenburg [53], Iwata [53], Seeley [61], Young [61c], Knittl [76], Chen [88] and Thelen [89]. The electrical transmission line theory of Ribblet [57], for example, has been applied to design optical antireflection coatings [Young 61c, Thelen 89]. Texts on microwave theory, such as Collin [72], summarize microwave design techniques. Table 2-3 summarizes the analogies, including a microwave waveguide. The optical analog of the load admittance is the optical constant $\hat{n} = n - j\,k$ of the substrate. A lossy inductor [see Eq. (2-15a)] is analogous to the use of a metallic substrate for an optical coating.

2.2.5.4 Reflection coefficients in terms of surface admittance

The amplitude reflection coefficient may be written in terms of the normalized surface admittance $\hat{Y}$

$$\hat{Y} = \hat{H} / (\hat{E}\,\Upsilon_0) \equiv Y_r + jY_j \tag{2-16b}$$

where $\hat{E}$ and $\hat{H}$ are the components parallel to the surface (i.e., interface), regardless of which polarization is chosen. $\hat{Y}$ is dimensionless. At normal incidence, the normalized admittance in Eq. (2-16a) and the surface admittance are identical. At nonnormal incidence the surface admittance of a medium of optical constant $\hat{n}$ is

$$\hat{Y} = \hat{n}/\hat{\gamma} \tag{2-16c}$$

for the *p* polarization and for the *s* polarization

$$\hat{Y} = \hat{n}\,\hat{\gamma}\ . \tag{2-16d}$$

The *internal* reflection coefficient $\hat{\rho}$ in an absorbing medium (of optical constant $\hat{n}_i$) is defined at normal incidence as

$$\hat{\rho} = \frac{\hat{n}_i - \hat{Y}}{\hat{n}_i + \hat{Y}}\ , \tag{2-16e}$$

at nonnormal incidence for *s* polarization as

$$\hat{\rho}_s = \frac{\hat{n}_i\hat{\gamma}_i - \hat{Y}}{\hat{n}_i\hat{\gamma}_i + \hat{Y}} \tag{2-16f}$$

and for *p* polarization as

$$\hat{\rho}_p = \frac{\hat{n}_i/\hat{\gamma}_i - \hat{Y}}{\hat{n}_i/\hat{\gamma}_i + \hat{Y}}\ . \tag{2-16g}$$

The symbol *r* is restricted to represent the reflection coefficient only at an interface. $\hat{\rho}$ is applicable to the more general case and can represent the reflection *inside* a multilayer at a point where the normalized admittance is $\hat{Y}$.

Consider the case where the admittance is *exterior* to a multilayer—i.e., at the surface adjacent to the incident medium. Then if n_0 is the incident medium index, the $\hat{\rho}$ appearing in Eqs. (2-16e), (2-16f) and (2-16g) is termed the *external* amplitude reflection coefficient. Although the above equations are not rewritten, at normal incidence we have

$$\hat{\rho}]_{external} = \frac{n_0 - \hat{Y}}{n_0 + \hat{Y}} \tag{2-16h}$$

2.2.5.5 Reflection coefficients in terms of sheet resistance

In appendix §2.11.5, an equation is shown that obtains at longer wavelengths in the infrared. The reflectance of a conducting thin lamina may be expressed in terms of its sheet resistance (see Eq. (1-23).

2.2.6 Addition of waves at an interface

This section considers the reflection of waves an interface between two nonabsorbing media. §2.2.7 extends this to the two interfaces of a single layer.

2.2.6.1 Waves at an air-glass interface

It is important to distinguish between the *components* of a wave and the *total amplitude* of a wave. A *component* is the part traveling in either the "plus zee" or "minus zee" direction. The total amplitude is the net sum (at any given instance of time) of the components.

An example is the wave reflected from an air-glass interface. The index of the glass is 1.52. From Eq. (2-8f) the amplitude reflection coefficient is

$$r = (1 - 1.52) / (1 + 1.52) = -0.206 \ . \qquad (2\text{-}17a)$$

The minus sign implies a "rare-to-dense" reflection. This terminology means that the index of the incident medium is less than that of the glass. The equation for the reflected wave is similar to Eq. (2-2a), with $k = 0$, since the medium is nonabsorbing.

$$E^{(-)}(z,t) = r\, E_0^{(+)} \cos(2\pi \nu t + 2\pi z n / \lambda) \ . \qquad (2\text{-}17b)$$

λ is the vacuum wavelength. Figure 2-20 plots three sets of curves—the electric field vs distance. First, consider panel (a), which shows the two components. At the interface and at this instance of time, the amplitude of the incident wave (*long dash*) is negative and the reflected wave (*short dash*) is on the opposite side of the z axis because r [see Eq. (2-17a)] is negative. The *solid* curve is the *total amplitude*, which is the algebraic sum of the incident and reflected components, which is $1 - |r| = 1 - 0.206 = 0.79$.

The waves in panel (b) of Fig. 2-20 are at a later time. The peak of the incident wave has moved towards the interface and the reflected wave has moved away from it. Finally, panel (c) of Fig. 2-20 shows the waves at a quarter cycle later in time. The peaks of the reflected and incident waves are on the same side of the z axis. In contrast to the total amplitude found in the previous paragraph, the net amplitude is now

$$1 + |r| = 1 + 0.206 = 1.206.$$

The *maximum* of the total amplitude is positioned at $\lambda_0/4$ *in front* of the glass.

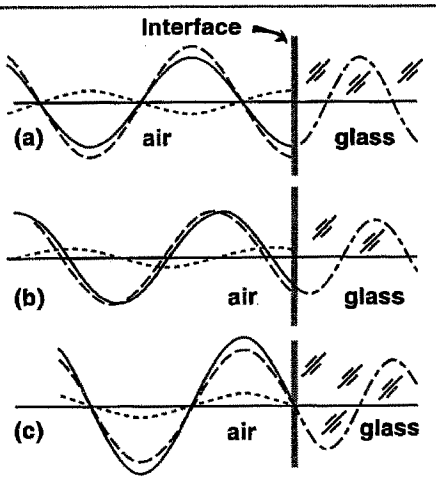


Fig. 2-20 A light wave impinges from the left upon an air-glass interface.

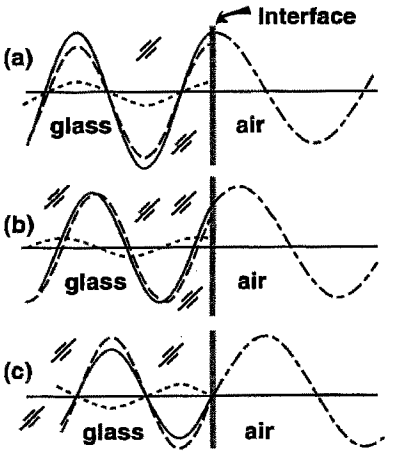


Fig. 2-21 A light wave impinges from the left upon a glass-air interface.

Both of the above Figs. plot the amplitude of the incident wave component (*long dash*), reflected wave component (*short dash*) and transmitted wave (*long-short dash*). The *solid* curve is the total amplitude of the wave. (a) Earlier instance of time; (b) later time; and (c) latest time.

2.2.6.2 Waves at a glass-air interface

Figure 2-21 depicts the amplitudes of the wave that reflect from a glass-air interface. The Fresnel amplitude reflection coefficient has the same magnitude as that in Eq. (2-17a),

but opposite sign. r is positive. In panel (a) of Fig. 2-21, at the interface the reflected wave component (*short dash*) is on the *same* side of the z axis as the incident component (*long dash*). The maximum of the total amplitude (*solid*) is at the glass surface. This is to be contrasted with its position when there is an air-to-glass reflection—this is a quarterwave in front of the surface

Panel (b) shows the waves at a later time and panel (c) at a quarter cycle later in time. In all cases, the amplitude of the wave in the emergent medium (which is air) does not change because it is a traveling wave. In contrast, there is a standing wave on the glass side of the interface.

2.2.6.3 Total electric field amplitude at an interface

Figure 2-22 depicts the *total* amplitude of a wave at two instances of time. It is reflecting at an interface between two nonabsorbing materials. A priori, it is not known whether the incident medium is *medium* 1 or *medium* 2. The following facts help to establish the direction of propagation and relative refractive indices:

- The amplitude of the wave in the emergent medium is independent of time.
- The ratio of the wavelengths in the two media establishes the refractive index ratio.

As an example, the emergent medium must be *medium* 1 in Fig. 2-22, since the amplitude changes in *medium* 2. A scaling of the wavelengths shows that the index ratio is 2.4:1. In other words, the light source is in air (on the right) and on the left side of the interface is a higher index material, such as zinc selenide (index = 2.40).

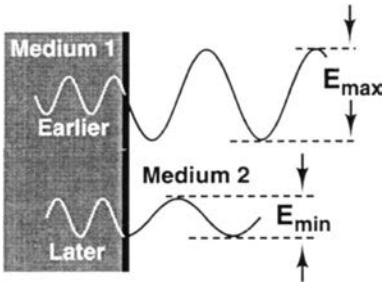


Fig. 2-22 Total electric field at two instances of time for a wave traveling from *medium* 2 (of refractive index 1.00) to *medium* 1 (of refractive index 2.40).

2.2.7 Waves reflected at interfaces of a layer

Figure 1-7 shows the waves reflected from a single layer on a substrate of higher index glass (n_s = 1.65). In one case the layer of index 1.38 reduces the reflectance, whereas a relatively higher index layer (n = 1.85) enhances the reflectance. The analyses in §2.2.7.1 and §2.2.7.2 are for a layer of quarterwave optical thickness—the incident wavelength is λ_0 and the optical thickness of the film is $\lambda_0/4$.

As discussed below, the *net* reflected wave consists of the sum of the waves that reflect at each interface. It is the amplitudes and phases of these waves that determine whether the reflectance is enhanced or decreased.

2.2.7.1 Low index layer

A layer of index 1.38 is deposited upon the glass substrate of index 1.65. The first step is to find the amplitude reflection coefficients at the two interfaces. At the air-film interface the amplitude reflection coefficient is

$$r_0 = (1 - 1.38) / (1 + 1.38) = -0.16 \qquad (2\text{-}17c)$$

and at the film-glass interface it is

$$r_1 = (1.38 - 1.65) / (1.38 + 1.65) = -0.09 \ . \qquad (2\text{-}17d)$$

Both r_0 and r_1 are negative because these are "rare-to-dense" reflections.

The reflected wave at the air-film interface is shown as the *long-dash* curve in Fig. 2-23. The incident wave (*solid* curve) penetrates the magnesium fluoride and reflects at the film-glass interface. According to Eq. (2-7d), the amplitude of the reflected wave (represented by a *short-dash* line in Fig. 2-23) is 9% of that of the incident wave. It may seem strange that the reflected wave is on the same side of the z axis as the incident wave. This is what was shown to be true in §2.2.6.1.

The salient point is that the waves that reflect from the two interfaces are 180° *out of phase* with each other. This is termed "destructive interference" in books on physical optics. The net effect is to *reduce* the net amplitude of the reflected wave. The net amplitude ρ of the reflected wave is approximately

$$\rho \approx r_0 - r_1 \approx (-0.16) - (-0.09) \approx -0.07 \ . \qquad (2\text{-}17e)$$

The negative sign indicates that the phase shift upon reflection is 180°. The radiant reflectance is computed below.

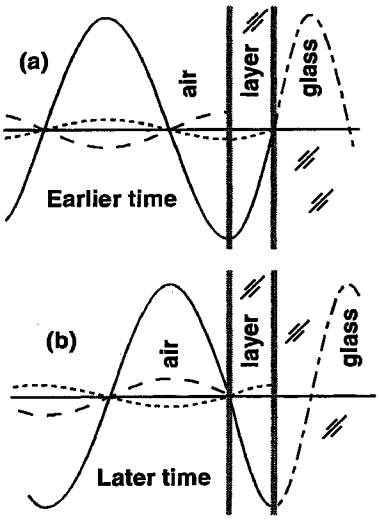


Fig. 2-23 A light wave impinges from the left upon glass coated with a layer of refractive index 1.38.

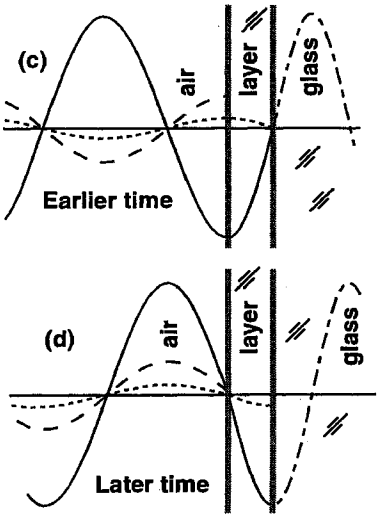


Fig. 2-24 A light wave impinges from the left upon glass coated with a layer of refractive index 1.85.

The Figs. above plot the amplitude of the incident wave component (*solid*), component reflected at the air-film interface (*long dash*), reflected component at the film-glass interface (*short dash*) and wave (*long-short dash*) transmitted into the glass substrate of index 1.65. The layer optical thickness is λ/4. (Panels *a* and *c*) At an earlier instance of time. (Panels *b* and *d*) At latest time.

2.2.7.2 High index layer

In this instance, a layer of index 1.85 is deposited on a substrate of low index. The amplitude reflection coefficient at the air-film interface is

$$r_0 = (1 - 1.85)/(1 + 1.85) = -0.30 \qquad (2\text{-}17f)$$

and at the film-glass interface it is

$$r_1 = (1.85 - 1.65)/(1.85 + 1.65) = +0.06 \ . \qquad (2\text{-}17g)$$

r_1 is positive because it is a "dense-to-rare" reflection.

As in the case of §2.2.7.1, the reflection from the air-film interface (*long-dash* in Fig. 2-24) is "rare-to-dense." The lower panel of Fig. 2-24 shows that the reflected wave (*short-dash*) at the film-glass interface is on the *same side* of the z axis. This causes it to interfere *constructively* with the wave reflected from the air-film interface.

$$\rho \approx r_0 - r_1 \approx (-0.30) - (+0.06) \approx -0.36. \qquad (2\text{-}17h)$$

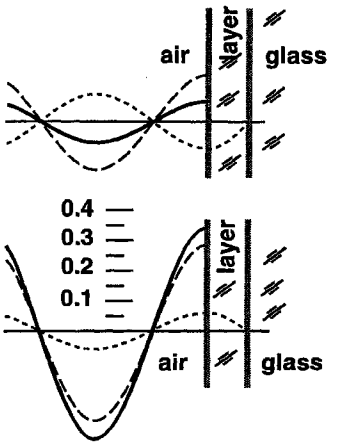


Fig. 2-25 The captions to Figs. 2-23 and 2-24 obtain, with the exceptions that the incident wave is omitted and the net reflected wave is *solid*. The vertical scale is the amplitude.

2.2.7.3 Comparison of radiant reflectances

Figure 2-25 redraws Figs. 2-23 and 2-24 on an expanded scale, showing only the reflected wave components (both *dashed*) and the net amplitude of the reflected wave (*solid*). The magnitude of the latter can be determined from the scale. The radiant reflectance is obtained from Eq. (2-12c) and is

$$R \approx (-0.06)^2 \approx 0.005 \approx 0.5\% \qquad (2\text{-}17i)$$

for the low index layer ($n = 1.38$) and

$$R \approx (-0.36)^2 \approx 0.013 \approx 13\% \qquad (2\text{-}18a)$$

for the 1.85 index layer.

2.3 Phase shift upon reflection and node of the standing wave

The phase shift upon reflection is important for several reasons. The first is when two multilayers are added together. The phase shift determines whether the net reflectance of the composite stack is larger or smaller.

The phase shift is also important when multilayers are incorporated into an interferometer, laser cavity or liquid crystal display system. The concept of a standing wave provide a useful way of visualizing the phase.

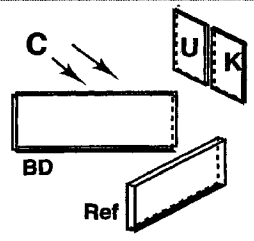


Fig. 2-26 The phase shift upon reflection of an unknown U sample is measured by comparing it with the phase of a known material K. The displacement of the fringes produced by a Twyman-Green interferometer is measured.

2.3.1 Definition of phase shift

r and t are the symbols representing the amplitude reflection and amplitude transmission coefficients at an interface. In the general case, $\hat{\rho}$ is used as the amplitude reflection coefficient for either a multilayer or interface. This defines the phase shift upon reflection δ_i. In the ith medium it is written

$$\hat{\rho}_i \equiv |\hat{\rho}_i| \exp(j\,\delta_i) \equiv \frac{\hat{E}_i^{(-)}}{\hat{E}_i^{(+)}} \,. \qquad (2\text{-}18b)$$

The sign of δ has been discussed in reference to ellipsometry; Muller [69] summarizes the sign conventions. Because t in Eq. (2-69a) and δ have the same sign, an increase in δ is equivalent to an advance in phase. §2.3.6 discusses the sign conventions that appear in Born & Wolf [75] and similar books relating to ellipsometry.

The phase shift is important when a multilayer is combined coherently with another multilayer. Figure 2-26 shows one of the many ways phase shift is measured. The multilayer is inserted as a reflector in one arm of a Twyman-Green interferometer. The surface of a reflector of known phase shift (for example, an air-dielectric interface) is positioned so that it is coincident with the surface of a multilayer. The phase is determined from the shift of the fringes. Interferometers such as the Fabry-Perot [Stanley 64], Fizeau [Jenkins 57b, 59], FECO [Bennett 64] and Koster [Barakat 63] have been used as well.

2.3.1.1 Explicit equations for the phase shift

The amplitude reflection coefficient, whether computed at a surface or a multilayer, may be expressed as the quotient of complex numbers

$$\hat{\rho} = \frac{a + jb}{c + jd} \quad . \tag{2-18c}$$

It follows that δ is

$$\delta = \arctan\left(\frac{bc - ad}{ac + bd}\right) . \tag{2-18d}$$

The proper quadrant for δ is found by examining the signs of the numerator and denominator. An example is that a negative numerator and negative denominator in Eq. (2-18d) place δ in quadrant III.

In the special case of an absorbing substrate at normal incidence, $a \equiv n_0 - n_s$, $b \equiv |k_s|$, $c \equiv n_0 + n_s$ and $d \equiv -|k_s|$. This simplifies to

$$\delta = \arctan\left(\frac{2\, n_0\, |k_s|}{n_0^2 - n_s^2 - k_s^2}\right) . \tag{2-18e}$$

2.3.2 Standing waves normal to an interface

The phase shift upon reflection may be visualized using the concept of standing waves. Given a reflector characterized by a $\hat{\rho}$ defined in Eq. (2-18b), the total E field in front of it is the sum of the incident wave and reflected wave

$$E(z) = \Re\left\{\hat{E}_0^{(+)} [\exp(-jk\hat{\gamma}z) + |\hat{\rho}| \exp(j\delta) \exp(jk\hat{\gamma}z)]\right\} \quad . \tag{2-18f}$$

$\hat{\gamma}$ reduces to $\cos\theta$ when the medium is nonabsorbing. Manipulation of Eq. (2-18f) yields

$$E(z) = \Re\left\{\hat{E}_0^{(+)} \exp(j\delta/2)\, [(1 - |\hat{\rho}|) \exp(-j\beta") + 2|\hat{\rho}| \cos\beta"]\right\} \tag{2-18g}$$

where

$$\beta" \equiv 2\pi\sigma n z + \tfrac{1}{2}\delta \quad . \tag{2-18h}$$

The formulation in Eq. (2-18g) presents the physical picture. The first term is proportional to $1 - |\hat{\rho}|$ and represents the traveling wave component. If the H field were found and the irradiance computed using Eq. (2-71i), it can be shown that the Poynting vector component S is independent of z. The remaining term is proportional to $\hat{\rho}$. It is identified as the standing wave amplitude. At a given position z in space, this amplitude varies between zero and $2|\hat{\rho}|$, depending upon the phase shift δ. When $\hat{\rho} = 0$ there is no reflection and no standing wave component exists. As $|\hat{\rho}| \rightarrow 1.0$, the traveling component vanishes.

The concept of a standing wave is important for several reasons. First, it helps to visualize the phase shift upon reflection. Second, the standing wave is increasingly important in laser damage studies.

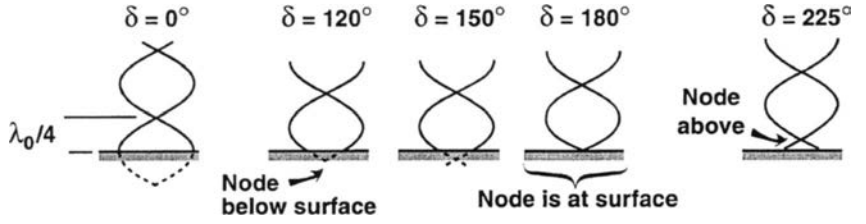


Fig. 2-27 Standing wave electric field of a light wave incident upon an interface characterized by a phase shift upon reflection δ [see Eq. (2-18b)] at time $t = 0$ (*solid*) and later in time when $\omega t = \pi$ (*dashed*). The node of the standing wave is shown.

Figure 2-27 depicts the standing wave component of Eq. (2-18g), i.e., the cosβ" term for several values of δ. The *solid* line represents the total E field at $t = 0$. The *dashed* line represents the same E field a half cycle later when $\omega t = \pi$—the algebraic sign of the field is reversed. When $\delta = 0$, the node of the standing wave is displaced 90° in phase (or $\lambda/4$) in front of the reflector. The node below the surface moves closer to it as δ increases. At $\delta = 150°$, the node is just below the surface. It is precisely at the interface when $\delta = 180°$. When $\delta = 225°$, the node of the standing wave is above the surface.

The standing wave ratio is illustrated by Fig. 2-22, which depicts the *total* electric field (at two instances of time) of the wave reflected at an interface. The field increases to a maximum value of E_{max} and then a quarter cycle later shrinks to E_{min}. The standing wave ratio (acronym SWR) is defined

$$V \equiv \frac{\text{maximum of total } E \text{ field}}{\text{minimum of total } E \text{ field}} \equiv \frac{E_{max}}{E_{min}} \tag{2-19a}$$

or

$$V \equiv \frac{1 + |\hat{\rho}|}{1 - |\hat{\rho}|} \; . \tag{2-19b}$$

SWR is identical to the *voltage standing wave ratio* in microwave technology. In the optical portion of the spectrum it could be termed the *E field standing wave ratio* or *amplitude standing wave ratio*. Starting with the identity $|\hat{\rho}| = \sqrt{R}$, Eq. (2-19b) is written

$$V = \frac{1 + \sqrt{R}}{1 - \sqrt{R}} \tag{2-19c}$$

where R is the radiant reflectance. V is never less than one

$$1 \le V \le \mathcal{LN} \tag{2-19d}$$

where $\mathcal{LN}$ stands for *large number*—of the order of 10^6 for optical coatings.

Consider the total E fields at the air-dielectric interface in Fig. 2-22. §2.2.6.3 makes it clear that this is drawn for $n_s = 2.40$. From Eq. (2-8f) $|r| = |\hat{\rho}| = 0.412$. It follows that $V = (1 + 0.412)/(1 - 0.412) = 2.40$, this the ratio of E_{max} to E_{min} shown in that Fig.

Equation (2-19c) is inverted to yield

$$R = (1 - V)^2 / (1 + V)^2 \; . \tag{2-19e}$$

The *loss* $\mathcal{L}$—regardless of whether it is due to absorption A or transmittance T—is

$$\mathcal{L} \equiv 1 - R \equiv A + T = 4V/(1+V)^2 = 4/(2+V+V^{-1}) \ . \qquad (2\text{-}19f)$$

Equation (2-19f) reduces to the transmittance T for a nonabsorbing system. If $V >> 1$ then

$$T \approx 4/V. \qquad (2\text{-}19g)$$

It is useful to express T in terms of the absorbance D (see Eq. (1-6a)]

$$D \approx \log_{10}(V) - 0.6 \ . \qquad (2\text{-}19h)$$

V and $\hat{Y}$ are related when the normalized admittance $\hat{Y}$ is pure real. The reduced admittance $\hat{y}$ (which, like normalized admittance, is dimensionless) is defined as

$$\hat{y}_i \equiv \hat{Y}/n_i \equiv y_r + jy_j \qquad (2\text{-}20a)$$

where n_i is the index of the medium where $\hat{Y}$ is measured. In the case where $\hat{y}$ is pure real, it follows from Eqs. (2-16h) and (2-19b) that

$$V = y_r \qquad \text{or} \qquad V = y_r^{-1} \ , \qquad (2\text{-}20b)$$

depending upon whether y_r is greater than or less than unity.

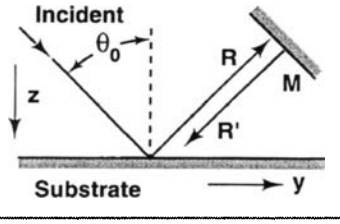


Fig. 2-28 A standing wave in the y direction occurs when a spatially coherent incident beam reflects from substrate at an angle θ_0. The reflected beam R is redirected so that it and the redirected beam (R') interfere.

2.3.3 Lateral standing waves

Although it is rarely encountered, it is possible to produce standing waves in the lateral, i.e., y direction. This occurs in a coherent system, such as a ring laser cavity. Figure 2-28 shows that a well-collimated laser beam is incident upon the substrate at an angle θ_0. It refracts in a layer of index n_1 at an angle of θ_1. The mirror M redirects the reflected beam R back to the surface (an angle θ_0). The latter is shown as R'. The incident wave and redirected wave interfere. From Eqs. (2-3a) and (2-4a) the standing wave component is proportional to

$$E_{total}(y) \propto \exp(-j2\pi y n_1 \sin\theta_1 / \lambda) + \exp(j2\pi y n_1 \sin\theta_1 / \lambda) \qquad (2\text{-}20c)$$

or

$$E_{total}(y) \propto \cos(2\pi y n_0 \sin\theta_0 / \lambda) \qquad (2\text{-}20d)$$

and depends *only* upon the index n_0 and angle θ_0 in the incident medium. If a maximum of E_{total} occurs at a position y_1, the argument of Eq. (2-20d) is $m\pi$ or an odd multiple thereof. The adjacent maximum of $|\hat{E}|$ occurs at $y_1 + d$, where

$$d = \lambda_0 (2 n_0 \sin\theta_0)^{-1} \ . \qquad (2\text{-}20e)$$

d could be termed the "fringe spacing" and is one-half the distance between the grooves of a diffraction grating that diffract in first order the flux at an angle θ_0.

2.3.4 Differential phase shift and reflectance vs angle of incidence

The difference in the phase shift upon reflection δ_s for s polarized flux and δ_p for p polarization is Δ

$$\Delta \equiv \delta_p - \delta_s \ . \qquad (2\text{-}20f)$$

Δ is frequently specified when the polarization state of the flux reflected from mirrors or other multilayer components must be either preserved or altered. Δ is often measured to determine the n and k of a metallic reflector. An alternative sign convention for the phase shift upon reflection is discussed in §2.3.6.

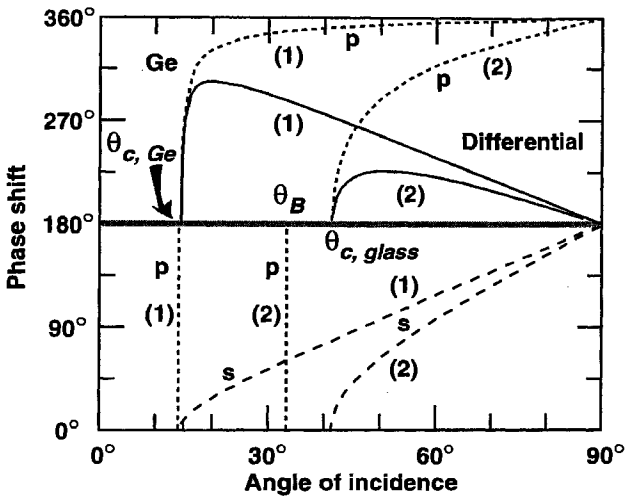


Fig. 2-29 Phase shift upon reflection for *s* polarization (*long dash*), *p* polarization (*short dash*) and differential phase shift (*solid*) vs angle of incidence at an interface between an incident medium of index 4.00 and air (*curve* 1). For *curve* 2, the incident medium index is 1.52. The Brewster angle θ_B and critical angle θ_c are shown. Abelès sign conventions are used.

As an example, Fig. 2-29 plots δ_p , δ_s and Δ vs angle of incidence when the incident medium is a dielectric material (of index n_0) and the emergent medium (index n_s) is air. Using the phase shift convention in this book, *both* r_p and r_s are positive (dense-to-rare media) at normal incidence. As discussed in §2.2.3.3.2, r_p changes sign at θ_B. This appears as a phase jump of 180° in Fig. 2-29. It occurs at $\theta_0 = 14°$ for germanium ($n_0 = 4.00$) and at 33° for glass ($n_0 = 1.52$).

Total internal reflection occurs when θ_0 exceeds the critical angle [see Eq. (2-4e)]. In that case ζ in Eq. (2-12h) is pure imaginary and $R_p = R_s = 1.0$. An example is that the prisms in Fig. 1-116 reflect internally with $R = 100\%$. As the angle increases from θ_c to 90°, δ_s monotonically increases to 180° and δ_p increases to 360°. The differential phase shift Δ attains a maximum when

$$\theta_0 = \arcsin \left[\sqrt{2} (1 + n_0^2 n_s^{-2})^{-1/2} \right] . \qquad (2\text{-}20g)$$

For germanium, Δ is 270° at θ_0 near 20°, which means that a single internal reflection converts linearly polarized flux into circularly polarized flux.

Figure 2-30 depicts R_p, R_s, δ_p, δ_s and Δ for germanium in the visible. It is absorbing in this part of the spectrum; n_s exceeds k_s. R is slightly greater than 50% at normal incidence. R_s grows steadily larger as θ_0 increases from zero to 90°. R_p is nearly zero at a pseudo-Brewster which is close to 80°. The phase shift is almost 180° at normal incidence; δ_p steadily declines until it reaches zero at $\theta_0 = 90°$. Δ is negative at all angles of incidence.

Figure 2-31 shows the same data for a transition metal (such as nickel) in which $n_s \approx k_s$. The minimum in R_p at a relatively high angle of incidence of 70° is characteristic of these metals.

Figure 2-32 depicts the R_s, R_p and phase shifts for a metal with high electrical conductivity (such as silver) in the visible part of the spectrum. The ordinate for R extends from 90% to 100%. A consequence of the fact that $n_s << k_s$ for this material is a reflectance that approaches 95% at normal incidence. Although R_p dips to a minimum at $\theta_0 \approx 70°$ incidence, R_p still exceeds 90% at all angles.

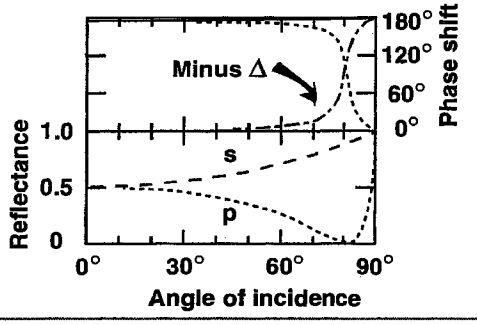


Fig. 2-30 Light reflects from an opaque medium of optical constant $\hat{n} = 5.7 - j1.25$.

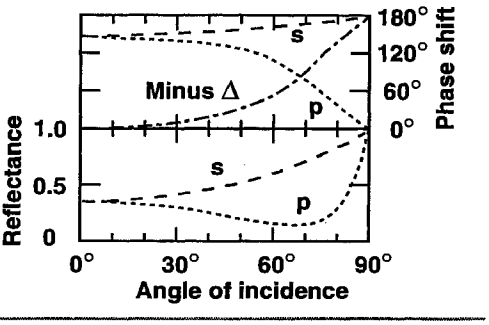


Fig. 2-31 Light reflects from an opaque medium of optical constant $\hat{n} = 1.79 - j1.86$.

For both Figures above, the angle of incidence is in air. Plotted are the reflectance for *s* polarization (*long dash*, left ordinate), *p* polarization (*short dash*, left ordinate), phase shift upon reflection for *p* polarization ((*short dash*, right ordinate), phase shift upon reflection for *s* polarization ((*long dash*, right ordinate) and differential phase shift upon reflection Δ (*long-short dash*). The Abelès sign convention is used.

2.3.5 Constraints on the reflection and transmission coefficients

§2.3.1 distinguishes between the symbols $\hat{r}$ and $\hat{\rho}$, and $\hat{t}$ and $\hat{\tau}$. Either $|\hat{r}|$ or $|\hat{t}|$ may exceed 1.00. A condition of $t > 1.00$ occurs at the interface between dielectric media in which $n_i > n_{i+1}$, as, for example, when a wave emerges from glass into air. The t that exceeds 1.00 does not violate the laws of physics. According to Eq. (2-71i), the irradiance depends upon the product of E and H. Computation of the amplitude transmission coefficient t_H for H would show that $t_H < 1.00$ at this glass-air interface. Conservation of energy is thus guaranteed. $|\hat{r}|$ never exceeds 1.00 at any dielectric interface but often exceeds 1.00 when either the incident medium or emergent medium are absorbing. Even though the amplitude reflection coefficient for the *electric* field may exceed 1.00, the amplitude reflection coefficient for the *magnetic* field is such that the radiant reflectance never exceeds 1.00.

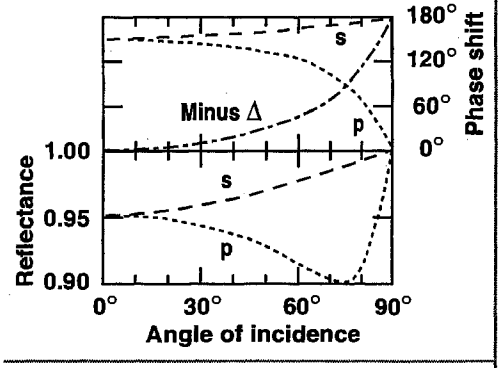


Fig. 2-32 The captions to Figs. 2-30 and 2-31 obtain, with the exception that the optical constant is $\hat{n} = 0.18 - j3.64$.

The comments in the foregoing paragraph about $\hat{r}$ also apply to the *internal* reflection $\hat{\rho}$ (see §2.2.5.4). In fact Table 2-19 lists a $|\hat{\rho}|$ that exceeds 1.00 in an absorbing medium. A consequence of t being greater than 1.00 is that the total E field is higher at the emergent surface of a glass slab, as compared to the incident surface. Boling [73] attributes a lower laser damage threshold at the emergent interface to this effect.

2.3.6 Conventions relating to the phase shift upon reflection

There has been a relentless change of notation over the last century. An example is that a physical optics book of the early twentieth century established the "state of polarization" as the direction of the magnetic vector, rather than the electric vector. Six decades later the roles of the electric and magnetic vectors are reversed. The notation in this book conforms to that of Muller [69], with one exception—the *assumed* direction of reflected E field for p polarization. The convention of Abelès, rather than that of Muller, is used. The convention of Abelès is used in deriving the Fresnel coefficients in §2.2.3.1. The sign convention of Muller is:

> For p polarization, the Fresnel reflection coefficient is positive if the E field component along the z axis is in the *same* direction for the incident and reflected waves.

In either system, when the equation for the Fresnel reflection coefficient produces a positive sign, the *assumed* direction and *actual* direction are the same. Regardless of which convention is used, the following steps should be followed:

(1) *Assume* a coordinate system that specifies the directions of the E and H fields both before and after a reflection at an interface occurs.
(2) Derive the equations for reflection at an interface, similar to the procedure in §2.2.3.
(3) If, for a given configuration, the aforementioned equation produces a negative sign for the reflection coefficient for E, then the direction of the E field of the reflected wave should be inverted. The same applies to H in the reflected wave.

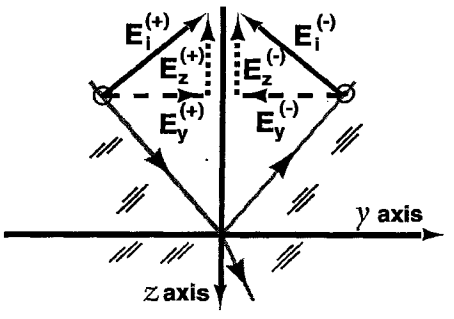


Fig. 2-33 Actual directions of the incident and reflected components of the electric field E for p polarized light at an interface (coincident with the y axis) between air and glass of index 1.517 at 30° angle of incidence. The components of E along the y axis (*long dash*) and z axis (*short dash*) are shown.

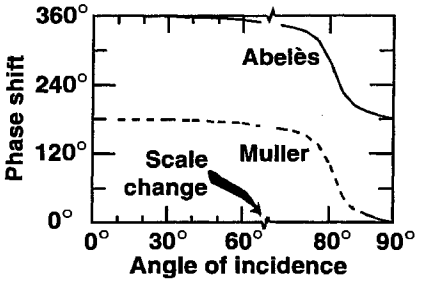


Fig. 2-34 Differential phase shift upon reflection [see Eq. (2-20f)] — modulo 360° — at an interface between air and a medium of optical constant $\hat{n} = n - jk = 5.7 - j1.25$. The *solid* curve uses the Abelès sign convention and the *dashed*, Muller [69].
The abscissa changes scale at 70°.

Figure 2-33 shows the *actual* directions of the E fields for a plane wave that reflects at an interface between *air* (incident medium) and *glass* (emergent medium). The incidence angle is less than Brewster's angle. The Abelès sign convention is embodied in

Eq. (2-9c) — it produces a negative Fresnel coefficient, which means that the direction of the assumed E field is reversed in the reflected wave from the direction shown in the left panel of Fig. 2-12. For the TM polarized wave, the Fresnel coefficient using the Muller notation is

$$r_p = \frac{n_{i+1}\sec\theta_{i+1} - n_i \sec\theta_i}{n_{i+1}\sec\theta_{i+1} + n_i \sec\theta_i} \tag{2-20h}$$

and is positive (see Born & Wolf [75] section 1.5.2).

The sign convention is illustrated in Fig. 2-33, which depicts the y and z components of the E fields. For the reflected wave, the y component is along the negative y axis, which, according to §2.2.3.1, implies a negative sign in the Fresnel equation — for the Abelès notation. In the Muller notation, however, the direction of the z component of the E field in the incident and reflected wave is the same. Ergo, the sign of Eq. (2-20h) must be positive, which it is. To put it another way, the *assumed E* field direction and *actual E* field direction in Fig. 2-33 are the same.

There is a simple expedient that converts from the Abelès to the Muller notation: add 180° to the phase shift upon reflection for p. This tacitly assumes that the phases are computed modulo 360°. Figure 2-34 shows, for example, the differential phase shift Δ at an interface between air and absorbing germanium (see Fig. 2-30). The upper curve (*solid*) is based upon Abelès and the lower (*dashed*) follows the Muller sign convention.

A final comment is made about the two conventions. The advantage of the Abelès system is that one set of equations may be used, and, with certain restrictions, *effective indices* may be used to adapt equations that pertain at normal incidence to nonnormal incidence. The advantage of the Muller system is the reversion of the coordinate system for the reflected wave. A clockwise rotation of the direction of the E field in the incident beam is also clockwise in the reflected beam.

2.4 Properties of a multilayer

2.4.1 Introduction

§2.4 develops the explicit equations for the reflectance R and transmittance T of both a single film and a multilayer. Although this book focuses upon homogeneous layers, these methods can be extended to optically inhomogeneous films.

2.4.2 Radiant reflectance and transmittance of a single layer

2.4.2.1 Balance of flow at an interface

§2.1.2 establishes that there are two components of the electric field when a wave is propagating in a homogeneous medium.

As shown in Fig. 2-35, these are $\hat{E}_i^{(+)}$ propagating in the +z direction and $\hat{E}_i^{(-)}$ in the –z direction in the ith medium. These E fields represent the components tangential to the surfaces at nonnormal incidence. At the interface between the ith and $(i+1)^{th}$ surfaces, a balance of flow requires that

$$\hat{E}_i^{(-)} = r_i \hat{E}_i^{(+)} + t_{i+1,i}\hat{E}_{i+1}^{(-)} \tag{2-21a}$$

$$\hat{E}_{i+1}^{(+)} = t_{i,i+1}\hat{E}_i^{(+)} + r_i' \hat{E}_{i+1}^{(-)} \tag{2-21b}$$

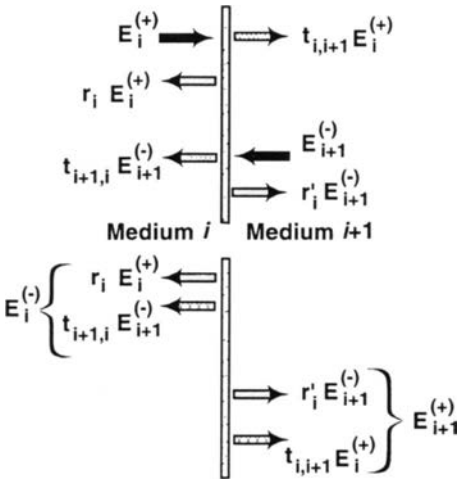


Fig. 2-35 (*Upper*) $\hat{E}_i^{(+)}$ and $\hat{E}_{i+1}^{(+)}$ are the components of the wave traveling from left-to-right in medium i and medium i+1, respectively. The "(–)" superscripts denote the corresponding components flowing from right-to-left. The *solid black* vectors denote an input to the interface. The *cross hatched* vectors are transmitted through the interface and *dotted* vectors are reflected from the interface. (*Lower*) Balance of flow at the interface, as embodied in Eqs. (2-21a) and (2-21b).

where r_i, r'_i, $t_{i,i+1}$, $t_{i+1,i}$ are the Fresnel coefficients of reflection and of transmission for the incident and emergent sides of the interface, respectively. In other words, the r and t in Eq. (2-8f) — or their counterparts at nonnormal incidence — are used.

The foregoing two equations are recast using matrix theory, the rudiments of which are summarized in §2.6.2. Those equations are expressed as the product of a matrix and column vector

$$\begin{bmatrix} \hat{E}_i^{(+)} \\ \hat{E}_i^{(-)} \end{bmatrix} = \frac{1}{t_{i,i+1}} \begin{bmatrix} 1 & -r'_i \\ r_i & t_{i,i+1}\, t_{i+1,i} - r_i\, r'_i \end{bmatrix} \begin{bmatrix} \hat{E}_{i+1}^{(+)} \\ \hat{E}_{i+1}^{(-)} \end{bmatrix}. \tag{2-21c}$$

Applying Eqs. (2-6g) and (2-6h), this reduces to

$$\begin{bmatrix} \hat{E}_i^{(+)} \\ \hat{E}_i^{(-)} \end{bmatrix} = \frac{1}{t_{i,i+1}} \begin{bmatrix} 1 & r_i \\ r_i & 1 \end{bmatrix} \begin{bmatrix} \hat{E}_{i+1}^{(+)} \\ \hat{E}_{i+1}^{(-)} \end{bmatrix}. \tag{2-21d}$$

The foregoing equation establishes that the electric field components $\hat{E}_i^{(+)}$ and $\hat{E}_i^{(-)}$ at the left of an interface are linear combinations of $\hat{E}_{i+1}^{(+)}$ and $\hat{E}_{i+1}^{(-)}$ on its emergent side.

2.4.2.2 Propagation and phase of retardation

Given that $E^{(+)}$ and $E^{(-)}$ at the interface are known, what are these components in the interior of the ith layer—at a distance h_i below the surface? Equation (2-4c) describes

the propagation of a plane wave. The origin of the coordinate system is at the top surface of the layer. The y coordinate is frozen and the fields are examined at a given instance of time. The y and t in Eq. (2-4c) are constants. For the wave propagating in the +z direction

$$\hat{E}^{(+)}(h_i) = \hat{E}_{0,i}^{(+)} \exp(-j\hat{\beta}_i) \ . \tag{2-21e}$$

$\hat{E}_{0,i}^{(+)}$ is the amplitude (i.e., electric field) at the origin. The "$\hat{}$" is subsequently omitted from the β. It is tacitly assumed that β is complex [see Eq. (2-5b)] when the medium is absorbing. Inversion of Eq. (2-21e) provides

$$\hat{E}_{0,i}^{(+)} = \hat{E}^{(+)}(h_i) \exp(j\beta_i) \tag{2-21f}$$

for the incident component and

$$\hat{E}_{0,i}^{(-)} = \hat{E}^{(-)}(h_i) \exp(-j\beta_i) \tag{2-21g}$$

for the reflected component. The two previous equations are written in matrix form

$$\begin{bmatrix} \hat{E}_{0,i}^{(+)} \\ \hat{E}_{0,i}^{(-)} \end{bmatrix} = \begin{bmatrix} \exp(j\beta_i) & 0 \\ 0 & \exp(-j\beta_i) \end{bmatrix} \begin{bmatrix} \hat{E}^{(+)}(h_i) \\ \hat{E}^{(-)}(h_i) \end{bmatrix} = \mathbf{P}_i \begin{bmatrix} \hat{E}^{(+)}(h_i) \\ \hat{E}^{(-)}(h_i) \end{bmatrix} . \tag{2-21h}$$

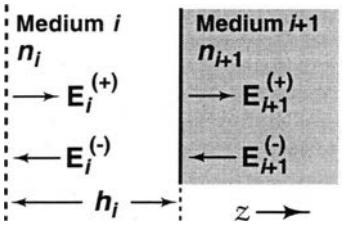


Fig. 2-36 Equation (2-22a) transforms the electric field components $\hat{E}_{i+1}^{(+)}$ and $\hat{E}_{i+1}^{(-)}$ in the $(i+1)^{th}$ medium to $\hat{E}_i^{(+)}$ *and* $\hat{E}_i^{(-)}$ at a distance h_i in the negative z direction in the ith medium.

The foregoing transforms the components $\hat{E}_i^{(+)}$ and $\hat{E}_i^{(-)}$ by distance h_i in the $-z$ direction. This also defines the matrix $\mathbf{P}_i$.

2.4.2.3 E^+ E^- matrix

The matrix is found that transfers field components on the right of the ith interface to a distance h_i on its left, as shown in Fig. 2-36. The left column vector in Eq. (2-21d) is identical to the right column vector of Eq. (2-21h). The subscript "$_{0,i}$" is contracted to "$_i$"

and the resultant matrix is

$$\begin{bmatrix} \hat{E}_i^{(+)} \\ \hat{E}_i^{(-)} \end{bmatrix} = \frac{1}{t_{i,i+1}} \begin{bmatrix} \exp(j\beta_i) & r_i \exp(j\beta_i) \\ r_i \exp(-j\beta_i) & \exp(-j\beta_i) \end{bmatrix} \begin{bmatrix} \hat{E}_{i+1}^{(+)} \\ \hat{E}_{i+1}^{(-)} \end{bmatrix} . \tag{2-22a}$$

The foregoing matrix is used in this book. It contains the reflection coefficient of an interface and the phase thickness of the medium to the right of that interface.

2.4.2.4 Matrix elements of a single layer and substrate

The E^+E^- matrix for a single layer is developed. Its amplitude reflection coefficient $\hat{\rho}$ and its amplitude transmission τ coefficient are found. §2.4.3.3 extends this to a multilayer.

There is yet another way to write this matrix. This alternate form transfers the fields $\hat{E}_{i-1}^{(-)}$ and $\hat{E}_{i-1}^{(-)}$ at the left of the interface to a distance h_i on its right, as shown in Fig. 2-37. The matrix is

$$\begin{bmatrix} \hat{E}_{i-1}^{(+)} \\ \hat{E}_{i-1}^{(-)} \end{bmatrix} = \frac{1}{t_{i-1,i}} \begin{bmatrix} \exp(j\beta_i) & r_{i-1}\exp(-j\beta_i) \\ r_{i-1}\exp(j\beta_i) & \exp(-j\beta_i) \end{bmatrix} \begin{bmatrix} \hat{E}_i^{(+)} \\ \hat{E}_i^{(-)} \end{bmatrix}. \quad (2\text{-}22b)$$

In treating the reflection at an interface in §2.2.3.2, it is assumed that there is no reflected component of the electric field in the emergent medium. The same constraint is imposed here. The partial waves are shown in Fig. 2-38. $\hat{E}_0^{(+)}$ and $\hat{E}_0^{(-)}$ are the components in the incident medium. Within the layer itself there are the components $\hat{E}_1^{(+)}$ and $\hat{E}_1^{(-)}$. In the substrate there is only the component $\hat{E}_2^{(+)}$ that represents a wave in the +z direction. Applying Eq. (2-21d) to the substrate film interface yields

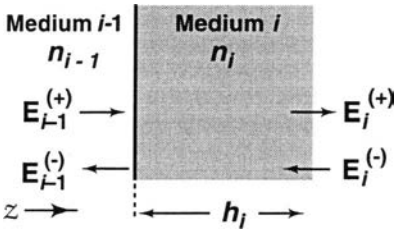


Fig. 2-37 Equation (2-22b) transforms the electric field components $\hat{E}_i^{(+)}$ and $\hat{E}_i^{(-)}$ in the ith medium to $\hat{E}_{i-1}^{(+)}$ and $\hat{E}_{i-1}^{(-)}$ at a distance h_i in the negative z direction in the ith medium.

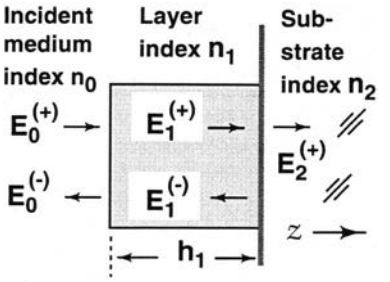


Fig. 2-38 Incident and reflected components of the electric field are $\hat{E}_0^{(+)}$ and $\hat{E}_0^{(-)}$ in the incident medium, $E_1^{(+)}$ and $\hat{E}_1^{(-)}$ on the emergent side of the layer and $\hat{E}_2^{(+)}$ in the emergent medium (i.e., substrate). The layer-substrate interface is shaded.

$$\begin{bmatrix} \hat{E}_1^{(+)} \\ \hat{E}_1^{(-)} \end{bmatrix} = \frac{1}{t_{1,2}} \begin{bmatrix} 1 & r_1 \\ r_1 & 1 \end{bmatrix} \begin{bmatrix} \hat{E}_2^{(+)} \\ 0 \end{bmatrix}. \quad (2\text{-}22c)$$

Multiplication of this matrix by the column vector results in

$$\hat{E}_1^{(+)} = \hat{E}_2^{(+)} / t_{1,2}$$
$$\hat{E}_1^{(-)} = \hat{E}_2^{(+)} r_1 / t_{1,2} \quad (2\text{-}22d)$$

which are substituted in the column vector on the right of Eq. (2-22b)

$$\begin{bmatrix} \hat{E}_0^{(+)} \\ \hat{E}_0^{(-)} \end{bmatrix} = \frac{1}{t_{0,1}} \frac{1}{t_{1,2}} \begin{bmatrix} \exp(j\beta_1) & r_0\exp(-j\beta_1) \\ r_0\exp(j\beta_1) & \exp(-j\beta_1) \end{bmatrix} \begin{bmatrix} \hat{E}_2^{(+)} \\ r_1\hat{E}_2^{(+)} \end{bmatrix}. \quad (2\text{-}22e)$$

2.4.2.4.1 Amplitude transmission coefficient

Given $\hat{E}_0^{(+)}$ and $\hat{E}_0^{(-)}$ in the foregoing, the reflection coefficient for the electric field $\hat{\rho}_0$ (amplitude reflection coefficient) is found from Eq. (2-18b). As discussed below, the reference position for the measurement of the phase is at the top surface of the outermost layer of the stack.

The transmission coefficient for the electric field (the amplitude transmission coefficient) $\hat{\tau}$ is

$$\hat{\tau} \equiv \hat{E}_s^{(+)} / \hat{E}_0^{(+)} = |\tau| \exp(-j\nu) \tag{2-22f}$$

where $\hat{E}_s^{(+)}$ is the traveling wave component of the electric field in the substrate — the emergent medium — which is *medium* 2 in this example. The negative sign in the exponent of Eq. (2-22f) is chosen so that the magnitude of ν (the transmissive phase retardance) increases as a multilayer grows thicker.

2.4.2.5 Explicit equation for the amplitude reflection coefficient ρ of a single layer

The equations for R and T of a single layer are summarized. The particular form that might be chosen depends upon the available computational facilities. An example is that small calculators often have polar-to-Cartesian conversion routines that facilitate the calculation of the complex quantities in Eq. (2-22e).

From Eqs. (2-22e), (2-18b) and (2-6h), the amplitude reflection coefficient $\hat{\rho}_0$ of a single layer is

$$\hat{\rho}_0 = \frac{\hat{r}_0 + \hat{r}_1 \exp(-j2\beta_1)}{1 + \hat{r}_0 \hat{r}_1 \exp(-j2\beta_1)} . \tag{2-23a}$$

In the case where the layer is nonabsorbing and β_1 is pure real, $\hat{\rho}_0$ can be plotted on the complex plane parametrically vs the phase thickness β_1. It is easily proven that this is a circle. The general form of Eq. (2-23a) is

$$\hat{w} = \frac{a + b\hat{z}}{c + d\hat{z}} \tag{2-23b}$$

where $\hat{w}$ and $\hat{z}$ are complex variables. a, b, c and d are complex constants. The foregoing is the bilinear transformation that maps circles into circles. Because a plot on the Argand plane of the equation $\hat{z} = \exp(-j2\beta)$ represents a circle, $\hat{\rho}_0$ must also be a circle. For this reason, a parametric plot of $\hat{\rho}(\beta)$ is termed a *circle diagram* [Berning 62, Apfel 72].

2.4.2.6 *R* and *T* of nonabsorbing layer, normal incidence

The preferred formulation of the equations depends upon whether the film and/or substrate are nondispersive or dispersive. In either case, it is computationally more efficient to compute either T or R/T. For the sake of completeness, the reflectance R of a single nonabsorbing film is

$$R = 1 - T = \frac{a_1 \cos^2\beta_1 + a_3 \sin^2\beta_1}{a_2 \cos^2\beta_1 + a_4 \sin^2\beta_1} \tag{2-23d}$$

where

$$a_1 \equiv (n_0 - n_s)^2 \qquad a_2 \equiv (n_0 + n_s)^2 \tag{2-23e}$$

$$a_3 \equiv (n_1 - n_0 n_s / n_1)^2 \qquad a_4 \equiv (n_1 + n_0 n_s / n_1)^2 \tag{2-23f}$$

As shown in Figs. 2-37 and 2-39, the extrema in reflectance occur when β_1 equals 90°, 180°, 270°, and so on. When β_1 is 180° or an integral multiple thereof, only the coefficients in Eq. (2-23e) determine R, which is the Fresnel reflection at the surface of the uncoated substrate [see Eq. (2-12f)]. As discussed in §2.6.3.1, such a layer is termed *absentee* when β_1 is 180°. This is termed the *multiple halfwave point* because the optical

thickness is a halfwave — or an integer multiple thereof. Mathematically, if λ_m and σ_m are wavelengths or wave numbers at which such a halfwave occurs,

$$2 n_1 h_1 / \lambda_m = 2 n_1 h_1 \sigma_m = m \qquad (2\text{-}23g)$$

where m is an integer.

When the phase thickness is 90° or an odd multiple thereof, the optical thickness is a *quarterwave point.* Mathematically, if λ_m and σ_m are wavelengths or wave numbers at which such a quarterwave occurs,

$$4 n_1 h_1 / \lambda_m = 4 n_1 h_1 \sigma_m = m \qquad (2\text{-}23h)$$

where m is an *odd* integer. The condition for minimum reflectance at $\beta_1 = 90°$ is

$$n_0 < n_1 < n_s \qquad \text{or} \qquad n_0 > n_1 > n_s \ , \qquad (2\text{-}23i)$$

which is termed the *intermediate* index condition. It follows that for zero reflectance, a_3 in Eq. (2-23f) is zero. This requires that

$$n_1 = \sqrt{n_0 n_s} \ . \qquad (2\text{-}23j)$$

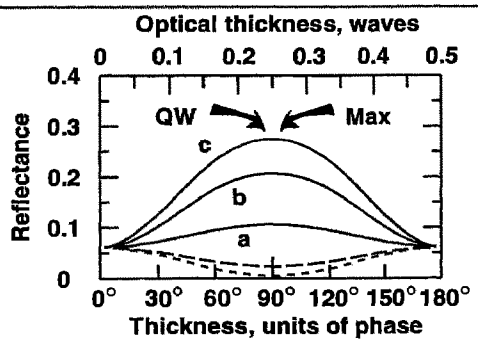


Fig. 2-39 Radiant reflectance vs phase thickness of nonabsorbing layers deposited upon a substrate of refractive index 1.65. The refractive indices of the layers are: 1.38 (*short dash*); 1.50 (*long dash*); (*a*) 1.80; (*b*) 2.10; (*c*) 2.30. The incident medium is air. At QW, the optical thickness of the layer is a quarterwave.

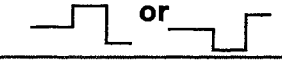
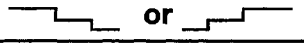
R is a maximum at $\beta_1 = 90°$ if it is not a minimum. An example is shown in Fig. 2-39, which depicts reflectance vs phase thickness for relatively low index ($n_1 = 1.38,\ 1.50$) films on a glass ($n_s = 1.65$) substrate. Both of the *dashed* curves have a minimum reflectance at $\beta_1 = 90°$, since the indices satisfy Eq. (2-23i). Curves *a, b* and *c* show the $R(\beta)$ as the film index gets progressively larger; R_{max} increases accordingly. Hence, the low index films are used as antireflection coatings and the higher index layers are incorporated into beamdividers and partial reflectors.

Regardless of whether a maximum or minimum occurs, when Eq. (2-23h) is satisfied, the extremum reflectance R_{ex} is

$$R_{ex} = \left[\frac{1 - \dfrac{n_1^2}{n_0 n_s}}{1 + \dfrac{n_1^2}{n_0 n_s}} \right]^2 \quad at\ \beta = 90° \ . \qquad (2\text{-}24a)$$

Fig. 2-40 isometrically plots reflectance vs phase thickness and layer index. A germanium layer (index = 4.2) on glass manifests a maximum reflectance of 70% when its optical thickness is a quarterwave. Table 2-4 summarizes the positions of the extrema for the *intermediate* index profile (lower row) and non-intermediate index profile (*upper* row). Halfwave and quarterwave refer to the optical thickness at the designated extremum.

Table 2-4 Reflectance extrema of a single layer

Index profile	Maxima	Minima
	Quarterwave	Halfwave
	Halfwave	Quarterwave

Index profiles (incident medium, layer and substrate) of nonabsorbing layers and the optical thickness of the layer at the extrema in reflectance.

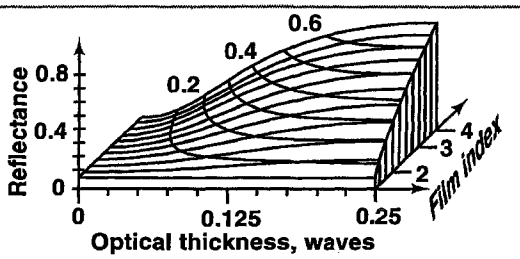


Fig. 2-40 Isometric plot of reflectance vs optical thickness of a single nonabsorbing layer of refractive index n deposited on a substrate of index 1.50, where n steps from 1.50 to 4.0 in increments of 0.25. The step of the isoreflectance contour is 0.1.

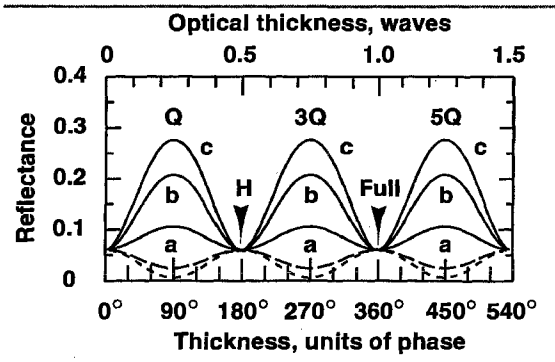


Fig. 2-41 R vs phase thickness and optical thickness of nonabsorbing layers deposited upon a substrate of refractive index 1.65. The layer refractive indices are: 1.38 (*short dash*); 1.50 (*long dash*); (*a*) 1.80; (*b*) 2.10; 2.30 (*long-short dash*). The incident medium is air. Q, 3Q, 5Q, H and Full are explained in the text.

The R/T formulation is computationally more efficient. Defining F as

$$F \equiv R/T \qquad (2\text{-}24b)$$

results in the identities

$$R = F(1+F)^{-1} \qquad (2\text{-}24c)$$

and

$$T = (1+F)^{-1} \; . \qquad (2\text{-}24d)$$

For a single layer,

$$F = (a_1 \cos^2\beta_1 + a_3 \sin^2\beta_1)(4\, n_0\, n_s)^{-1} \qquad (2\text{-}24e)$$

where the a coefficients appear in Eqs. (2-23e) and (2-23f).

The advantage of using β as a variable in plots such as Fig. 2-39 or 2-41 is that the maxima and minima occur at multiples of 90°. Due to the fact that the wave number σ is directly proportional to β in Eq. (2-5b), σ could be used equally well. When wavelength is used in lieu of σ, the symmetry so evident in Fig. 2-39 vanishes. For this reason, σ is often preferred to λ in multilayer design.

2.4.2.6.1 Thicker layers

The phase thickness β on the abscissa of Fig. 2-39 may be extended beyond 180°. Figure 2-41 shows that each curve is a repeat of its behavior when β is in the range $0° \le \beta \le 180°$. For the higher index films, R is a maximum when the optical thickness is a quarterwave Q (β = 90°), three quarterwave 3Q (β = 270°) or five quarterwave 5Q (β = 540°). Nested between Q and 3Q is H—the halfwave position—where the reflectance is independent

of the layer refractive index. Between the peaks at 3Q and 5Q the optical thickness is a full wave ($\beta = 360°$) and the reflectance plunges to that of the uncoated substrate. As long as the layer is optically homogeneous and nonabsorbing, the abscissa curve repeats every 180° as the phase thickness grows thicker and thicker.

2.4.2.6.2 Effects of dispersion

The symmetry of the $R(\beta)$ curve is altered if either the incident medium, film or substrate is dispersive. Figures 2-42 and 2-43 show an example in which $R(\sigma)$ and $R(\lambda)$ are plotted for a high index film on a glass substrate. The glass index is nondispersive. Because the film index exceeds both the index of the substrate and incident medium, a maximum reflectance occurs when its optical thicknesses is approximately a quarterwave.

The splitting between the reflectance curves in the aforementioned Figs. occurs because in one instance the film index is dispersive. The caption to Fig. 2-42 has information on the dispersion of the layer. The nondispersive layer has an index of 2.27. The metric thicknesses are adjusted so that each has an optical thickness of $\lambda_0/4$ at $\lambda_0 = 0.45\,\mu m$. The indices of the dispersive and nondispersive layers are the same at this wavelength. Dispersion has the following effects:

- The wavelengths at which the halfwave points occur (reflectance minima in the case $n_1 > n_s$) are not affected by dispersion of the film itself, although they affected by the dispersion of the substrate index.
- The wavelengths at which the quarterwave points occur are influenced by dispersion of the film; the reflectance maxima are shifted to higher frequencies (i.e., shorter wavelengths).

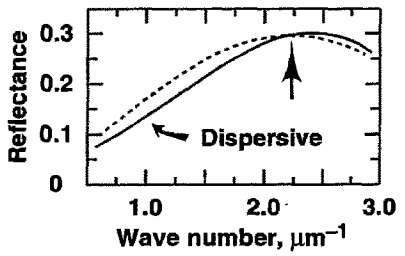


Fig. 2-42 Reflectance vs wave number σ of the coating described below.

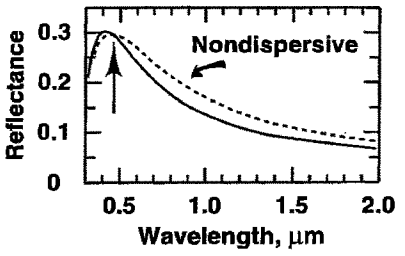


Fig. 2-43 Reflectance vs wavelength λ of the coating described below.

Layers with an optical thickness of $\lambda_0/4$ at $\lambda_0 = 450$ nm deposited upon a nondispersive substrate of index 1.52. The refractive index of the nondispersive film (*dashed* curves) is 2.27. The index of the dispersive layer (*solid* curves) is given by Eq. (2-2c) with $A = (2.10)^2$ and $B = 0.15\ (\mu m)^2$. The incident medium is air.
The arrow in Fig. 2-42 is the wave number equivalent of 450 nm.

In particular, Abelès [63] derives a transcendental equation for the β values at which $R(\sigma)$ of a dispersive film is a maximum. An alternative approach is to find the maximum by calculating $R(\sigma)$ near the maximum. A second-order interpolation polynomial (i.e., a parabola) is fitted through three data points (i.e., reflectance vs wave number σ)

$$R(\sigma) = C + D\sigma + E\sigma^2 \ . \qquad (2\text{-}24f)$$

Differentiation gives the wave number σ_{ex} of the extremum

$$\sigma_{ex} = -D/(2E) \ . \qquad (2\text{-}24g)$$

The shift of R_{max} to higher β values may be visualized if the amplitude reflection coefficient $\hat{\rho}$ of the dispersive layer is plotted parametrically vs β in the Argand plane, as in Fig. 2-44. The $\hat{\rho}(\beta)$ of a nondispersive film is shown for comparison. This is a circle, for reasons described in §2.4.2.5. When the film index is dispersive, the distance from the center of the circle to the locus is constantly increasing as β gets larger, hence $|\hat{\rho}|$ has not yet achieved a maximum on the negative real axis. The maximum of $|\hat{\rho}|$ occurs at the arrowed position, where $\beta = 99.3°$. For comparison, the maximum reflectance is 31.2% for the nondispersive layer and 31.6% for the dispersive film. Valeev [63, 64] discussed the effects of dispersion and absorption on the $R(\sigma)$ curve of a single layer.

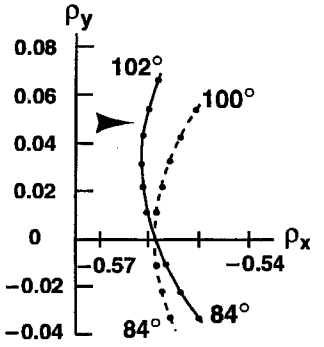


Fig. 2-44 Amplitude reflection coefficient plotted parametrically in Cartesian coordinates vs the phase thickness β (2° increments) for the layers captioned in Figs. 2-42 and 2-43— nondispersive (*dashed*) and dispersive (*solid*). For the dispersive layer, $|\hat{\rho}|$ is a maximum at the arrowed position.

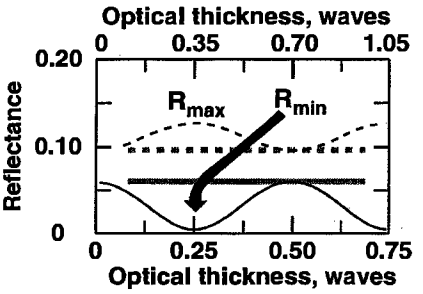


Fig. 2-45 Reflectance (for *p* polarization at $\theta_0 \neq 0$) vs optical thickness of a layer of refractive index 1.38 deposited on a substrate of index 1.65. At normal incidence (*solid*) the lower abscissa obtains and at 75° incidence (*dashed*), the upper. At 75° and at normal incidence, the shaded horizontal lines represent the reflectance of the uncoated substrate.

2.4.2.7 *R* and *T* of nonabsorbing layer, nonnormal incidence

With the proviso that the incident angle does not exceed the critical angle in either the substrate or film, any of the equations in §2.4.2.6 can be used at nonnormal incidence, provided the $\cos\theta_1$ in Eq. (2-4d) is included when β is computed. Effective indices [in Eqs. (2-8b) or (2-10b)] should be substituted for the appropriate polarization. *Caveat:* An effective index should *not* be substituted for the n in Eq. (2-4d). Figure 2-45 provides an illustration of effective indices. A magnesium fluoride film (index n_1) is deposited upon a glass substrate (index n_s), with air (index n_0) as the incident medium. Figure 2-45 graphs its reflectance vs optical thickness at both normal incidence and at 75° for the TM wave. At 75°, the extremum reflectance is a maximum, whereas at normal incidence a minimum occurs. This behavior is understood with the help of Table 2-5, which shows the indices of the three media and the corresponding effective indices at 75°. At normal incidence, the extremum of R when $\beta = 90°$ is a minimum, because Eq. (2-23i) is satisfied. At θ_0 of 75°, however, Table 2-5 shows that $n_0 > n_1$, where effective indices are considered.

This is the reason a maximum in R occurs when $\beta = 90°$, as is evident in Fig. 2-45. In fact, such a maximum is observed at any angle incidence in excess of Brewster's angle at the interface between air and the film.

2.4.2.8 Frustrated total reflection in a single layer

Frustrated Total Reflection (its acronym is FTR) occurs when the angle of incidence θ_0 exceeds the critical angle [see Eq. (2-4e)] in a nonabsorbing layer. The angle of incidence must not exceed the critical angle in the substrate, however.

Figure 2-46 shows the experimental setup for a FTR layer in which a relatively high index prism is the incident medium. A film of lower index is deposited upon its hypotenuse. If no prism were attached to the obverse side of the coating, then θ_0 would exceed the critical angle in the emergent medium (air) and TIR *Total Internal Reflection* occurs. The reflectance is 100% for both planes of polarization, as long as the film is nonabsorbing — regardless of its thickness.

FTR occurs when there is a prism on the emergent side of the layer, which is often encased in optical cement (see §1.9.2 and §6.6.1). The configuration is shown in Fig. 2-46. That layer should be sufficiently thin to allow the flux to penetrate into the emergent prism. In other words, the total reflection is *frustrated*. This phenomenon is equivalent to quantum mechanical tunneling through a potential barrier [Bohm 51]. Such a film could be called a *tunnel layer* because the exponentially attenuated E field "tunnels" through it. Equations for its T and R are developed below.

Table 2-5 Effective indices at 75° incidence

Index at normal incidence	Effective index for TM	θ_i
1.00	3.86	75°
1.38	1.93	44.4°
1.65	2.04	35.8°

Effective indices [see Eq. (2-38b)] for p polarized flux. The incidence angle of 75° in air and θ_i is the angle of refraction in the medium whose index is listed in *column* 1.

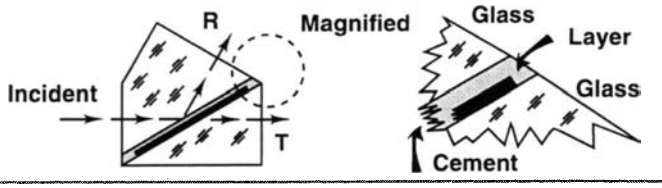


Fig. 2-46 (*Left*) Frustrated total reflection FTR occurs when a single layer of relatively low index is deposited on a higher index prism. (*Right*) The film is embedded in optical cement. The coating thickness is exaggerated for purposes of illustration.

As long as the media are nonabsorbing, $T = 1 - R$ and the R/T equation is valid. Within the FTR layer $\hat{\gamma}_1$ is pure imaginary and from Eq. (2-3h)

$$\gamma_1' \equiv \sqrt{L^2 n_1^{-2} - 1} \ . \tag{2-25a}$$

L appears in Eq. (2-4a). The parameter β_1 [see Eq. (2-4h)] is no longer the "the phase of retardation," since the amplitude of the wave is *attenuated*, but not *retarded* in phase, as

it penetrates the FTR layer. Because the media are nonabsorbing, the R/T equations are valid [Lees 79b]. For s polarization

$$F_s = R_s / T_s = (1 - A_s)^2 \cosh^2\beta_1' + (B_s + C_s)^2 \sinh^2\beta_1' \qquad (2\text{-}25b)$$

and for p polarization

$$F_p = R_p / T_p = (1 - A_p)^2 \cosh^2\beta_1' + (B_p + C_p)^2 \sinh^2\beta_1' \qquad (2\text{-}25c)$$

Table 2-6 Constants relating to a single FTR layer

Constant	p polarization	Constant	s polarization
A_p	$\dfrac{n_s \sec\theta_s}{n_0 \sec\theta_0}$	A_s	$\dfrac{n_s \cos\theta_s}{n_0 \cos\theta_0}$
B_p	$\dfrac{\gamma_1' n_s \sec\theta_s}{n_1}$	B_s	$\dfrac{n_s \cos\theta_s}{n_1 \gamma_1'}$
C_p	$\dfrac{n_1}{\gamma_1' n_0 \sec\theta_0}$	C_s	$\dfrac{n_1 \gamma_1'}{n_0 \cos\theta_0}$

Constants appearing in Eqs. (2-25b) and (2-25c), where n_0, n_1 and n_s are the indices of the incident medium, layer and substrate, respectively. θ_0 and θ_s are the angles of incidence in the incident medium and substrate, respectively. γ_1' appears in Eq. (2-25a).

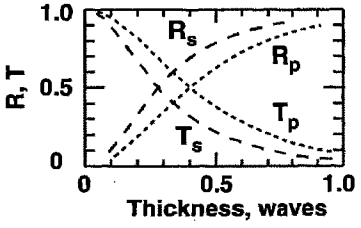


Fig. 2-47 At an incidence angle of 53°, R_p and T_p (*short dash*), as well as R_s and T_s (*long dash*), vs metric thickness (in waves) of a single nonabsorbing layer of index 1.35. The refractive index of the incident and emergent media is 1.72.

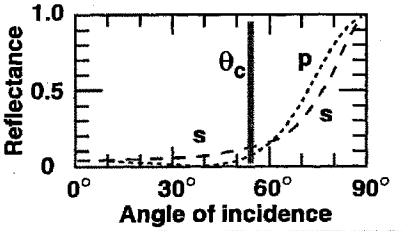


Fig. 2-48 Versus angle of incidence, R_p (*short dash*) and R_s (*long dash*) of a layer of index 1.35 and 0.2 waves thick. The refractive index of the incident and emergent medium is 1.72. θ_c is the critical angle [see Eq. (2-4e)].

where A, B, C appear in Table 2-6. When $n_0 = n_s$, $A = 1$ and the reflectance is zero as $h \to 0$. When $h/\lambda >> 1$, $\sinh\beta_1' \approx \cosh\beta_1' \approx \frac{1}{2}\exp(\beta_1')$ and $T \propto \exp(-2\beta')$.

One way of plotting R and T of the FTR layer is vs thickness, as shown in Fig. 2-47. The rate at which T decreases depends on γ', which in turn is determined by θ_0 and n_1/n_0. Another approach is to plot reflectance vs angle for a layer of given thickness, as depicted in Fig. 2-48. R_p decreases to zero at the internal Brewster angle (at 35°) and then monotonically increases. There are no discontinuities in the curves at the critical angle. R_s and R_p are equal at an angle

$$\theta_0 = \arcsin\left[\sqrt{2\,(n_0^2 n_1^{-2} + 1)^{-1}}\,\right] . \qquad (2\text{-}25d)$$

Astheimer [66] constructed a nonpolarizing FTR modulator at this angle. Turner [47] used a single cryolite layer immersed in a glass of relatively high refractive index as a beamdivider. Other applications include its use as a laser output coupler [Bergstein 62, Steele 66] and also as a long wavelength pass filter [Wettling 64]. Baumeister [67] reviews applications of FTR. Lees [79a] constructed a linear polarizer utilizing FTR.

Table 2-7 Coefficients relating to a single absorbing layer

$\begin{matrix} b \\ c \end{matrix} = (n_1 \pm n_s)^2 + k_1^2$	$\begin{matrix} a \\ d \end{matrix} = (n_1 \mp n_0)^2 + k_1^2$
$\begin{matrix} e \\ g \end{matrix} = 2\,[(n_0^2 + n_s^2)(n_1^2 + k_1^2) - (n_1^2 + k_1^2)^2 - n_0^2 n_s^2 \mp 4 n_0 n_s k_1^2]$	
$\begin{matrix} f \\ h' \end{matrix} = 4 (n_s \mp n_0)(n_1^2 + k_1^2 \pm n_0 n_s) \lvert k_1 \rvert$	

Constants appearing in Eqs. (2-25e) and (2-25f), where n_0, n_1 and n_s are the indices of the incident medium, layer and substrate, respectively. k_1 is the extinction coefficient of the layer. After Abelès [63].

2.4.2.9 Reflectance and transmittance of an absorbing layer

Any of several computational methods may be used. In a small calculator where storage is at a premium, for example, Eq. (2-23a) might be used and the radiant reflectance found from Eq. (2-12c). For a film of index n_1, extinction coefficient k_1 and metric thickness h_1, R and T are

$$R = \frac{eC + fS + abX + cdX^{-1}}{gC + h'S + bdX + acX^{-1}} \tag{2-25e}$$

and

$$T = \frac{16\, n_0 n_s (n_1^2 + k_1^2)}{gC + h'S + bdX + acX^{-1}} \tag{2-25f}$$

where

$$X \equiv \exp(4\pi \lvert k_1 \rvert h_1 / \lambda), \tag{2-25g}$$

$$C \equiv \cos(4\pi n_1 h_1 / \lambda), \quad S \equiv \sin(4\pi n_1 h_1 / \lambda) \tag{2-25h}$$

and the remaining constants *a, b, c* and so on appear in Table 2-7 [Abelès 63]. This obtains at normal incidence—the substrate is nonabsorbing.

2.4.2.9.1 Weakly absorbing layer

Weakly absorbing is defined rather arbitrarily as meaning that its k is 0.1 or less. An example of such a layer is germanium near its absorption edge at a wavelength of 1.5 μm. Insight into the behavior of this type of film is obtained by simplifying Eq. (2-25f). An expression for the absorption/transmittance ratio is derived. When $k_1 << 1$, expansion of the X function in Eq. (2-25g) into a MacLaurin series [i.e., $\exp(x) \approx 1 + x + \frac{1}{2}x^2$] produces

$$\frac{A}{T} \approx \frac{k_1 [(n_1^2 - n_s^2) S + 2\beta_1 (n_1^2 + n_s^2) + 2 \lvert k_1 \rvert n_s (C - 1 + 2\beta_1^2)]}{2 n_s (n_1^2 + k_1^2)} \tag{2-25i}$$

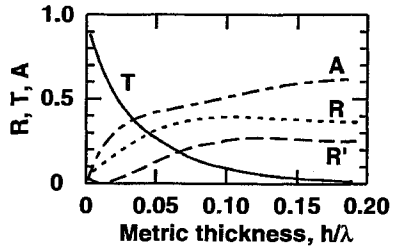


Fig. 2-49 Transmittance T (*solid*), reflectance from air R (*short dash*), reflectance from glass R' (*long dash*) and absorption from the air side A (*long-short dash*) vs metric thickness (in waves) of a film of optical constant $\hat{n}_1 = 1.79 - j\,1.86$ deposited on a glass substrate of index 1.52.

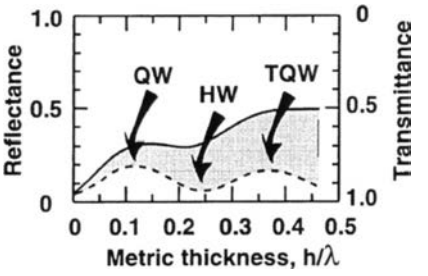


Fig. 2-50 Transmittance T (*solid* curve, right ordinate, plotted downwards) and reflectance R (*dashed* curve, left ordinate) vs metric thickness (in waves) of a film of optical constant $\hat{n}_1 = 2.0 - j\,0.1$ deposited on a substrate of index 1.52, with air incident medium. The shaded area is the absorption. Optical thicknesses: QW = quarterwave; HW = halfwave; and TQW = three quarterwave.

As an example, Fig. 2-49 plots R, T and absorption A vs thickness of a metallic film. Its optical constants are those of nickel in the visible part of the spectrum. Its transmittance monotonically decreases as the layer grows thicker—this is due to the predominance of the X term [i.e., Eq. (2-25g)] in the equations above. The reflectance and absorption both grow larger with increasing thickness until the curves flatten. There is a slight maximum in the reflectance curve attributed to optical interference.

where C and S appear in Eq. (2-25h) [Smith 78]. β_1 appears in Eq. (2-4d)—with $\cos\theta_i = 1$ and $i = 1$. This expression is useful in assessing the extinction coefficient k_1 of a weakly absorbing film from reflectance and transmittance measurements.

Figure 2-50 presents a convenient way of visualizing the absorption. The index of the layer is 2.00, its k is 0.1 and the substrate is BK-7. Reflectance is plotted upwards and transmittance, downwards. Notwithstanding the absorption (shaded area between the curves), optical interference is clearly taking place, as evidenced by the maxima and minima in the reflectance curve. The maxima at the quarterwave (QW) and three-quarterwave (TQW) positions are evident.

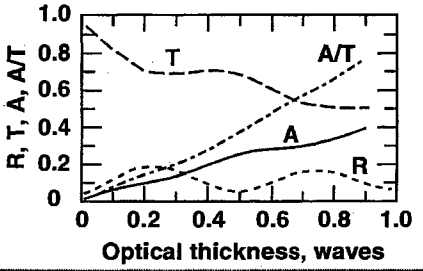


Fig. 2-51 The caption to Fig. 2-50 obtains, with the exception that the curves represent transmittance T (*long dash*), reflectance R (*short dash*), absorption A (*solid*) and A/T (*long-short dash*).

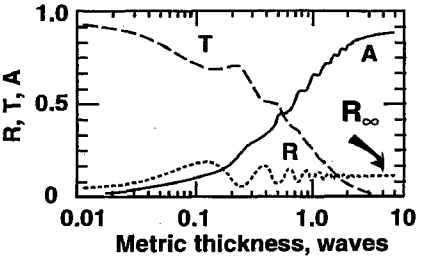


Fig. 2-52 The caption to Fig. 2-50 obtains, with the exceptions that the abscissa is metric thickness and its scale is logarithmic.

Fig. 2-51 graphs R, T, A and A/T vs h_1/λ of this layer. Its optical thickness is close to 90° when h_1/λ is 0.25 and its first maximum R_{max} occurs. As shown in Eq. (2-25i), the A/T curve is dominated by the linear dependence on β and hence A/T increases at least linearly with h_1/λ. In fact, at $\beta = 90°$, the S term in Eq. (2-25i) is zero and for $n_1 \approx 2$ and $n_s \approx 1.5$,

$$A/T \approx 1.5\, k_1 \ . \tag{2-26a}$$

Consider the case where a weakly absorbing layer of refractive index greater than 2.0 is deposited on a relatively low (i.e., BK-7 glass) substrate. When the $\beta = 90°$, the transmittance is close to 70%, which is nearly the reciprocal of 1.5. A rule of thumb is:

> For a film of quarterwave optical thickness whose refractive index is in the neighborhood of 2.0 deposited upon on a BK-7 substrate, its radiant absorption approximately equals its k value.

As an example, if the k_1 of the film is 0.01, then its absorption is close to 1%.

The final question relates to the R, T and A curves when the layer is many waves thick, as in Fig. 2-6. Fig. 2-52 illustrates that the optical interference effects in the reflectance are damped out and R approaches a saturation value R_{sat}, which is

$$R_{sat} \approx (1 - n_1)^2\,(1 + n_1)^{-2} \approx (1 - 2.0)^2\,(1 + 2.0)^{-2} \approx 0.111 \approx 11\% \tag{2-26b}$$

where the k in Eq. (2-12e) is neglected. In other words, the reflectance is that of a bulk substrate with an index of 2.00—that of the film. Figure 2-53 helps to explain this phenomenon. In order for optical interference to take place in this layer, there must be a wave that can reach the bottom interface, where its amplitude reflection coefficient is r_1. Due to the thickness and k value of the layer, however, that wave (*dashed* in Fig. 2-53) is severely attenuated. It follows that the principal reflection is at the outer interface.

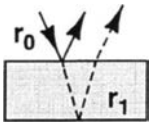


Fig. 2-53 Path of a ray that reflects from the upper interface. The *dashed* ray penetrates the layer to the bottom interface, reflects from that interface and emerges..

The wave that reflects at the film-substrate interface is so highly attenuated that its contribution to the reflectance is negligible.

Valeev [63, 64] discusses the effects of k_1 (for $k_1 << 1$) on R and T at the maxima and minima.

2.4.3 Use of the E^+E^- matrix

2.4.3.1 Introduction

§2.4.2.3 develops the matrix for a single layer. This matrix is generalized to represent the ith layer in a stack. The reflectance and transmittance of a multilayer are found.

2.4.3.2 Alternative form of the E^+E^- matrix

All parameters appearing in the E^+E^- matrix in Eq. (2-22b) have a physical interpretation, such as a Fresnel coefficient, phase of retardation and so on. A simplified form is obtained by substituting Eqs. (2-8f) into Eq. (2-22b). The result is

$$\mathbf{N}_i \equiv (2\,\hat{n}_{i-1})^{-1} \begin{bmatrix} (\hat{n}_{i-1} + \hat{n}_i)\exp(j\beta_i) & (\hat{n}_{i-1} - \hat{n}_i)\exp(-j\beta_i) \\ (\hat{n}_{i-1} - \hat{n}_i)\exp(j\beta_i) & (\hat{n}_{i-1} + \hat{n}_i)\exp(-j\beta_i) \end{bmatrix} . \tag{2-26c}$$

Physical insight is lost but computational efficiency is gained by avoiding the division in the computation of r and t. This efficiency is important when the layers are absorbing.

2.4.3.3 R and T of a stack using the E^+E^- matrix

Figure 2-54 illustrates the arrangement of the incident medium, ℓ layers (numbered from left to right) and the emergent medium, i.e., the substrate. The first step is to generate the matrix in Eq. (2-26c) for each of the ℓ layers. There is a total of $\ell+1$ matrices. The final matrix is

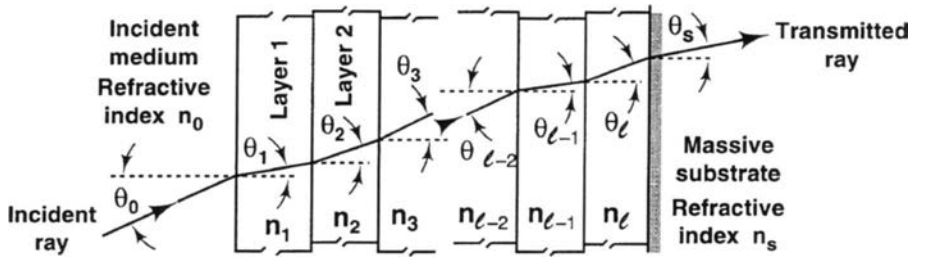


Fig. 2-54 A ray in the incident medium (at an angle of incidence θ_0) penetrates ℓ layers and thence the emergent medium (substrate). The reflections of the ray at the interfaces are omitted for clarity.

$$\mathbf{N}_{\ell+1} \equiv (2\,\hat{n}_\ell)^{-1} \begin{bmatrix} (\hat{n}_\ell + \hat{n}_s) & (\hat{n}_\ell - \hat{n}_s) \\ (\hat{n}_\ell - \hat{n}_s) & (\hat{n}_\ell + \hat{n}_s) \end{bmatrix} \tag{2-26d}$$

where $\hat{n}_s$ is the optical constant of the substrate. The *product matrix* $\mathbf{N}$ is generated

$$\mathbf{N} = \begin{bmatrix} \hat{q}_{11} & \hat{q}_{12} \\ \hat{q}_{21} & \hat{q}_{22} \end{bmatrix} = \prod_{i=1}^{\ell+1} \mathbf{N}_i \ . \tag{2-26e}$$

The manipulation of matrices is discussed in §2.6. In general, the matrix elements $\hat{q}_{ij}$ are complex. The $\mathbf{N}$ matrix links electric field components such as $\hat{E}_0^{(+)}$ in the incident medium to those in the emergent medium — the substrate

$$\begin{bmatrix} \hat{E}_0^{(+)} \\ \hat{E}_0^{(-)} \end{bmatrix} = \mathbf{N} \begin{bmatrix} \hat{E}_s^{(+)} \\ \hat{E}_s^{(-)} \end{bmatrix} = \mathbf{N} \begin{bmatrix} \hat{E}_s^{(+)} \\ 0 \end{bmatrix} . \tag{2-26f}$$

$\hat{E}_s^{(+)}$ is the wave amplitude in the substrate at the interface between it and the last layer. The physical interpretation of the boundary condition $\hat{E}_s^{(-)} = 0$ is that there is no reflected wave propagating in the substrate. The amplitude reflection coefficient $\hat{\rho}$ of the stack is

$$\hat{\rho} = \frac{\hat{E}_0^{(-)}}{\hat{E}_0^{(+)}} = \frac{\hat{q}_{21}}{\hat{q}_{11}} \tag{2-26g}$$

and its radiant reflectance is

$$R = \frac{\hat{q}_{21}\ \hat{q}_{21}^*}{\hat{q}_{11}\ \hat{q}_{11}^*} \tag{2-26h}$$

where the asterisk denotes the complex conjugate. Given the real and imaginary parts of $\hat{q}_{21}$ and $\hat{q}_{11}$ in Eq. (2-26g), the phase shift upon reflection is found from Eqs. (2-18c) and (2-18d).

The radiant transmittance is

$$T = \frac{n_s\, n_0^{-1}}{\hat{q}_{11}\ \hat{q}_{11}^*} \ . \tag{2-27a}$$

At nonnormal incidence the electric fields in Eq. (2-26f) represent the components parallel to the interfaces. Effective indices may be substituted for any $\hat{n}$ appearing in Eq. (2-26c) not associated with β.

$$\hat{n}_i \rightarrow \hat{n}_i / \hat{\gamma}_i \tag{2-27b}$$

for TM polarization and

$$\hat{n}_i \rightarrow \hat{n}_i \hat{\gamma}_i \tag{2-27c}$$

for TE polarization. $\hat{\gamma}_i$ appears in Eq. (2-3h). Substitution of *effective indices* provides the correct δ, ρ, R and T at $\theta_0 \neq 0$, but *not* the correct τ in Eq. (2-22f).

It is noted that in the computation of the matrix product [see Eq. (2-26e)], each element of $\mathbf{N}_i$ is, in general, complex. Hence, 32 multiplications and 24 additions are required to multiply a matrix by another matrix. If the layers are nonabsorbing, however, the computational labor is reduced by the relationships

$$\hat{q}_{12} = \hat{q}_{21}^* \qquad \text{and} \qquad \hat{q}_{11} = \hat{q}_{22}^* \ . \tag{2-27d}$$

In other words, cross-diagonal elements are complex conjugates of each other. A consequence is that only one row of the matrix need be generated; the other is found from the above.

2.4.3.4 Kard formulation of the E^+E^- matrix

This theory seems at first glance sophisticated and unnecessarily complicated since the formulas involve hyperbolic functions. The advantage of this method is that approximations are made with facility. There are numerous applications in antireflection coating design.

This theory was published by Kard [57] and is summarized in his text [71]. The starting point is Eq. (2-21d) that transforms the field components $\hat{E}^{(+)}$ and $\hat{E}^{(-)}$ across an interface. As in §2.4.3.2, it is rewritten to remove the ones from the diagonal. At the interface between *medium* 0 and *medium* 1 this matrix is

$$\mathbf{F}_{01} = \sqrt{\frac{n_1}{n_0}} \begin{bmatrix} \dfrac{n_0 + n_1}{2\sqrt{n_0 n_1}} & \dfrac{n_0 - n_1}{2\sqrt{n_0 n_1}} \\ \dfrac{n_0 - n_1}{2\sqrt{n_0 n_1}} & \dfrac{n_0 + n_1}{2\sqrt{n_0 n_1}} \end{bmatrix} . \tag{2-27e}$$

The matrix $\mathbf{F}_{ij}$ is formed at the interface between the ith and j^{th} layer by replacing 0 by i and 1 by j in the foregoing. The E^+E^- matrix for a single layer is written

$$\begin{bmatrix} \hat{E}_0^{(+)} \\ \hat{E}_0^{(-)} \end{bmatrix} = \mathbf{F}_{01}\,\mathbf{P}_1\,\mathbf{F}_{12} \begin{bmatrix} \hat{E}_2^{(+)} \\ \hat{E}_2^{(-)} \end{bmatrix} \tag{2-27f}$$

where the $\mathbf{P}_1$ matrix appears in Eq. (2-21h)—with $i = 1$. Lest there be any confusion, the E^+E^- matrix for a two-layer stack is written out in detail in Table 2-8.

From Eqs. (2-26h) and (2-27a), R/T is

$$F = R/T = \hat{c}_{21}\,\hat{c}_{21}^{*}\,n_0/n_s \tag{2-28a}$$

where the elements of the $\mathbf{C}$ matrix appear in Table 2-8. Kard noted that the factor $\sqrt{n_0/n_s}$ can be omitted and the $\mathbf{G}$ matrix is defined

Table 2-8 Kard matrices for a two-layer stack

$$\sqrt{\frac{n_3}{n_0}} \begin{bmatrix} \dfrac{n_0+n_1}{2\sqrt{n_0 n_1}} & \dfrac{n_0-n_1}{2\sqrt{n_0 n_1}} \\ \dfrac{n_0-n_1}{2\sqrt{n_0 n_1}} & \dfrac{n_0+n_1}{2\sqrt{n_0 n_1}} \end{bmatrix} \mathbf{P}_1 \begin{bmatrix} \dfrac{n_1+n_2}{2\sqrt{n_1 n_2}} & \dfrac{n_1-n_2}{2\sqrt{n_1 n_2}} \\ \dfrac{n_1-n_2}{2\sqrt{n_1 n_2}} & \dfrac{n_1+n_2}{2\sqrt{n_1 n_2}} \end{bmatrix} \mathbf{P}_2 \begin{bmatrix} \dfrac{n_2+n_3}{2\sqrt{n_2 n_3}} & \dfrac{n_2-n_3}{2\sqrt{n_2 n_3}} \\ \dfrac{n_2-n_3}{2\sqrt{n_2 n_3}} & \dfrac{n_2+n_3}{2\sqrt{n_2 n_3}} \end{bmatrix}$$

$$= \mathbf{F}_{01}\,\mathbf{P}_1\,\mathbf{F}_{12}\,\mathbf{P}_2\,\mathbf{F}_{23} = \begin{bmatrix} \hat{c}_{11} & \hat{c}_{12} \\ \hat{c}_{21} & \hat{c}_{22} \end{bmatrix}$$

The product of the five matrices relates the column vector $[E_0^{(+)}, E_0^{(-)}]$ representing the electric field components in the incident medium to the column vector $[E_3^{(+)}, E_3^{(-)}]$ representing the corresponding components in the substrate of index n_3. $\mathbf{P}_1$ and $\mathbf{P}_2$ appear in Eq. (2-21h), with $i = 1$ and $i = 2$, respectively.

$$\mathbf{G}_{ij} \equiv \begin{bmatrix} \dfrac{n_i+n_j}{2\sqrt{n_i n_j}} & \dfrac{n_i-n_j}{2\sqrt{n_i n_j}} \\ \dfrac{n_i-n_j}{2\sqrt{n_i n_j}} & \dfrac{n_i+n_j}{2\sqrt{n_i n_j}} \end{bmatrix}. \tag{2-28b}$$

This has the advantage that when the $\mathbf{G}$ matrices are substituted for the $\mathbf{F}$ matrices in Eq. (2-27f), R/T reduces to

$$F = R/T = \hat{g}_{21}\,\hat{g}_{21}^{*} \tag{2-28c}$$

where $\hat{g}_{21}$ is the lower antidiagonal element of the matrix product

$$\mathbf{G} = \begin{bmatrix} \hat{g}_{11} & \hat{g}_{12} \\ \hat{g}_{21} & \hat{g}_{22} \end{bmatrix}. \tag{2-28d}$$

The $\mathbf{G}$ matrix simplifies the equations at the expense of a loss of physical significance. Both the radiant reflectance and the amplitude transmission coefficient [see Eq. (2-22f)] can be extracted from the $\mathbf{F}$ matrix. This is precluded with the $\mathbf{G}$ matrix formulation.

The following relationships can be verified for the $\mathbf{G}$ matrices

$$\mathbf{G}_{i-1,i}\;\mathbf{G}_{i,i+1} = \mathbf{G}_{i-1,i+1} \tag{2-28e}$$

and

$$\mathbf{G}_{ik}^{-1} = \mathbf{G}_{ki} \ , \tag{2-29a}$$

where the superscript "–1" denotes the inverse matrix. In other words, the inverse is formed by interchanging rows and columns. These theorems are used to factor the matrices for a two-layer stack

$$\mathbf{G} = \mathbf{G}_{03}\,(\mathbf{G}_{31}\,\mathbf{P}_1\,\mathbf{G}_{13})\,(\mathbf{G}_{32}\,\mathbf{P}_2\,\mathbf{G}_{23}) \ . \tag{2-29b}$$

The natural log of the refractive index u_i is defined as

$$u_i \equiv \tfrac{1}{2} \ln (n_i \,/\, n_s) \tag{2-29c}$$

for $i = 0, 1, 2, \ldots \ell\text{-}1, \ell$. $n_s \equiv n_{\ell+1}$ is the index of the emergent medium — the substrate. The $\mathbf{G}$ matrices are written for the ith interface of a two-layer stack ($i = 0, 1, 2$)

$$\mathbf{G}_{i3} \equiv \begin{bmatrix} \cosh(u_i) & -\sinh(u_i) \\ -\sinh(u_i) & \cosh(u_i) \end{bmatrix} \tag{2-29d}$$

and

$$\mathbf{G}_{3i} \equiv \begin{bmatrix} \cosh(u_i) & \sinh(u_i) \\ \sinh(u_i) & \cosh(u_i) \end{bmatrix} . \tag{2-29e}$$

With the help of hyperbolic double angle formulas, it follows that for a single layer

$$\mathbf{G} = \mathbf{G}_{02}\ \mathbf{K}_1 \tag{2-29f}$$

and for two layers

$$\mathbf{G} = \mathbf{G}_{03}\,\mathbf{K}_1\ \mathbf{K}_2 \tag{2-29g}$$

where the Kard matrix $\mathbf{K}_i$ for the ith layer is

$$\mathbf{K}_i \equiv \mathbf{I}\cos\beta_i + j\sin\beta_i \begin{bmatrix} \cosh(2u_i) & \sinh(2u_i) \\ -\sinh(2u_i) & -\cosh(2u_i) \end{bmatrix} \tag{2-29h}$$

where $\mathbf{I}$ is the unit matrix

$$\mathbf{I} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} . \tag{2-30a}$$

For a stack of ℓ layers, there is a residual matrix $\mathbf{G}_{0\ell}$ containing the refractive index of the incident medium and substrate. Otherwise, there are the $\mathbf{K}$ matrices — ℓ of them — each of which contains the phase thickness and refractive index of an individual layer. The Kard matrix is utilized in §2.7.4.3 to find the R/T polynomial of a multilayer in which the phase thicknesses are equal.

2.4.3.4.1 Principle of refractive index scaling

This mathematical manipulation above has hopefully provided insight. Equation (2-29c) shows that all of the refractive indices of a nonabsorbing stack can be expressed as a *ratio* to that of the substrate. The reflectance of the stack is unchanged as long as this ratio remains constant. Consider a two-layer stack with refractive indices *air*, 1.5, 1.7, 2.1 (where 2.1 is the substrate index). If the phase thicknesses of the layers remain unchanged, the stack has the same reflectance as the design 2.0, 3.0, 3.4, 4.2 (where 2.0 and 4.2 are the indices of the incident medium and substrate, respectively). This is termed *scaling* of the indices.

2.4.4 Characteristic matrix of a homogeneous layer

2.4.4.1 Definition of the characteristic matrix

There is a matrix — the characteristic matrix — that relates the total electric field E and total magnetic field H. Consider first the case of TE polarized flux. By definition, the total electric field in the ith medium is the sum of the incident and reflected components in Eqs. (2-3a) and (2-3b). For the electric field $\hat{E}_i$ this means

$$\hat{E}_i \equiv \hat{E}_i^{(+)} + \hat{E}_i^{(-)} \ . \tag{2-30b}$$

A similar relationship holds for the magnetic field H

$$\hat{H}_i \equiv \hat{H}_i^{(+)} + \hat{H}_i^{(-)} \ . \tag{2-30c}$$

By the use of Eq. (2-15c), the foregoing is transformed to

$$\mathcal{Z}_0 \hat{H}_i = \hat{\gamma}_i n_i \left(\hat{E}_i^{(+)} - \hat{E}_i^{(-)} \right) \ . \tag{2-30d}$$

Equations (2-30b) and (2-30d) are written in matrix form

$$\begin{bmatrix} \hat{E}_i \\ \mathcal{Z}_0 \hat{H}_i \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \hat{\gamma}_i n_i & -\hat{\gamma}_i n_i \end{bmatrix} \begin{bmatrix} \hat{E}_i^{(+)} \\ \hat{E}_i^{(-)} \end{bmatrix} \ . \tag{2-30e}$$

The foregoing equation is inverted and the subscript "i" is replaced by "$i+1$"

$$\begin{bmatrix} \hat{E}_{i+1}^{(+)} \\ \hat{E}_{i+1}^{(-)} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & (\hat{\gamma}_{i+1} n_{i+1})^{-1} \\ 1 & -(\hat{\gamma}_{i+1} n_{i+1})^{-1} \end{bmatrix} \begin{bmatrix} \hat{E}_{i+1} \\ \mathcal{Z}_0 \hat{H}_{i+1} \end{bmatrix} \ . \tag{2-30f}$$

The following steps are executed: (1) both sides of Eq. (2-22a) are multiplied by the matrix in Eq. (2-30e); and (2) the column vector on the left of Eq. (2-22a) is replaced with the right of Eq. (2-30f). The net effect of multiplying these three matrices is to produce a matrix that relates the *total* electric and magnetic fields $\hat{E}_i$ and $\hat{H}_i$ to $\hat{E}_{i+1}$ and $\hat{H}_{i+1}$. At nonnormal incidence, the H field components are the tangential components parallel to the interface. The result is

$$\begin{bmatrix} \hat{E}_i \\ \mathcal{Z}_0 \hat{H}_i \end{bmatrix} = \begin{bmatrix} \cos\beta_i & j(\hat{\gamma}_i \hat{n}_i)^{-1} \sin\beta_i \\ j\hat{\gamma}_i \hat{n}_i \sin\beta_i & \cos\beta_i \end{bmatrix} \begin{bmatrix} \hat{E}_{i+1} \\ \mathcal{Z}_0 \hat{H}_{i+1} \end{bmatrix} \ . \tag{2-31a}$$

The foregoing is called the characteristic matrix because it is *characteristic* of the ith layer. This is evident by inspection of Eq. (2-31a). It contains only the phase thickness β_i, index $\hat{n}_i$ and $\hat{\gamma}_i$ of that layer.

Consider next the TM polarized wave. In comparing Eq. (2-7c) with Eq. (2-10d), it is evident that the tangential components of the E and H fields appear in Eq. (2-31a). A consequence is that Eq. (2-31a) is valid if the substitution $\hat{n}_i \hat{\gamma}_i \rightarrow \hat{n}_i / \hat{\gamma}_i$ is made

$$\begin{bmatrix} \hat{E}_i \\ \mathcal{Z}_0 \hat{H}_i \end{bmatrix} = \begin{bmatrix} \cos\beta_i & j\hat{\gamma}_i \hat{n}_i^{-1} \sin\beta_i \\ j\hat{\gamma}_i^{-1} \hat{n}_i \sin\beta_i & \cos\beta_i \end{bmatrix} \begin{bmatrix} \hat{E}_{i+1} \\ \mathcal{Z}_0 \hat{H}_{i+1} \end{bmatrix} \ . \tag{2-31b}$$

At nonnormal incidence, the E field or H fields in Eq. (2-31a) or (2-31b) are the tangential components — i.e., those parallel to the interface.

2.4.4.2 Characteristic matrix elements for a single layer

The matrix elements for a single absorbing layer are written out at normal incidence. Given its index n_i, extinction coefficient k_i and thickness h_i, its phase thickness $\hat{\beta}_i$ is computed using Eq. (2-5b). β_j [see Eq. (2-5c)] is negative.

The elements of the characteristic matrix are

$$\mathbf{M} = \begin{bmatrix} \hat{m}_{11} & \hat{m}_{12} \\ \hat{m}_{21} & \hat{m}_{22} \end{bmatrix} \equiv \begin{bmatrix} a_{11}+jb_{11} & a_{12}+jb_{12} \\ a_{21}+jb_{21} & a_{22}+jb_{22} \end{bmatrix} \tag{2-31c}$$

where

$$a_{11} = a_{22} = \cos\beta_r \cosh\beta_j \tag{2-31d}$$

and

$$b_{11} = b_{22} = -\sin\beta_r \sinh\beta_j \ . \tag{2-31e}$$

The antidiagonal elements in Eq. (2-31a) are multiplied by j. Additional definitions are

$$A \equiv \sin\beta_r \cosh\beta_j \tag{2-32a}$$

and

$$B \equiv \cos\beta_r \sinh\beta_j \ . \tag{2-32b}$$

It follows that

$$a_{21} = -n_i B + k_i A \tag{2-32c}$$

$$b_{21} = n_i A + k_i B \tag{2-32d}$$

$$a_{12} = (-n_i B - k_i A) / (n_i^2 + k_i^2) \tag{2-32e}$$

$$b_{12} = (n_i A - k_i B) / (n_i^2 + k_i^2) \ . \tag{2-32f}$$

In all cases $k = |k|$ in the foregoing equations. For a nonabsorbing layer, these reduce to

$$\mathbf{M}_i = \begin{bmatrix} m_{11} & jm_{12} \\ jm_{21} & m_{22} \end{bmatrix} = \begin{bmatrix} \cos\beta_i & jn_i^{-1}\sin\beta_i \\ jn_i \sin\beta_i & \cos\beta_i \end{bmatrix} \tag{2-32g}$$

where m_{11}, m_{12}, m_{21} and m_{22} are real. Due to the constant j in Eq. (2-32g), the antidiagonal elements are pure imaginary and the diagonal elements are pure real.

2.4.4.3 Reflectance and transmittance of a multilayer

The characteristic matrix $\mathbf{M}_i$ for each layer is generated and thence the product of the matrices is computed

$$\mathbf{M} = \prod_{i=1}^{\ell} \mathbf{M}_i = \begin{bmatrix} \hat{m}_{11} & \hat{m}_{12} \\ \hat{m}_{21} & \hat{m}_{22} \end{bmatrix} = \begin{bmatrix} a_{11}+jb_{11} & a_{12}+jb_{12} \\ a_{21}+jb_{21} & a_{22}+jb_{22} \end{bmatrix} . \tag{2-33a}$$

A separate matrix product is generated for each linear polarization if the incident angle is nonzero. The reflectance is found most easily in terms of the normalized surface admittance $\hat{Y}$. The characteristic matrix of the stack — [see Eq. (2-33a)] — is

$$\begin{bmatrix} \hat{E}_0 \\ \mathcal{Z}_0 \hat{H}_0 \end{bmatrix} = \mathbf{M} \begin{bmatrix} \hat{E}_s^{(+)} \\ \hat{n}_s \hat{\gamma}_s \hat{E}_s^{(+)} \end{bmatrix} \tag{2-33b}$$

for the TE polarized wave and for the TM

$$\begin{bmatrix} \hat{E}_0 \\ \mathcal{Z}_0 \hat{H}_0 \end{bmatrix} = \mathbf{M} \begin{bmatrix} \hat{E}_s^{(+)} \\ \hat{n}_s \hat{\gamma}_s^{-1} \hat{E}_s^{(+)} \end{bmatrix} \tag{2-33c}$$

where $\hat{E}_s^{(+)}$ is the E field component parallel to the substrate-stack interface. It is necessary to use the boundary condition that there is no reflected wave in the emergent medium—the substrate.

The next step is to find the reflectance using the admittance equation. The matrix elements are computed for the TE wave . $\hat{H}_0$ and $\hat{E}_0$ are found from Eq. (2-33b) and their ratio $\hat{H}_0 / \hat{E}_0$ is the surface admittance for s polarization

$$\hat{Y} \equiv \hat{\mathcal{Z}}_0 H_0 / \hat{E}_0 = \frac{\hat{m}_{21} + \hat{m}_{22} \hat{n}_s \hat{\gamma}_s}{\hat{m}_{11} + \hat{m}_{12} \hat{n}_s \hat{\gamma}_s} \tag{2-33d}$$

whence from Eq. (2-16f) the amplitude reflection coefficient is

$$\hat{\rho}_s = \frac{n_0 \cos\theta_0 \, (\hat{m}_{11} + \hat{m}_{12} \hat{n}_s \hat{\gamma}_s \,) - (\hat{m}_{21} + \hat{m}_{22} \hat{n}_s \hat{\gamma}_s \,)}{DEN} \tag{2-33e}$$

where DEN is complex

$$DEN \equiv n_0 \cos\theta_0 \, (\hat{m}_{11} + \hat{m}_{12} \hat{n}_s \hat{\gamma}_s \,) + \hat{m}_{21} + \hat{m}_{22} \hat{n}_s \hat{\gamma}_s \ . \tag{2-33f}$$

The radiant reflectance $R_s = \hat{\rho}_s \hat{\rho}_s^*$ and the phase shift upon reflection is found from Eqs. (2-18d) and (2-33e). The modulus squared of DEN appears in §2.10.4.3.

Starting with Eqs. (2-71i) and (2-12b), the radiant transmittance is

$$T_s = \frac{\cos\theta_s \, \Re \left[\hat{E}_s^{(+)} \left(\hat{n}_s \hat{E}_s^{(+)} \right)^* \right]}{\cos\theta_0 \, \Re \left[E_0^{(+)} \Re \left(n_0 \hat{E}_0^{(+)} \right)^* \right]} \ . \tag{2-33g}$$

Because n_0 and n_s are pure real, this reduces to

$$T_s = \frac{n_s \cos\theta_s \hat{E}_s^{(+)} \left(\hat{E}_s^{(+)} \right)^*}{n_0 \cos\theta_0 \hat{E}_0^{(+)} \left(\hat{E}_0^{(+)} \right)^*} \ . \tag{2-33h}$$

$\hat{E}_0^{(+)}$ is found from the simultaneous solution of Eqs. (2-30b) and (2-30d)

$$2 \, n_0 \cos\theta_0 \hat{E}_0^{(+)} = \mathcal{Z}_0 \hat{H}_0 + n_0 \cos\theta_0 \hat{E}_0 \tag{2-34a}$$

where "$\cos\theta_0$" is substituted for $\hat{\gamma}_0$, since the incident medium is nonabsorbing. The result is

$$T_s = \frac{4 n_0 \cos\theta_0 \, n_s \cos\theta_s}{DEN \, DEN^*} \ . \tag{2-34b}$$

The product ($DEN \, DEN^*$) appears in §2.10.4.3. Explicit equations for R_s appear in an appendix (§2.10.4).

The computation of R_p and T_p is simplified by observing that "effective indices" appear in the foregoing equations. β_r and β_j appearing in Eq. (2-32a) are precisely the same for both s and p polarizations. Only the $n/\hat{\gamma}$ or $\hat{\gamma}/n$ in the antidiagonal elements of Eq. (2-31b) are different for p and s. These elements are converted to the TM polarization as follows: (1) the matrix for each layer is generated using Eq. (2-31a); (2) the matrix product in Eq. (2-33a) is formed; (3) the "$\hat{n}_s \hat{\gamma}_s$ in Eq. (2-33e) is replaced with "$\hat{n}_s / \hat{\gamma}_s$"; and (4) the "$n_0 \cos\theta_0$" in Eq. (2-33e), Eqs. (2-33f) and (2-33h) is replaced with

"$n_0/\cos\theta_0$". The modulus squared of *DEN* appears in §2.10.4.3. The amplitude reflection coefficient for *p* polarization is

$$\hat{\rho}_p = \frac{n_0 \sec\theta_0 \,(\hat{m}_{11} + \hat{m}_{12}\,\hat{n}_s\,\hat{\gamma}_s^{-1}) - (\hat{m}_{21} + \hat{m}_{22}\,\hat{n}_s\,\hat{\gamma}_s^{-1})}{DEN} \tag{2-34c}$$

where

$$DEN \equiv n_0 \sec\theta_0 \,(\hat{m}_{11} + \hat{m}_{12}\,\hat{n}_s\,\hat{\gamma}_s^{-1}) + (\hat{m}_{21} + \hat{m}_{22}\,\hat{n}_s\,\hat{\gamma}_s^{-1}) \quad . \tag{2-34d}$$

The corresponding equation for transmittance [by analogy with Eq. (2-34b)] is

$$T_p = \frac{4 n_0 \sec\theta_0 \; n_s \sec\theta_s}{DEN \; DEN^*} \tag{2-34e}$$

where *DEN* appears in Eq. (2-34d). The product (*DEN DEN**) appears in §2.10.4.3. Explicit equations for R_p appear in an appendix (§2.10.4).

The explicit equations for *R* and *T* at normal incidence are

$$R = \frac{(A-B)^2 + (C-D)^2}{(A+B)^2 + (C+D)^2} \tag{2-34f}$$

$$T = \frac{4 n_0 \, n_s}{(A+B)^2 + (C+D)^2} \tag{2-34g}$$

where

$$A \equiv n_0 \, a_{11} + n_0 \, n_s \, a_{12} + n_0 \, |k_s| \, b_{12} \tag{2-35a}$$

$$B \equiv a_{21} + n_s \, a_{22} + |k_s| \, b_{22} \tag{2-35b}$$

$$C \equiv n_0 \, b_{11} - n_0 \, |k_s| \, a_{12} + n_0 \, n_s \, b_{12} \tag{2-35c}$$

$$D \equiv b_{21} - |k_s| \, a_{22} + n_s \, b_{22} \tag{2-35d}$$

and where a_{11}, b_{11} and so on appear in Eq. (2-31c). In the case where the substrate is absorbing, *T* represents the fraction of flux flowing into the substrate. In this way, the computer code can find both the total loss [i.e., $(1 - R)$] and the loss in the substrate.

In the special case where all layers and the substrate are nonabsorbing, the eight numbers $a_{11}\; b_{11} \; \ldots \; a_{22}\; b_{22}$ in the foregoing reduce to the four in Eq. (2-32g), with $m_{11} \equiv a_{11}$; $m_{12} \equiv b_{12}$; $m_{21} \equiv b_{21}$; $m_{22} \equiv a_{22}$. The transmittance simplifies to

$$T = \frac{4}{2 + m_{11}^2 \, n_0 \, n_s^{-1} + m_{22}^2 \, n_s \, n_0^{-1} + m_{21}^2 \, n_0^{-1} \, n_s^{-1} + m_{12}^2 \, n_0 \, n_s} \tag{2-35e}$$

and thus depends on the sum of squares of the matrix elements. The foregoing equation utilizes the fact that the matrix is unimodular [see §2.6.2.3)].

2.4.4.4 General form of the characteristic matrix

Abelès [50] has shown that the characteristic matrix can be generalized to a system of layers that are either homogeneous or inhomogeneous. As an illustration, one film might have an index profile given by the equations in Table 2-22, another is homogeneous and still another inhomogeneous with a different profile. Regardless of whether analytical or numerical methods are used to find the wave functions, the general relationship is

$$\begin{bmatrix} \hat{E}_0 \\ \mathcal{Z}_0 \hat{H}_0 \end{bmatrix} = \begin{bmatrix} f_1(z) & f_2(z) \\ f_3(z) & f_4(z) \end{bmatrix} = \begin{bmatrix} \hat{E}_s \\ \mathcal{Z}_0 \hat{H}_s \end{bmatrix} \tag{2-35f}$$

and is similar to Eq. (2-33b), with the exception that f_1 , f_2 , f_3 and f_4 are not necessarily combinations of sines and cosines, as is the case when the films are homogeneous. The

foregoing equation is derived from Maxwell's equations and establishes that the fields $\hat{E}_s$ and $\hat{H}_s$ at the emergent surface of the multilayer are linear combinations of $\hat{E}_0$ and $\hat{H}_0$ at the incident surface. Table 2-22 and Jacobsson [75] list the wave functions in Eq. (2-35f) for several inhomogeneous index profiles.

2.4.5 Recursion methods of computing the reflectance

There are several recursion relationships that provide the amplitude reflection coefficient of a multilayer, which can be either absorbing or nonabsorbing. The reflection recursion and admittance recursion are applied to three nonabsorbing layers deposited on a metallic substrate, as shown in Fig. 2-55.

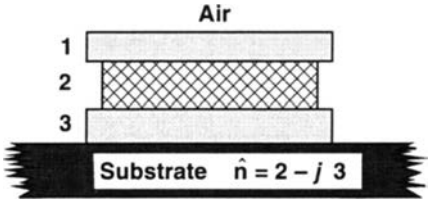


Fig. 2-55 Three nonabsorbing layers reduce the reflectance of a metallic substrate. The layers are drawn proportional to optical thicknesses. The design is listed in Table 2-9.

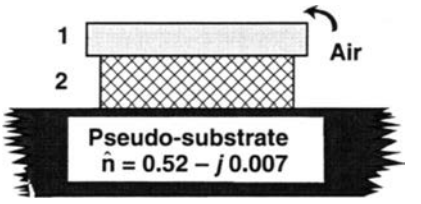


Fig. 2-56 Admittance recursion — the bottom layer and substrate depicted in Fig. 2-55 are replaced with a fictitious pseudo-substrate of admittance $\hat{Y}_2$.

2.4.5.1 Recursion of the admittance

Admittance recursion is widely used to compute the standing wave ratio of an electrical transmission line [see §2.2.5.3). Its optical analog provides the amplitude reflection coefficient of a multilayer. Equation (2-31a) is expanded into linear equations (at normal incidence, with $\hat{\gamma} = 1 + j0$)

$$\hat{H}_{i-1}\,\mathcal{Z}_0 = \hat{E}_i\; j\,n_i \sin\beta_i + \hat{H}_i \cos\beta_i\,\mathcal{Z}_0 \tag{2-36a}$$

$$\hat{E}_{i-1} = \hat{E}_i \cos\beta_i + \hat{H}_i\, j\, n_i^{-1} \sin\beta_i\,\mathcal{Z}_0\ . \tag{2-36b}$$

Equation (2-36a) is divided by Eq. (2-36b). The numerator and denominator of the quotient are divided by $\hat{E}_i$. The result is the normalized admittance [see Eq. (2-16a)]

$$\hat{Y}_{i-1} = \frac{\hat{Y}_i \cos\beta_i + j\,n_i \sin\beta_i}{\cos\beta_i + j\,(\hat{Y}_i / n_i)\sin\beta_i} \tag{2-36c}$$

where $\hat{Y}_i$ is the normalized admittance of the medium (or multilayer) on the emergent side of the film of index n_i and phase thickness β_i ; $\hat{Y}_{i-1}$ is the normalized admittance on the incident side of that layer.

2.4.5.2 Recursion of the amplitude reflection coefficient

The equation for the amplitude reflection coefficient is used recursively. The left of Eq. (2-23a) is the amplitude reflection coefficient $\hat{\rho}_0$ of a single layer (of phase thickness β_1) that has an amplitude reflection coefficient r_1 (i.e., Fresnel reflection) at its lower interface. This is generalized to the ith layer (of phase thickness β_i)

$$\hat{\rho}_{i-1} = \frac{\hat{r}_{i-1} + \hat{\rho}_i\ \exp(-j2\beta_i)}{1 + \hat{r}_{i-1}\,\hat{\rho}_i\ \exp(-j2\beta_i)} \tag{2-36d}$$

where $\hat{r}_{i-1}$ is the Fresnel amplitude reflection coefficient at the upper interface of that layer and $\hat{\rho}_i$ is the internal amplitude reflection coefficient at its lower interface. The latter coefficient is a function of both the (complex) optical constant of that layer and the design of the stack that lies below it—including the substrate. §2.4.5.4 illustrates the use of Eq. (2-36d).

2.4.5.3 Example of recursion — admittance

The foregoing equations are applied to the stack depicted in Fig. 2-55. The application of Eq. (2-36c) begins in the substrate and proceeds through each layer in a progression towards the incident medium. Table 2-9 lists the coating design and summarizes the results at each step.

The normalized admittance of the substrate is $\hat{Y}_3 = 2 - j3$. The bottom layer has an index of 2.05 and phase thickness of 63°. Equation (2-36c) produces a normalized admittance $\hat{Y}_2$

$$\hat{Y}_2 = \frac{(2-j3)\cos 63° + j\,2.05 \times \sin 63°}{\cos 63° + j\,((2-j3)/2.05)\sin 63°} = 0.52 + j0.007 \ . \qquad (2\text{-}36e)$$

As shown in Fig. 2-56, the normalized admittance $\hat{Y}_2$ is used recursively. Thus $\hat{Y}_2$ serves as a "pseudo-substrate" for the next layer (of index 1.45), which is the center layer. Its phase thickness is 101°. The normalized admittance $\hat{Y}_1$ resulting from the addition of this layer is

$$\hat{Y}_1 = \frac{(0.52 + j\,0.007)\cos 101° + j\,1.45 \times \sin 101°}{\cos 101° + j\,((0.52 + j\,0.007)/1.45)\sin 101°} = 3.3 - j1.5 \ . \qquad (2\text{-}36f)$$

The foregoing procedure is repeated until the normalized admittance $\hat{Y}_0$ of the entire stack is found. Equation (2-16h) is used to find $\hat{\rho}_0$. Equation (2-12c) provides the radiant reflectance.

2.4.5.4 Example of recursion — reflection coefficient

The first step is to find the Fresnel amplitude reflection coefficient at the interface between the bottom layer (of index 2.05) and the substrate (optical constant $\hat{n}_s = 2 - j3$). From Eq. (2-8f), this is

$$\hat{\rho}_3 = \frac{2.05 - (2 - j3)}{2.05 + (2 - j3)} = 0.35 - j0.48 \qquad (2\text{-}36g)$$

and the Fresnel coefficient $\hat{r}_2$ at the upper interface of that layer is

$$\hat{r}_2 = \frac{1.45 - 2.05}{1.45 + 2.05} = -0.171 + j0 \ . \qquad (2\text{-}36h)$$

The amplitude reflection coefficient $\hat{\rho}_2$ produced by the bottom layer is

$$\hat{\rho}_2 = \frac{(-0.171 + j0) + (0.35 - j0.48)\exp(-j125.3°)}{1 + (-0.171 + j0)(0.35 - j0.48)\exp(-j125.3°)} = 0.47 - j0.005 \ . \qquad (2\text{-}36i)$$

The center layer is added—$\hat{\rho}_1$ is the amplitude reflection coefficient after this operation. $\hat{\rho}_2$, as given in Eq. (2-36i), is the amplitude reflection coefficient at the bottom interface of that layer. The Fresnel coefficient at the upper interface of the center layer is the negative of the value given in Eq. (2-36h)

$$\hat{\rho}_1 = \frac{(0.171 + j0) + (0.47 - j0.005)\exp(-j202°)}{1 + (0.47 - j0.005)(0.171 + j0)\exp(-j202°)} = -0.28 + j0.2 \ . \qquad (2\text{-}36j)$$

Table 2-9 Admittance recursion and reflection recursion

Nr	Thickness	$\hat{n}$	Mod of $\hat{\rho}$	Phase of $\hat{\rho}$	Admittance
0*b*	*massive**	1.00	0.002	285.3°	$0.999 + j\,0$
1*u*			0.345	359.7°	*same*
1	$0.202\,\lambda_0$	$2.05 + j\,0$			
1*b*			0.345	145.2°	$3.27 - j\,1.46$
2*u*			0.472	158.5°	*same*
2	$0.280\,\lambda_0$	$1.45 + j\,0$			
2*b*			0.472	0°	$0.52 + j\,0.007$
3*u*			0.595	0°	*same*
3	$0.174\,\lambda_0$	$2.05 + j\,0$			
3*b*			0.595	125.6°	$2 - j\,3$
4	*massive*§	$2 - j\,3$	0		*same*

Internal reflection coefficient $\hat{\rho}$ (polar form, modulus *Mod* and phase) and normalized admittance of a three-layer stack at a wavelength λ_0. The optical thickness and optical constant $\hat{n}$ are listed in *columns* 2 and 3. The **boldface** in *column* 1 refers to the layer, *u* to its upper interface and *b* to its lower interface.

*Incident medium §Substrate

Table 2-10 Summary of computational methods

Physical parameters	Matrix formulation	Ratio used in recursion
Total E and H fields	Characteristic matrix. See Eq. (2-31a).	Admittance $\equiv \hat{Y} \equiv \hat{H}/\hat{E}$. See Eq. (2-36c).
Components of electric field $E^{(+)}$ and $E^{(-)}$	E^+E^- matrix. See Eq. (2-26c).	Amplitude reflection coefficient $\hat{\rho} \equiv \hat{E}^{(-)}/\hat{E}^{(+)}$. See Eq. (2-36d).

In Table 2-9 the foregoing procedure is repeated until the incident medium is reached. Given $\hat{\rho}_0$ in that medium, Eq. (2-12c) provides the radiant reflectance.

2.4.6 Comparison of computational methods

2.4.6.1 Overview of methods

Four methods of computing the reflectance of a multilayer are developed in Chap. 2. These are: (a) E^+E^- matrix (§2.4.3); (b) characteristic matrix (§2.4.4); (c) reflection recursion (§2.4.5.2); and (d) admittance recursion (§2.4.5.1). All methods can be applied to stacks containing absorbing layers. The usefulness of these computational methods is assessed for both nonabsorbing and absorbing multilayer stacks. Table 2-10 summarizes the methods.

2.4.6.2 The multilayer contains only nonabsorbing layers

The recursion methods compute the radiant reflectance R and also the phase shift upon reflection δ. Because the stack is nonabsorbing, its radiant transmittance is $T = 1 - R$.

Recursion cannot compute the phase shift upon transmission [alias phase retardation, see Eq. (2-22f)]. Matrix methods must be used to find the latter.

2.4.6.3 The multilayer contains one or more absorbing layers

The recursion methods compute the radiant reflection and phase shift upon reflection. They cannot provide the radiant transmittance, although Berning [63] derives an expression by which the absorption of each layer is computed from the admittance.

The matrix methods allow the computation of: (a) radiant reflectance of *both* sides of the multilayer (these are not necessarily identical if the stack is absorbing); and (b) radiant transmittance.

Advantages of the characteristic matrix method are:

(a) It allows the designer to find the R and T of the same stack deposited upon different substrates, with little additional computational labor.

(b) It is relatively easy to compute the properties of a stack repeated many times (i.e., exponentiated — see §5.8.1).

(c) The fact that the determinant of the matrix equals one is a useful computational check that furnishes an estimate of the truncation errors perpetrated by the arithmetic unit of the computer.

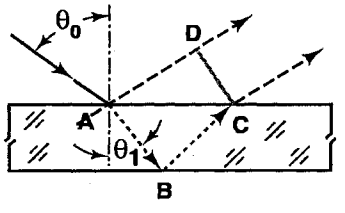


Fig. 2-57 Ray is incident at an angle θ_0 on a layer. The *long-dash* line is the ray reflected at the upper surface. The *short-dash* line is the ray that penetrates to the lower interface, where it is reflected at *B*. It emerges into the incident medium at *C*.

2.5 Design concepts used at nonnormal incidence

2.5.1 Introduction

The formalism for the computation of the reflectance of a multilayer is developed in §2.4. Because this includes nonnormal incidence, why is any further discussion required? There are concepts, however, such as *effective thickness* and *effective index* that furnish insight. These are enumerated below.

2.5.2 Optical thickness and effective thickness

Physical optics texts define optical thickness OT of the ith layer as the product of its index n_i and metric thickness h_i

$$OT \equiv n_i h_i \ . \tag{2-37a}$$

From Eq. (2-4d) the phase of retardation β of a nonabsorbing layer is

$$\beta = (OT_{effective})\, 2\pi / \lambda \tag{2-37b}$$

where the *effective optical thickness* (or *effective thickness*, for short) is

$$OT_{effective} \equiv n_i h_i \cos\theta_i \tag{2-37c}$$

where θ_i is the angle of refraction in that layer. The *reduced* effective thickness is dimensionless. It is the cosine of the angle of refraction in a layer and is a measure of the fractional change of thickness with increasing angle of incidence:

$$OT / (n_i h_i) = \cos\theta_i \ . \tag{2-37d}$$

2.5.2.1 Geometrical interpretation of effective thickness

It is curious that the path in a bulk absorbing medium (see Eq. (2-14a) in §2.2.4.4) *increases* at nonnormal incidence, whereas the *effective thickness* decreases. A geometric picture furnishes insight into this phenomenon.

Consider the case when the wave is propagating in a bulk medium—no optical interference is taking place. The path does, indeed, *increase* at nonnormal incidence, as shown by the slant path $\overline{AB}$ in Fig. 2-57. In the case of optical interference, a simple model [Jenkins 57a] explains why the OT *decreases*. Figure 2-57 shows a normal to the wave front (i.e., ray) of the wave incident on a layer of index n_1. The ray reflected at the top interface is represented by a *dashed* line. Shown as a dotted line is the path of the wave that penetrates the top interface at A, reflects from the bottom interface of the layer at B and emerges into the incident medium at C. The *difference* in optical path determines the relative phase of the aforementioned two components of the reflected wave. Those components interfere—as explained in §2.8.3.1. The optical path difference is

$$OT = [\, n_1 \,(\overline{AB} + \overline{BC}) - n_0 \,\overline{AD}\,] \qquad (2\text{-}37e)$$

where n_0 is the index of the incident medium.

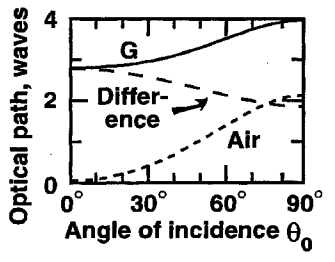


Fig. 2-58 Versus angle of incidence θ_0 in air, optical path in air (*short dash*) and optical path in the film of index 1.38 (*solid, curve G*) for a film of metric thickness λ. The *long-dash* curve is the difference in optical paths.

Figure 2-58 graphs vs θ_0 the optical path in the incident medium "$n_0\,\overline{AD}$" and the optical path in the layer "$n_1\,(\overline{AB} + \overline{BC})$" for $n_0 = 1.00$ and film index n_1 of 1.38. With increasing incidence angle, *both* paths increase, but the *difference* between them decreases—hence, the $\cos\theta_1$ fall-off of optical thickness. In fact, from geometry and Snell's law [see, Eq. (2-4a)], Eq. (2-37c) follows from Eq. (2-37e).

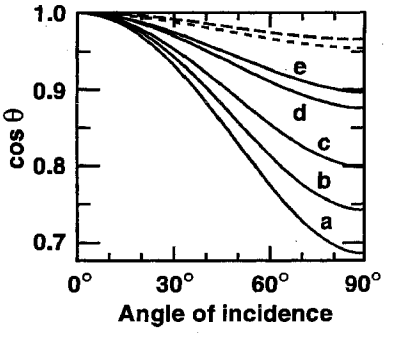


Fig. 2-59 Reduced effective thickness [see Eq. (2-37f)] vs the angle of incidence in air for media of refractive index: (a) 1.38; (b) 1.50; (c) 1.68; (d) 2.10; (e) 2.30; (*short-dash*) 3.48; and (*long-dash*) 4.0.

2.5.2.2 Change of effective thickness with angle

From Eqs. (2-37d) and (2-4a), the reduced effective thickness of the ith medium is

$$\cos\theta_i = \sqrt{1 - (n_0\, n_i^{-1} \sin\theta_0)^2}\ . \qquad (2\text{-}37f)$$

It is plotted vs θ_0 in Fig. 2-59. It is unity at normal incidence and decreases with increasing θ_0. Its change is comparatively small for high index layers and much larger for low index layers such as magnesium fluoride ($n = 1.38$).

In almost all instances the decrease of effective thickness at nonnormal incidence shifts the features of a multilayer coating to shorter wavelengths. By *feature* it is meant a maximum or minimum in its reflectance or the

wavelength at which T equals 50%. This shift occurs for the following reason. Suppose this feature occurs at a wavelength λ_1 at normal incidence. β has a *specific value* in a given layer at that wavelength. The cosine in Eq. (2-4d) decreases as θ_0 grows larger. The wavelength in the denominator must also decrease if β is to maintain the *specific value* mentioned previously. This "angle shift" of bandpass and edge filters to a shorter wavelength is important in filter design and is discussed in §1.6.4.1, §1.8.5.3.1 and §5.3.8. Table 2-5 lists the effective indices of the incident medium, substrate and the single layer.

2.5.3 Effective index

2.5.3.1 Useful applications of effective index

The concept of *effective index* is introduced for two reasons:

(a) Certain equations used at normal incidence are readily converted to nonnormal.

(b) Useful insight is gained.

As an example of the latter, the Fresnel amplitude reflection coefficient [see Eqs. (2-8c), (2-8f) and (2-9c)] at the interface between layers a and b is generalized to

$$r = \frac{1 - \dfrac{n_{a,effective}}{n_{b,effective}}}{1 + \dfrac{n_{a,effective}}{n_{b,effective}}} . \qquad (2\text{-}37g)$$

The ratio of effective indices in the foregoing equation emphasizes that the magnitude of r is relatively small if the indices are close together. In such a case, it would be difficult to construct a reflector with a reflectance of nearly 100% without using many, many layers. This is discussed in more detail in §5.3.8.1.

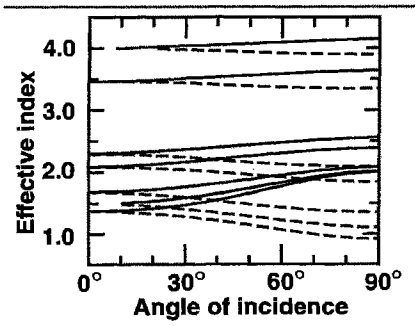


Fig. 2-60 Effective index [see Eq. (2-38b)] for p polarization (*solid*) and effective index [see Eq. (2-38a)] for s polarization (*dashed*) vs angle of incidence in air.

In a sense, *effective index* is a misnomer. Equation (2-70d) defines the refractive index as a velocity ratio. *Effective index* is clearly not such a ratio. Effective index appears in the equations containing the boundary conditions that require the tangential projections of the E and H fields to be continuous at an interface. As an illustration, the substitution of effective index [Eq. (2-8b) or Eq. (2-10b)] into Eq. (2-8a) allows an equation at normal incidence to be adapted to nonnormal incidence.

2.5.3.2 Definitions of effective index

The starting point is the angle θ_0 in the incident medium of index n_0 from which L in Eq. (2-4a) and θ_i in Eq. (2-4b) are computed for the ith medium. As graphed in Fig. 2-60, the effective index N for s polarization is

$$N_{s,i} \equiv n_i \,]_{effective} = n_i \cos\theta_i = n_i \sqrt{1 - L^2 n_i^{-2}} \qquad (2\text{-}38a)$$

and for the p polarization

$$N_{p,i} \equiv n_i \,]_{effective} = n_i \sec\theta_i = n_i \left[\sqrt{1 - L^2 n_i^{-2}} \right]^{-1} . \tag{2-38b}$$

2.5.3.3 Variation with angle, nonimmersed multilayer

Nonimmersed implies that n_0 never exceeds the index of the substrate or any layer. Mathematically, this means that $\theta_i < \theta_{i,c}$, where the critical angle $\theta_{i,c}$ appears in Eq. (2-4e). At normal incidence the effective index reduces to the actual index of the medium. As θ_0 grows larger, N increases for p polarization and decreases for the s, as shown in Fig. 2-60. The percentage increase depends upon the refraction of the ray in the medium. Consider, for example, the N of a low index medium ($n = 1.38$) as θ_0 (in air) increases from $0 \rightarrow 90°$. N increases about 30% for the p polarization and decreases by the same amount for the s polarization. The ray refracts comparatively little in a high index layer such as germanium ($n = 4.00$). The N changes by less than 3% as θ_0 ranges from zero to 90°.

The effective indices approach finite limits as $\theta_0 \rightarrow 90°$

$$N]_{\theta_0 \rightarrow 90°} = \sqrt{n^2 - n_0^2} \tag{2-38c}$$

for s polarization and

$$N]_{\theta_0 \rightarrow 90°} = \frac{n^2}{\sqrt{n^2 - n_0^2}} \tag{2-38d}$$

for p polarization.

2.5.3.4 Inversion of effective index equations

There are coating designs, such as beamdividers, that function at nonnormal incidence. These often require a prescribed effective index N. Then, given N, what is the actual index n that will produce it? For s polarization, inversion of Eq. (2-38a) yields

$$n = \sqrt{N^2 + L^2} \tag{2-38e}$$

and is single valued. For the p polarization, a quadratic relationship is obtained

$$n^4 - n^2 N^2 + L^2 N^2 = 0 . \tag{2-38f}$$

Solving for n provides

$$n = N \sqrt{\tfrac{1}{2}\left[1 \pm \sqrt{1 - 4L^2 N^{-2}}\right]} . \tag{2-38g}$$

In nonimmersed systems — with $L < 1.00$ — the positive root is selected.

Of special importance is the case in which the discriminant in Eq. (2-38g) is zero. For a given L, this provides the *lowest* available N_p . This effective index is

$$N_p = \sqrt{2}\, n \tag{2-39a}$$

and the index n at which this occurs is the single valued function

$$n = \sqrt{2}\, L = \sqrt{2}\, n_0 \sin\theta_0 . \tag{2-39b}$$

Suppose, for example, a designer desires to produce the highest possible reflectance at an incidence angle of 82°. The design consists of alternating high index and low index layers — any index in the range 1.35 to 2.30 is permissible. A given reflectance is attained with the fewest layers when the spread between the effective indices of the high and low is as large as possible. In other words, the magnitude of the Fresnel coefficients [see Eq. (2-37g)] is maximized when the spread is the largest. This,

in turn, provides for the maximum reflectance. Substitution of 82° into Eq. (2-39b) shows that N_p is a minimum when n is 1.40. Therefore, if an index of 1.40 were available, it should be used in preference to 1.35 in order to optimize the p reflectance.

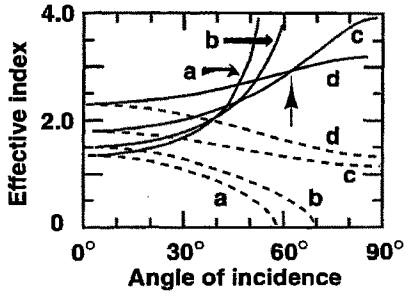


Fig. 2-61 Effective index [see Eq. (2-38b)] for p polarization (*solid*) and effective index [see Eq. (2-38a)] for s polarization (*dashed*) vs angle of incidence θ_0 in a medium of index 1.60. An *internal Brewster angle* occurs at the intersection (arrowed) of two curves.

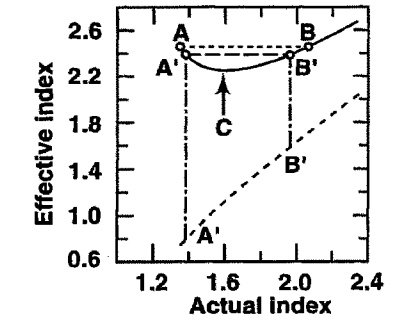


Fig. 2-62 Effective index for s polarization (*dashed*) and p polarization (*solid*) vs the actual index n, as computed from Eqs. (2-38a) and (2-38b) for $L = 1.13$.

2.5.3.5 Effective indices for an immersed system

This section pertains to coatings in a cemented cube, shown in Fig. 1-168, Fig. 1-183 or Fig. 1-196. Dramatic polarization effects usually occur in such systems.

A plot of effective index vs θ_0 for a higher index prism ($n_0 = 1.60$) serves as an example. This is shown in Fig. 2-61. The angle of incidence θ_0 can be varied by changing the external angle (see Fig. 1-196) or by altering the prism angle. The curves of N vs θ_0 in Fig. 2-61 are for films of index 1.35, 1.50, 1.80 and 2.30. The indices of 1.35 and 1.50 are less than n_0 and consequently the critical angle occurs at $\theta_0 = 58°$ and 70°, respectively. The N is zero for the s polarization and approaches infinity for the p, when $\theta_0 \rightarrow \theta_c$. There are angles at which two materials have the same N_p, as shown by the arrowed intersection in Fig. 2-61. This leads to the MacNeille equation. The angle at which this occurs is termed an *internal Brewster angle.*

2.5.3.6 MacNeille equation and the internal Brewster angle

Figure 2-61 is but one format for the display of effective index N vs angle for given values of n. An alternative is to plot N vs n for a given value of L. Figure 2-62 shows such a plot for $L = 1.13$, which corresponds to the configuration in Fig. 2-61 in which $n_0 = 1.60$ and $\theta_0 = 45°$.

As depicted in Fig. 2-62, the N for p polarization (*upper curve*) attains a value of nearly 2.5 when the film index is 1.35. It decreases (with increasing n) to a minimum at C at an n given by Eq. (2-39b). It then increases again and eventually approaches a straight line asymptotically $N \approx n$ for $n >> 1$. For a given N, the n values A and B, or A' and B', correspond to the two roots in Eq. (2-38g). If the N values are equal for n_a and n_b, this requires that

$$n_a^2\left(1-L^2 n_a^{-2}\right)^{-1} = n_b^2\left(1-L^2 n_b^{-2}\right)^{-1} \qquad (2\text{-}39c)$$

which reduces to

$$n_a^{-2} + n_b^{-2} = L^{-2} = \left(n_0 \sin\theta_0 \right)^{-2} . \qquad (2\text{-}39d)$$

Given n_a and n_b, the θ_0 (measured in the incident medium) that satisfies the foregoing (called the *MacNeille equation*) is termed the *internal Brewster angle*. That equation, published by MacNeille [46], is used in §6.3.2.1 to design a linear polarizer.

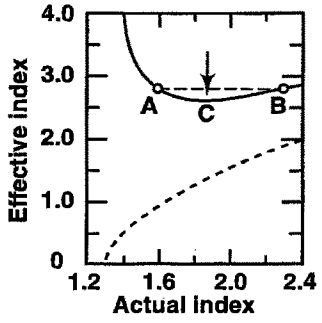


Fig. 2-63 The caption to Fig. 2-62 obtains, with the exception that $L = 1.31$.

The general shape of the $N(n)$ curve is unaltered as L increases. An illustration is that Fig. 2-63 shows $N(n)$ for $L = 1.31$, which corresponds to an n_0 of 1.60 and θ_0 of 55°. The minimum of the $N(n)$ curve shifts to a larger n, as predicted by Eq. (2-39b). Also, the range of n that satisfies Eq. (2-39d) is considerably narrower — given that the maximum attainable index in the visible part of the spectrum is 2.35.

There are a few attainable indices near 1.35 that satisfy Eq. (2-39d) when the incident medium is air, as noted by Catalán [65]. There are few applications for such coatings, however, since the internal Brewster angle occurs when θ_0 is in excess of 80°.

Brewster's angle can be derived using the concept of *effective index*. For TM polarized flux, r_p is zero at the interface between the incident medium (of index n_0) and the emergent medium (index n_s) when the effective indices are equal, as is evident from Eq. (2-37g). Then starting with

$$\frac{n_0}{\cos\theta_0} = \frac{n_s}{\cos\theta_s} \qquad (2\text{-}39e)$$

and Snell's law, Eq. (2-10g) follows. As mentioned earlier, Eq. (2-39e) can be visualized graphically by the intersections of the N_p curves, as depicted in Fig. 2-61.

2.6 Aids to computation

2.6.1 Introduction

§2.1 and §2.4 contain most equations requisite to compute the reflectance and transmittance of a multilayer. This section presents theorems that simplify the calculations and furnish insight.

2.6.2 Matrix relationships

A letter in the type style **A** or **B** represents a two-by-two matrix. a_{ij} and b_{ij} are the elements of the corresponding matrix.

2.6.2.1 (a) Scalar multiplication

If c is a scalar, the multiplication of c by **A** implies that

$$\mathbf{B} = c\,\mathbf{A} \qquad \text{or} \qquad b_{ij} = c\,a_{ij} \qquad (2\text{-}40a)$$

for all i and j. Thus each element of the matrix is multiplied by the scalar.

2.6.2.1 (b) Matrix multiplication

Multiplication is defined as

$$\mathbf{C} = \mathbf{A}\,\mathbf{B} \rightarrow c_{jm} = a_{ji}\, b_{im} \text{ or } c_{jm} = \sum_{i=1}^{2} a_{ji}\, b_{im} \tag{2-40b}$$

where the convention is used: *Sum over any subscript that appears twice.* **B** is termed a *column vector* when the subscript m in Eq. (2-40b) is 1.

2.6.2.1 (c) Distributive property of multiplication

Multiplication of matrices is distributive. This implies that the multiplication can be grouped, as, for example,

$$\mathbf{A}\,\mathbf{B}\,\mathbf{C}\,\mathbf{D} = \mathbf{A}\,(\mathbf{B}\,\mathbf{C}\,\mathbf{D}) = (\mathbf{A}\,\mathbf{B})(\mathbf{C}\,\mathbf{D}) = (\mathbf{A}\,\mathbf{B}\,\mathbf{C})\,\mathbf{D}\ . \tag{2-40c}$$

2.6.2.1 (d) Matrix multiplication is noncommutative

The order in which matrices are multiplied in Eq. (2-40c) is preserved—just the grouping is changed. The matrix product is not the same if the order of multiplication is altered, i.e., commuted. Mathematically, this means that

$$\mathbf{A}\,\mathbf{B} \neq \mathbf{B}\,\mathbf{A}\ . \tag{2-40d}$$

How does this relate to a multilayer? Suppose that two films (represented by the symbols **A** and **B**) are deposited on a metal substrate. If the design is

air **A B** *metal*

then the corresponding matrices are written **AB** and multiplied in that order. Reversal of the order of multiplication to **BA** represents a different design

air **B A** *metal* .

2.6.2.2 Commutative property of the unit matrix and scalars

The unit matrix **I** [see Eq. (2-30a)] has the property that

$$\mathbf{I}\,\mathbf{A}\,\mathbf{B} = \mathbf{A}\,\mathbf{I}\,\mathbf{B} = \mathbf{A}\,\mathbf{B}\,\mathbf{I} \tag{2-40e}$$

and thus commutes with all matrices. The scalar b likewise commutes

$$b\,\mathbf{A} = \mathbf{A}\,b\ . \tag{2-40f}$$

There are several applications. The first is a simplification in the multiplication of the E^+E^- matrices. The scalar $(2\ n_{i-1}\ )^{-1}$ in Eq. (2-26c) is factored out of the matrix product in Eq. (2-26e). If only the reflectance of a stack (either nonabsorbing or absorbing) is to be found, then this scalar is forgotten, since it cancels out.

Bastien [62] suggested how highly absorbing layers are handled. This occurs when the elements of the characteristic matrix for an absorbing film are large because h/λ is large. In such instances $|\beta_j|$ in Eq. (2-5c) often exceeds 11. The result is that the hyperbolic functions in Eqs. (2-31d), (2-31e), (2-32a) and (2-32b) are sufficiently large so that, to a precision of 10 decimal digits, $\cosh(x) \approx \sinh(x) \approx 0.5\ \exp(x)$. The scalar "0.5 $\exp(x)$" is factored from the matrix. A consequence is that the reflectance and phase shift upon reflection are computed correctly. The transmittance is, of course, vanishingly small. The scalar that is factored can be converted into an absorbance using Eq. (1-6a).

2.6.2.3 Determinate and trace of a matrix

The determinant is useful for two reasons. First, its use can simplify equations involving the matrices. It also serves as a computational check. The determinant M of a 2×2 matrix is

$$M = a_{11}\, a_{22} - a_{12}\, a_{21}\ . \tag{2-40g}$$

The determinant of the E^+E^- matrix representing the ith layer [Eq. (2-26c)] is

$$M = n_i / n_{i-1} \ . \tag{2-41a}$$

M is complex when either the ith or the $(i-1)^{th}$ media is absorbing. For both the Kard matrix [see Eq. (2-29h)] and the characteristic matrix [see Eq. (2-31a)],

$$M = 1 + j0 \ . \tag{2-41b}$$

Such matrices are termed *unimodular.*

There are several consequences. First, consider a stack containing absorbing layers. Eight numbers — the real and imaginary parts of $\hat{m}_{11}$, $\hat{m}_{12}$, $\hat{m}_{21}$ and $\hat{m}_{22}$ in Eq. (2-31c) — are needed to represent its characteristic matrix at λ_0. Only six of the eight are independent, however, because of the relationship in Eq. (2-41b).

The determinant is useful as a computational check. For the matrix product

$$\mathbf{Q} = \mathbf{A}\,\mathbf{B}\,\mathbf{C} \tag{2-41c}$$

the relationship between the determinants is

$$M_q = M_a\, M_b\, M_c \tag{2-41d}$$

where the subscripts refer to the appropriate matrices in Eq. (2-41c).

The trace x is the sum of the diagonal elements

$$x \equiv a_{11} + a_{22} \tag{2-41e}$$

and it is used in the exponentiation of a matrix (see §5.8.1).

2.6.2.4 Inverse matrix and stack reversal

The inverse matrix $\mathbf{A}^{-1}$ has the property that when it is multiplied by $\mathbf{A}$, the result is the unit matrix $\mathbf{I}$ [see Eq. (2-30a)]

$$\mathbf{I} = \mathbf{A}^{-1}\,\mathbf{A} \ . \tag{2-41f}$$

The inverse of a 2×2 matrix is

$$\mathbf{M}^{-1} = \begin{bmatrix} \hat{m}_{22} & -\hat{m}_{12} \\ -\hat{m}_{21} & \hat{m}_{11} \end{bmatrix} \frac{1}{M} \tag{2-41g}$$

which is obtained by swapping diagonal elements and changing the signs of the antidiagonal elements. A scalar — the reciprocal of the determinant M — multiplies the matrix.

It is tempting to think that the inverse matrix $\mathbf{M}^{-1}$ is the characteristic matrix representing a stack in which the order of the layers is reversed. For reasons described below, this is true, provided the sign of each diagonal element of the characteristic matrix is reversed.

As an example, Fig. 2-64 shows two layers **A** and **B** whose characteristic matrices are $\mathbf{A}$ and $\mathbf{B}$, respectively. The matrix product

$$\mathbf{M} = \begin{bmatrix} \hat{m}_{11} & \hat{m}_{12} \\ \hat{m}_{21} & \hat{m}_{22} \end{bmatrix} = \mathbf{A}\,\mathbf{B} \tag{2-41h}$$

represents the characteristic matrix of the stack in which the layers are ordered **AB**. The matrix for the stack in which the layers are ordered **BA** is

$$\mathbf{M}_r = \begin{bmatrix} \hat{m}_{22} & \hat{m}_{12} \\ \hat{m}_{21} & \hat{m}_{11} \end{bmatrix} = \mathbf{B}\,\mathbf{A} \tag{2-42a}$$

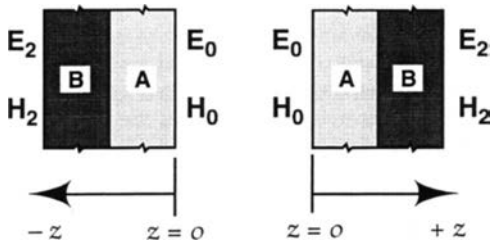


Fig. 2-64 (*Right*) E_0, H_0, E_2 and H_2 are, respectively, the total electric and magnetic fields in the incident medium and emergent medium for the stack *AB*. (*Left*) The order of the layers is reversed and the coordinate system is inverted.

where matrix elements appear in Eq. (2-41h). Equation (2-42a) is obtained from Eq. (2-41h) by swapping diagonal elements. A comparison of these equations shows that the signs of the antidiagonal elements are reversed, for the following reason. It is necessary to reverse not only the *order* of the layers but, in addition, to reverse the direction of the coordinate system. As shown in Fig. 2-64, the matrix **M** transforms the fields $\hat{E}_2$, $\hat{H}_2$ into $\hat{E}_0$, $\hat{H}_0$ at the origin

$$\begin{bmatrix} \hat{E}_0 \\ \mathcal{Z}_0 \hat{H}_0 \end{bmatrix} = \mathbf{M} \begin{bmatrix} \hat{E}_2 \\ \mathcal{Z}_0 \hat{H}_2 \end{bmatrix} . \tag{2-42b}$$

Multiplication of Eq. (2-42b) by $\mathbf{M}^{-1}$ reverses the order of the layers, but transforms the $\hat{E}_0$, $\hat{H}_0$ in the *negative* z direction. In order to transform $\hat{E}_0$, $\hat{H}_0$ to the *positive* z direction, the sign of β in Eq. (2-32g) is made negative for each layer. It can be shown that the net effect is to reverse the signs of the diagonal elements in Eq. (2-41g).

Suppose the stack **AB** is concatenated with a stack **BA**— which is **AB** in reverse order. The resultant design is **ABBA** and contains three layers because the two center **B** layers coalesce into a single layer. Given the characteristic matrix elements of a stack **AB** in Eq. (2-41h), the resultant matrix representing the stack of design **ABBA** is

$$\mathbf{M} = \begin{bmatrix} \hat{m}_{11}\hat{m}_{22} + \hat{m}_{12}\hat{m}_{21} & 2\hat{m}_{11}\hat{m}_{12} \\ 2\hat{m}_{21}\hat{m}_{22} & \hat{m}_{11}\hat{m}_{22} + \hat{m}_{12}\hat{m}_{21} \end{bmatrix} . \tag{2-42c}$$

This illustrates that a symmetrical stack has identical diagonal elements.

2.6.2.5 Derivatives of *R* and *T* using the characteristic matrix

The rules of distributive multiplication are applied to compute the derivatives of *R* or *T* with respect to either thickness or refractive index [Baumeister 62a]. This is illustrated using a stack of six layers, represented by the six layer sequence **ABCDEF**. A nonabsorbing stack is considered, for simplicity. The more complex equations for an absorbing stack are cited by van der Laan [78a]. The first step is to compute the characteristic matrix product **M** of the entire stack

$$\mathbf{M} = \begin{bmatrix} m_{11} & jm_{12} \\ jm_{21} & m_{22} \end{bmatrix} = \mathbf{ABCDEF} . \tag{2-42d}$$

Suppose the derivative of transmittance *T* with respect to the thickness of the first layer is to be found. By the chain rule, this is

$$\frac{\partial T}{\partial h_1} = \frac{\partial T}{\partial m_{11}} \frac{\partial m_{11}}{\partial h_1} + \frac{\partial T}{\partial m_{12}} \frac{\partial m_{12}}{\partial h_1} + \frac{\partial T}{\partial m_{21}} \frac{\partial m_{21}}{\partial h_1} + \frac{\partial T}{\partial m_{22}} \frac{\partial m_{22}}{\partial h_1} . \tag{2-42e}$$

Given the denominator *DEN* of Eq. (2-35e)

$$\frac{\partial T}{\partial m_{11}} = \frac{-8\, n_0\, n_s^{-1}}{(DEN)^2}\; m_{11} \tag{2-42f}$$

and similar expressions are obtained for the derivatives of T with respect to other matrix elements. The calculation of $\partial m_{11}/\partial h_1$ is simplified by noting that h_1 appears *only* in matrix **A** in Eq. (2-42d). This introduces the derivative matrix **M'** (for layer 1)

$$\mathbf{M}' = \begin{bmatrix} \dfrac{\partial m_{11}}{\partial h_1} & j\dfrac{\partial m_{12}}{\partial h_1} \\ j\dfrac{\partial m_{21}}{\partial h_1} & \dfrac{\partial m_{22}}{\partial h_1} \end{bmatrix} \tag{2-42g}$$

where

$$\mathbf{M}' = \mathbf{A}'\,\mathbf{B}\,\mathbf{C}\,\mathbf{D}\,\mathbf{E}\,\mathbf{F} \tag{2-42h}$$

and **A'** is obtained by differentiating each element of the matrix in Eq. (2-32g) with respect to thickness

$$\mathbf{A}' \equiv 2\pi\sigma n_1 \begin{bmatrix} -\sin\beta_1 & j\, n_1^{-1}\cos\beta_1 \\ j n_1 \cos\beta_1 & -\sin\beta_1 \end{bmatrix}. \tag{2-42i}$$

The foregoing procedure is generalized. Equation (2-42h) is written

$$\mathbf{M}' = \begin{bmatrix} \text{pre-matrix} \\ \text{product} \end{bmatrix} \begin{bmatrix} \text{derivative} \\ \text{matrix} \end{bmatrix} \begin{bmatrix} \text{post-matrix} \\ \text{product} \end{bmatrix}. \tag{2-43a}$$

The derivative with respect to layer 3 serves as another example. Layer 3 is the *C* layer, where the derivative matrix is written symbolically as

$$\mathbf{M}' = \mathbf{A}\,\mathbf{B}\,\mathbf{C}'\,\mathbf{D}\,\mathbf{E}\,\mathbf{F} \tag{2-43b}$$

where

$$\text{pre-matrix product} = \mathbf{A}\,\mathbf{B} \tag{2-43c}$$

$$\text{post-matrix product} = \mathbf{D}\,\mathbf{E}\,\mathbf{F} \tag{2-43d}$$

and the matrix **C'** is obtained by replacing the subscript "1" by "3" in Eq. (2-42i).

Now proceed to find the derivative matrix of layer 4. An efficient way of generating the pre-matrix and post-matrix is to multiply

$$\begin{bmatrix} \text{pre-matrix} \\ \text{product} \end{bmatrix}_{new} = \begin{bmatrix} \text{pre-matrix} \\ \text{product} \end{bmatrix}_{old} \mathbf{C} \tag{2-43e}$$

and

$$\begin{bmatrix} \text{post-matrix} \\ \text{product} \end{bmatrix}_{new} = \mathbf{D}^{-1} \begin{bmatrix} \text{post-matrix} \\ \text{product} \end{bmatrix}_{old}. \tag{2-43f}$$

If the generation of the original matrix product **M** is included, a total of $5\ell - 4$ matrix multiplications is required to find the derivatives of all ℓ layers in the stack. $\ell(\ell+1)$ multiplications are required if the derivatives are found from a finite difference approximation. The latter requires more computational labor if the number of layers exceeds three. The same method is used to find the derivatives with respect to the refractive index and extinction coefficient.

The calculation of the derivatives using matrices is but one method. The recursion procedure (see §2.4.5.2) has been implemented [Dupoisot 79].

2.6.2.6 Factoring of a matrix into quaternions

In short, this is a technique of factoring any of the three types of matrices — characteristic, Kard matrix or E^+E^- — into the sum

$$\mathbf{M} = a\,\mathbf{I} + b\,\mathbf{B} + c\,\mathbf{C} + d\,\mathbf{D} \qquad (2\text{-}43g)$$

where **I** is the unit matrix. **A**, **B** and **C** are the Pauli spin matrices, as used in quantum mechanics [Messiah 62]. a, b, c and d are scalars related to the layer thickness and index. Young [61b] and Osterberg [62] discuss quaternion multiplication. This representation has no advantage computationally.

2.6.3 Special layers

From the matrix theory developed previously it follows that the R and T of a multilayer depend upon the thicknesses and indices of each of the layers. There are combinations of layers, however, in which T and R may be independent of the thickness of a particular layer. R and T may also may be independent of the index of a particular layer.

2.6.3.1 Refractive index-independent (absentee) layer

A definition of the absentee layer is:

> An absentee layer is a nonabsorbing film whose optical thickness is $\lambda_0/2$, $2\lambda_0/2$, $3\lambda_0/2$, ... $m\lambda_0/2$ (where m is an integer).

The reflectance and transmittance of the stack are independent of the index of that layer when $\lambda = \lambda_0$. This is proven as follows. At $\lambda = \lambda_0$, the characteristic matrix representing the absentee layer is

$$(-1)^m\,\mathbf{I}\ . \qquad (2\text{-}44a)$$

where **I** is the unit matrix. The foregoing matrix is buried in the product of the matrices representing the layers in the remainder of that stack. According to §2.6.2.2, however, both the unit matrix and the scalar commute with all other matrices and hence the net effect is to multiply the characteristic matrix of the remaining layers of the stack by $(-1)^m$. Equations (2-34f) and (2-35e) show that R and T depend upon the square of the matrix elements and hence altering the sign of all elements in no way affects R or T. In effect, such a layer is *absentee*— a term coined by Osterberg [62] — at this particular wavelength, at least. The fact that a layer is *absentee* depends upon only its optical thickness and the wavelength. It is *independent* of the index of the layer. Only the absentee layer need be nonabsorbing; the remainder of the stack may contain absorbing layers.

Suppose, for example, there is a stack of five layers

air **H L H" L H** *glass*

where **H** and **L** all have a phase thickness β of 90° at λ_0. The β of **H"** at that wavelength is 180°. The characteristic matrix product [see Eq. (2-33a)] is

$$\mathbf{H\,L\,H"\,L\,H}\ .$$

At $\lambda = \lambda_0$ the matrix **H"** reduces to Eq. (2-44a)—with $m = 1$. It is absentee. Starting from that fact that both the scalar and **I** commute (see §2.6.2.2), the foregoing product becomes

$$(-1)\,\mathbf{I\,H\,L\,L\,H}\ .$$

The multilayer now contains two **L** contiguous layers of 90° phase thickness in its center. The contiguous **L** matrices are multiplied and double angle formulas are applied to the

circular functions. The result is that a pair of contiguous **L** layers coalesce into a single layer of 180° phase thickness, which is absentee. The foregoing matrix product is

$$(-1)\,(-1)\,\mathbf{I}\,\mathbf{I}\,\mathbf{H}\,\mathbf{H}$$

where both unit matrices and both scalars are retained for pedagogical reasons. The two remaining **H** matrices, each representing an **H** layer of 90° phase thickness, coalesce to form a 180° thickness layer—it is absentee. At this single wavelength, at least, all layers in the stack are absentee and the transmittance of the multilayer is that of the uncoated substrate. This applies, of course, only if all films are nonabsorbing.

2.6.3.2 Equi-admittance (buffer) layer

This film is the thickness counterpart of the absentee layer. Altering the *index* of the absentee does not change the R or T of the multilayer at a particular wavelength. The equi-admittance layer has the property that the reflectance (but not necessarily the transmittance) of the stack does not depend upon its *thickness*, provided its index is chosen properly. Mouchart [78a] calls this a buffer layer.

> An equi-admittance layer exists: (1) when its refractive index (or optical constant) matches the normalized admittance of the stack (or substrate) on the emergent side of that layer; or (2) when the refractive index of the layer matches that of the incident medium.

The foregoing is true at nonnormal incidence (for a given polarization), provided effective indices are used for the layer (and incident medium) and the normalized admittance is a surface admittance (see §2.2.5.4).

The first two examples are, perhaps, trivial. In both cases, the incident medium is optical cement of index 1.52 and the substrate is a higher index glass of index 1.65. A single film is deposited. If that layer has an index of $n_1 = 1.65$, the *lower* boundary (between the film and substrate) vanishes and its reflectance is independent of the layer thickness.

In the second example, the layer has an index of $n_1 = 1.52$, which matches the index of the incident medium. Because the upper boundary of the layer no longer exists, a change in its thickness cannot alter the reflectance.

As another illustration, a layer (optical thickness $n_2\ h_2 = \lambda_0/4$ and index $n_2 = 2.29$) is deposited on a substrate of index 1.52, as shown in Fig. 2-65. From Eq. (2-36c) its normalized admittance $\hat{Y}$ at $\lambda = \lambda_0$ is pure real and is

$$\hat{Y} = (2.29)^2\,(1.52)^{-1} = 3.45 + j\,0\ . \qquad (2\text{-}44b)$$

When another layer is added on top of it, this normalized admittance is viewed as a *pseudo-substrate* consisting of nonabsorbing material of index 3.45.

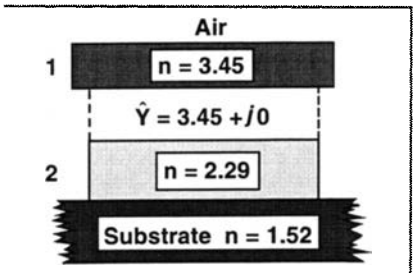


Fig. 2-65 Illustration of a buffer layer. A silicon layer of index 3.45 surmounts a single titania layer of index 2.29. At a wavelength λ_0 the normalized admittance $\hat{Y}_1$ of the titania matches the index of the silicon, thus converting the silicon into a buffer layer.

The next step is to deposit the second layer—a silicon film ($n_2 = 3.45$ in the IR). This is the equi-admittance layer. The reflectance of the two-layer stack is independent of the thickness of the silicon layer, since no Fresnel reflection occurs at the interface between this layer and the pseudo-substrate of the same "refractive index," i.e., normalized admittance. If no further layers were added to the emergent side of the silicon, the R at $\lambda = \lambda_0$ would be that of the single layer of index 2.29. If more layers were added on the *incident* (not the *emergent*) side of the silicon film, the reflectance at $\lambda = \lambda_0$ is independent of its thickness, since it fulfills the equi-admittance conditions stated above.

The final example relates to a method developed by Abelès [49] of measuring the index of a nonabsorbing film. It is usually deposited upon a nonabsorbing substrate and its p reflectance is measured at an angle θ_0. Suppose that θ_0 is chosen so that the effective indices of the incident medium and the film are identical—for p polarization. The layer then fulfills the requirements for equi-admittance. R_p must equal that of the uncoated substrate, regardless of the thickness of the layer. This is equivalent to stating that Brewster's angle occurs at the air-film interface.

As an example, suppose that a layer of index 2.828 is deposited upon a germanium substrate. The Brewster angle θ_B for a medium of air and index 2.828 is [from Eq. (2-10g)] $\theta_B = \arctan(2.828) = 70.5°$. At an angle of 70.5°, the p reflectance of that single layer equals the p reflectance of the uncoated substrate, regardless of the composition of the substrate and thickness of the layer. This angle is shown in panel 1 of Fig. 6-14.

An equi-admittance layer can be absorbing. The (complex) optical constant of the absorbing film must match the normalized admittance $\hat{Y}$ of the medium on the emergent side of the film. The transmittance is *not* independent of the thickness of the equi-admittance layer, although its reflectance is independent.

2.6.4 Partitioning of a multilayer

A multilayer is *partitioned* by slicing it into two sections. The admittance or reflectance of each is determined separately. The sections are combined to determine the transmittance of the original multilayer. This is usually done to gain insight into the performance of the design. Smith [58], for example, designs two cavity bandpass filters by partitioning. Thetford [69] views the three-layer antireflection coating as a partitioned stack.

2.6.4.1 Methods of partitioning a multilayer

Two methods of partitioning a multilayer are illustrated with reference to a specific example depicted in Fig. 2-66.

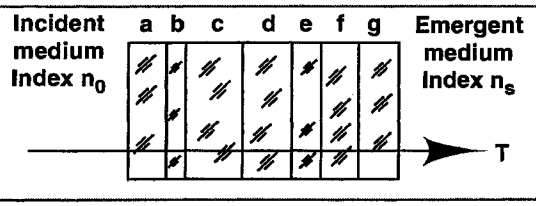


Fig. 2-66 Transmittance T of a seven layer coating is determined by partitioning.

Seven layers (designated by the symbols *abcdefg*) are deposited on a substrate. In the most general case, each layer has a different optical thickness and refractive index. Also, with certain restrictions, any of the layers may be absorbing.

- The first method of partitioning begins by choosing a layer—*c* in this example — and designating it as the *spacer*. The analysis becomes complex if this film is absorbing; for that reason it is assumed to be nonabsorbing. The remaining stacks *ab* and *defg* are designated as *partition* 1 and *partition* 2, respectively.
- In the second approach the layers are lumped into two partitions . The layer sequence *abc* forms *partition* 1 and the stack *defg* is *partition* 2. There is no spacer.

Regardless of which method is used, the transmittance of the original stack can be deduced from the properties of the partitions.

There is a remaining question. Is it necessary that the partitions be symmetrical about the spacer? Clearly, the partitions in the foregoing examples lack such symmetry because they contain unequal numbers of layers—2 and 4, or 3 and 4, depending upon the method of partitioning. This example was chosen to emphasize that it is not mandatory that the stack be partitioned symmetrically. In many instances, however, there are advantages of splitting the layers so that the partitions are symmetrical about a central point.

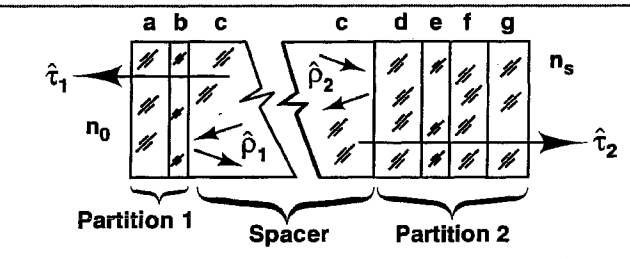


Fig. 2-67 The *c* layer is the spacer; the amplitude reflection coefficients $\hat{\rho}_1$ and $\hat{\rho}_2$, as well as the transmission coefficients $\hat{\tau}_1$ and $\hat{\tau}_2$, of each partition are computed.

2.6.4.2 Partitioning with a spacer layer

This subject is treated in physical optics texts [Stone 63]. Equation (2-45b) describes the transmittance of the Fabry-Perot interferometer (see §7.6 of Born & Wolf [75]) and, hence, a coating analyzed in this manner is often termed a *Fabry-Perot.*

The first step is to find the amplitude reflection coefficient $\hat{\rho}$ and amplitude transmission coefficient $\hat{\tau}$ for each of the partitions. Figure 2-67 establishes the conventions for determining the $\hat{\rho}$ and $\hat{\tau}$ of each partition. For *partition* 1, $\hat{\rho}_1$ is measured using the spacer layer (index n_c) as the incident medium and n_0 as the emergent medium. $\hat{\tau}_1$ and $\hat{\tau}_1'$ are determined in this manner. The substrate is the emergent medium for *partition* 2 and the spacer the incident medium. When this has been done, the partitions

consisting of the two-layer and four-layer stacks are resolved into two fictitious surfaces. They are characterized by the parameters $\hat{\rho}_1$, $\hat{\rho}_2$, $\hat{\tau}_1$, $\hat{\tau}_2$ and so on, as shown in Fig. 2-67.

Because each partition has been reduced to a fictitious surface (called an *effective interface* by Smith [58]), the E^+E^- matrices for such surfaces are represented by Eq. (2-21c). Note, however, that $\hat{\rho}_1$ is on the left of *partition* 1; this is why it appears in the lower antidiagonal element of the following equation

$$\mathbf{N} = \frac{1}{\hat{\tau}_1 \hat{\tau}_2} \begin{bmatrix} 1 & -\hat{\rho}_1 \\ \hat{\rho}'_1 & \hat{\tau}_1\hat{\tau}'_1 - \hat{\rho}_1\hat{\rho}'_1 \end{bmatrix} \mathbf{P} \begin{bmatrix} 1 & -\hat{\rho}'_2 \\ \rho_2 & \hat{\tau}_2\hat{\tau}'_2 - \hat{\rho}_2\hat{\rho}'_2 \end{bmatrix} \tag{2-44c}$$

where $\mathbf{P}$ is obtained by dropping the subscript from β_i in Eq. (2-21h). β is the phase thickness of the spacer and appears below.

The foregoing represents the E^+E^- matrix for the entire stack—i.e., both partitions. Equation (2-27a) provides its radiant transmittance

$$T = \frac{\hat{\tau}_1 \hat{\tau}_1^* \hat{\tau}_2 \hat{\tau}_2^*}{|\exp(j\beta) - \hat{\rho}_1 \hat{\rho}_2 \exp(-j\beta)|^2} \ . \tag{2-44d}$$

If the numerator and denominator of the foregoing are multiplied by the spacer index n_c, this reduces to

$$T = \frac{T_1 T_2}{[1 - \hat{\rho}_1 \hat{\rho}_2 \exp(-j2\beta)]\ [1 - \hat{\rho}_1^* \hat{\rho}_2^* \exp(j2\beta)]} \tag{2-44e}$$

where T_1 and T_2 are the radiant transmittance from the incident medium into the spacer, and from the spacer into the substrate, respectively.

The foregoing equation may be manipulated in several ways. One is to write the $\hat{\rho}$ coefficients in polar form [see Eq. (2-18b)]. Equation (2-44e) becomes

$$T = \frac{T_1 T_2}{1 - 2R_m \cos(2\beta'') + R_m^2} \tag{2-44f}$$

where

$$\beta'' \equiv 2\pi\sigma n_c h_c \cos\theta_m - \delta_m \ . \tag{2-44g}$$

δ_m is the arithmetic mean phase shift upon reflection

$$\delta_m \equiv (\delta_1 + \delta_2)/2 \tag{2-44h}$$

and n_c, h_c and θ_m are the refractive index, metric thickness and angle of incidence in the spacer layer, respectively. R_m is the geometric mean of the radiant reflectance of each partition, as measured from inside the spacer

$$R_m \equiv \sqrt{R_1 R_2} \ . \tag{2-45a}$$

After further manipulation, Eq. (2-44f) becomes

$$T = \frac{T_{max}}{1 + \mathcal{F}\sin^2\beta''} \tag{2-45b}$$

where

$$T_{max} \equiv T_1 T_2 (1 - R_m)^{-2} \tag{2-45c}$$

and

$$\mathcal{F} \equiv 4R_m (1 - R_m)^{-2} \ . \tag{2-45d}$$

In the special case where both partitions are nonabsorbing and equal in reflectance, T_{max} is 100%. The R/T form of Eq. (2-45b) simplifies to

$$F = R/T = F_{max} \sin^2\beta" \ . \tag{2-45e}$$

F_{max} is expressed in terms of the standing wave ratio V of the partition

$$F_{max} = (1 + V)^2 / (4\ V) \ . \tag{2-45f}$$

Equation (2-45b) appears in books on physical optics, whereas Eq. (2-45e) is found in the microwave literature [Matthaei 64]. The foregoing equations are used in an appendix (see §2.11.4.2) to compute the bandwidth of a single-cavity filter.

The equation for the reflectance is considerably more complicated because it involves the phase shift upon reflection of $\hat{\rho}_1'$. It appears in texts such as Born & Wolf [75] and the monograph of Macdonald [71].

2.6.4.3 Partitioning without a spacer layer

In this case, the spacer layer is included in *partition* 1. This is illustrated in Fig. 2-68. $\hat{\rho}_1$, $\hat{\rho}_2$ are measured from a fictitious medium of index n_f. At this point, the question arises — can this index n_f be chosen arbitrarily? In the case where $|\hat{\rho}_1| = |\hat{\rho}_2|$, the answer is yes. If the phase thickness β of the fictitious spacer is zero and if $|\hat{\rho}_1|$ equals $|\hat{\rho}_2|$, Eq. (2-44e) reduces to

$$T = \frac{T_1\ T_2}{\{1 - R\exp[-j(\delta_1 + \delta_2)]\}\ \{1 - R\exp[j(\delta_1 + \delta_2)]\}} \tag{2-45g}$$

where $R \equiv |\hat{\rho}_1|^2$. If the phase shifts upon reflection of the partitions are identical, this is transformed to

$$T = \frac{T_1\ T_2}{1 + Y_j^2\ Y_r^{-2}} \tag{2-45h}$$

where Y_r and Y_j are, respectively, the real and imaginary parts of the admittance of the partitions [Berning 57].

Both formulations — Eqs. (2-45g) and (2-45h) — are useful. Eq. 2-45g shows that if the transmittance is to be optimized, then

$$\hat{\rho}_1 = \hat{\rho}_2^* \ . \tag{2-45i}$$

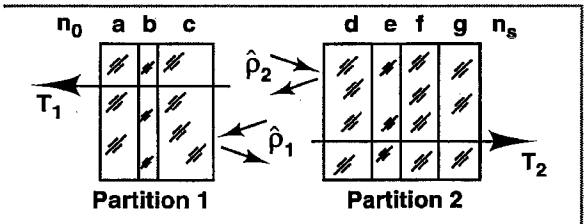


Fig. 2-68 The multilayer in Fig. 2-66 is partitioned without a spacer.

In other words, the amplitude reflection coefficients of the partitions should be complex conjugates of each other. An analysis of Eq. (2-45h) shows that if the transmittance of the partitioned multilayer is to attain a maximum, the admittance of each partition should lie on the real axis of the complex ρ plane, where $Y_j = 0$. The formulation above also answers the question about the refractive index n_f of the fictitious "spacer;" its thickness is set at zero to obtain Eq. (2-45g). The normalized admittance appearing in Eq. (2-45h) is a property of each partition and is independent of the index of the fictitious spacer.

2.6.5 Admittance transformation layers

2.6.5.1 Single layer

Suppose there is a metal surface or possibly a multilayer with a complex normalized admittance $\hat{Y}$. It is overcoated with a single nonabsorbing layer that transforms this complex admittance into a pure real normalized admittance. Given that $\hat{Y}_{i-1} = Y_r + j\,Y_j$ in Eq. (2-36c) and the index n_1 of the layer, setting Y_j to zero provides an equation for its phase thickness β_1

$$\beta_1 = 0.5\left\{\kappa + \arctan\left[\frac{-2\,n_1\,Y_j}{n_1^2 - Y_r^2 - Y_j^2}\right]\right\} \tag{2-46a}$$

where κ is discussed below. The normalized admittance $\hat{Y}_{ov}$ of the overcoated system is

$$\hat{Y}_{ov} = \frac{a \pm \sqrt{a^2 - 4\,n_1^2\,Y_r^2}}{2\,Y_r} + j\,0 \tag{2-46b}$$

where

$$a \equiv n_1^2 + Y_r^2 + Y_j^2 \ . \tag{2-46c}$$

The task remaining is to find the correct β_1 from Eq. (2-46a) and also the proper sign in Eq. (2-46b). The argument of the arctan in Eq. (2-46a) lies in a particular quadrant in the Cartesian plane. An example is that the angle is in quadrant III if the numerator and denominator are both negative. If the argument of the arctan is negative, κ is 180°; otherwise κ is zero. If the argument lies in quadrants I or II, the negative root in Eq. (2-46b) is chosen; otherwise, the positive root is used.

2.6.5.2 Admittance transformer containing two layers

The objective is to transform a real normalized admittance into another real admittance (of a different value) using two nonabsorbing layers. This is equivalent to matching the index of a substrate (index n_s) to an incident medium of index n_0, with zero reflectance. One approach is to select the indices n_1 and n_2 of the two layers. Their phase thicknesses are unknowns and are found from the R/T polynomial for two layers

$$F = R/T = (a_0 \cos\beta_1 \cos\beta_2 - a_2 \sin\beta_1 \sin\beta_2)^2 + (a_{01} \sin\beta_1 \cos\beta_2 + a_{02} \sin\beta_2 \cos\beta_1)^2 \tag{2-46d}$$

where β_1 and β_2 are the phase thicknesses of the outer and inner layers, respectively, and

$$a_{01} \equiv \sinh(u_0 - 2\,u_1) \tag{2-46e}$$
$$a_{02} \equiv \sinh(u_0 - 2\,u_2)$$
$$a_0 \equiv \sinh(u_0)$$
$$a_2 \equiv \sinh(u_0 - 2\,u_1 + 2\,u_2) \ . \tag{2-46f}$$

These equations follow from Eq. (2-29h). The u coefficients appear in Eq. (2-29c). This subject is further developed in §4.10.1.

2.6.6 Optimization

Optimization, which is synonymous with *refining* or *automatic correction*, is the generic term for an algorithm (usually programmed on a digital computer) that improves the spectral performance of a multilayer. This subject is developed in §10.2.

2.6.7 Equivalent layers

2.6.7.1 Introduction

An equivalent layer represents the optical properties of an odd-numbered group of nonabsorbing layers symmetrical about the center film: **A B A**, **A B C B A**, **A B C D C B A** and so on. This concept is useful for two reasons: (1) Analysis—a complex multilayer can often be reduced to one or more equivalent layers, thus enabling the designer to gain insight into its behavior. (2) Synthesis—with certain restrictions, equivalent layers may be synthesized with refractive indices as low as 0.5 and as large as 8 (see §2.6.7.3). Epstein [52] terms this a *Herpin* equivalent layer.

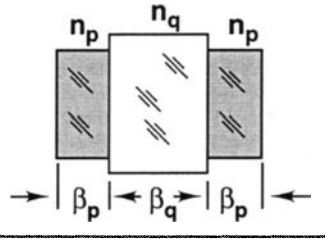


Fig. 2-69 Nomenclature describing a three-layer symmetrical stack in terms of the refractive index n and phase thickness β of each layer.

2.6.7.2 Equivalent layers used in analysis

Fig. 2-69 shows the simplest group (three layers) that may be represented by an equivalent layer. Each outer layer has an index of n_p and phase thickness (see §2.2.2.1) β_p . The phase thickness β_q and index n_q pertain to the inner layer. The total phase thickness β_t of the group is

$$\beta_t \equiv 2\beta_p + \beta_q \ . \qquad \text{(2-46g)}$$

Figure 2-70 illustrates that at a given wavelength λ the three-layer group is represented by a single equivalent layer of phase thickness β_e and index n_e. This is also true at other wavelengths, although its β_e and n_e are different. Figure 2-71 plots the index of the equivalent layer vs frequency (i.e., β_t). Figure 2-72 graphs the thickness β_e (in units of phase) of the equivalent layer. These are computed for indices of 1.38 and 2.30—and a ratio $2\,\beta_p/\beta_q$ of 1.00. If **H** and **L** represent high index and low index layers of quarterwave optical thickness (see §2.7.2.1), respectively, then the foregoing designs are $\frac{1}{2}$**L H** $\frac{1}{2}$**L** and $\frac{1}{2}$**H L** $\frac{1}{2}$**H**. For both the designs, n_e equals $\sqrt{n_H\, n_L}$ at $\beta_t = 0°$. The n_e of the *LHL* design rises with increasing β_t. The converse is true for *HLH*. A useful mnemonic is:

> As β_t increases from zero to 180°, the increase or decrease of the equivalent index "follows" the center index of any *LHL* or *HLH* group.

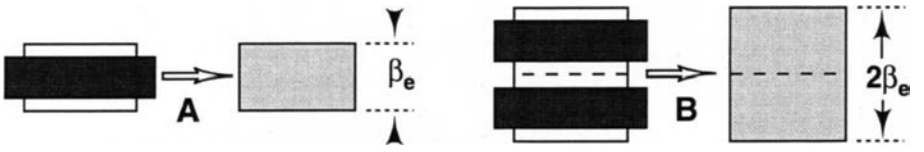


Fig. 2-70 A three-layer symmetrical stack, as represented by an equivalent layer (*left* panel) , is cascaded to form a thicker equivalent layer (*right* panel)

In other words, if the center index is high, n_e rises; if it is low, n_e decreases to zero at the stop band edge. Letters such as *LHL* or *HLH* represent a generic three-layer symmetrical nonabsorbing group—the thicknesses are not necessarily quarterwaves.

As β_t increases, the equivalent index of the *LHL* group continues to rise until it becomes infinite at the stop band edge (defined in §5.2.3). Does such an infinite index

produce a discontinuity in the reflectance curve? Figure 2-72 shows that its phase thickness at the edge is 180°—which in §2.6.3.1 is termed an *absentee* layer. This equivalent layer is thus a seeming anomaly—absentee, yet with infinite refractive index. There is no discontinuity in the reflectance at the stop band edge.

The stop band extends from β_t equals 151° to 209°; in this region the refractive index of the equivalent layer is pure imaginary and is not plotted. §5.4.3.1 and §5.8.10.1 discuss the stop band and the behavior of n_e in that spectral region.

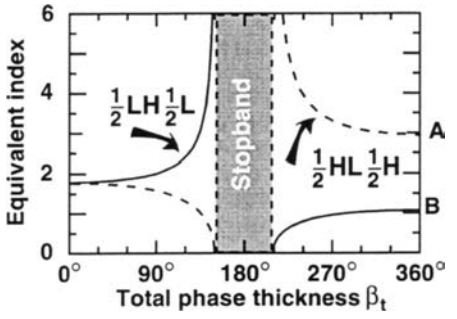


Fig. 2-71 Versus total phase thickness of the group [see Eq. (2-46g)], equivalent index of $\frac{1}{2}$**L H** $\frac{1}{2}$**L** and $\frac{1}{2}$**H L** $\frac{1}{2}$**H**, with $n_L = 1.38$ and $n_H = 2.30$.

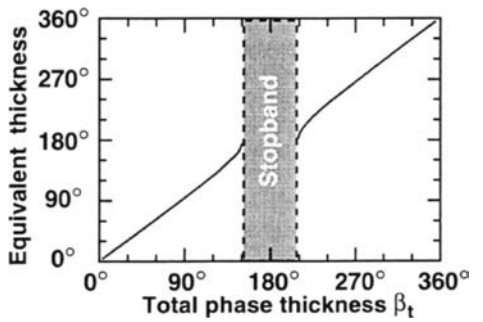


Fig. 2-72 The caption to Fig. 2-71 obtains, with the exception that the ordinate represents the phase thickness of the equivalent layer for both designs.

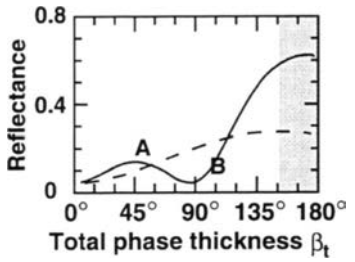


Fig. 2-73 Reflectance vs total phase thickness of a basic period [see Eq. (2-46g)] of (*dashed*)
air 0.5**L H** 0.5**L** *substrate*
and (*solid*)
air (0.5**L H** 0.5**L**)^2 *substrate*
deposited on a substrate of index 1.52. n_L is 1.38 and n_H is 2.30. The shaded area is the stop band.

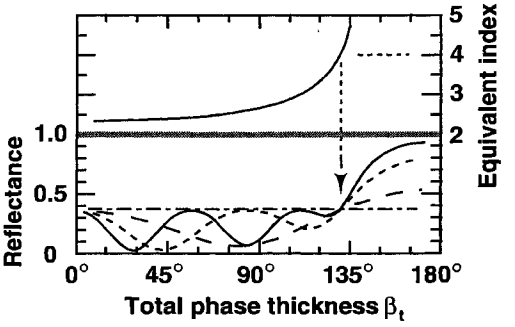


Fig. 2-74 Reflectance (*lower panel*) and index of the equivalent layer (*upper panel*) vs total phase thickness of a basic period [see Eq. (2-46g)] of:
air (0.667**L** 0.667**H** 0.667**L**)^q *Ge*
where indices of **L**, **H** and *Ge* are 1.85, 3.66, and 4.0, respectively. $q = 1$ (*long dash*), $q = 2$ (*short dash*) and $q = 3$ (*solid*). At the arrowed β_t, the equivalent layer index equals that of the substrate. The horizontal *long-short dashed* line is the reflectance of the uncoated substrate.

On the high frequency side of the stop band, the n_e of the $\frac{1}{2}$**L H** $\frac{1}{2}$**L** group starts at zero and increases to a value of $n_L^{3/2}\, n_H^{-1/2} = 1.07$ at $\beta_t = 360°$. Note that n_e is less than 1.0

over an extended bandwidth — a fact often exploited in antireflection coating design. For the $\frac{1}{2}$H L $\frac{1}{2}$H group, n_e decreases from infinity at the stop band edge to $n_H^{3/2}\, n_L^{-1/2} = 2.97$ at $\beta_t = 360°$.

2.6.7.3 Examples: equivalent layers in analysis

2.6.7.3.1 Single and double periods

The first example is a single equivalent layer — the group $\frac{1}{2}$L H $\frac{1}{2}$L — on a relatively low index ($n_s = 1.52$) substrate. Because the index of the equivalent layer exceeds that of the substrate, §2.4.2.6 predicts that a maximum reflectance R_{max} should occur when β_e is 90°, which is attained when β_t is approximately 86°. Fig. 2-73 shows, however, that R_{max} occurs at β_t close to 150°. This shift is attributed to the strong dispersion of the refractive index of the equivalent layer (see Fig. 2-42 in §2.4.2.6.2).

The same type of equivalent layer, but with double thickness, is used for the next example. This is the five-layer stack in Fig. 2-70. When the phase thickness of the basic period is 45°, the thickness of two such periods in tandem is 90°. Then neglecting dispersion of n_e, the maximum reflectance is attained at point *A* in Fig. 2-73. A separate calculation shows that n_e is 1.82 and from Eq. (2-24a) R_{max} is 13.8%. When β_t of a single period is 90°, β_t of the double layer system is 180° and the equivalent layer is absentee. The reflectance is that of the uncoated substrate, near point *B* in Fig. 2-73.

2.6.7.3.2 Variable number of periods

The final example is similar to those above, with the exception that the equivalent layer is composed of nonquarterwaves. An *LHL* group

air (0.667L 0.667H 0.667L)^q *germanium*

is repeated q times. When $\beta_t = 132°$, the n_e of the foregoing matches that of the substrate ($n_s = 4.00$). As shown in Fig. 2-74, the reflectance at $\beta_t = 132°$ is independent of the number of periods. The reason is that the addition of a period is equivalent to thickening the germanium substrate — its reflectance is 36% regardless of its thickness.

Table 2-11 Equivalent layers produced by synthesis

Layer	n	OT$]_A$	OT$]_B$
1	1.38	213 nm	287 nm
2	2.30	331 nm	169 nm
3	1.38	213 nm	287 nm

Designs of equivalent layers in terms of optical thickness OT and index n that produce $n_e = 0.88$ and $\beta_e = 270°$ at $\lambda_0 = 1.0$ μm.

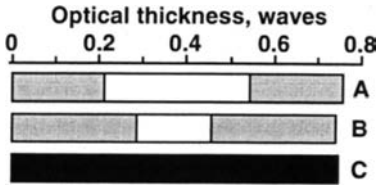


Fig. 2-75 Optical thicknesses of the equivalent layers in Table 2-11. *C* is an actual layer of three quarterwave optical thickness.

2.6.7.4 Synthesis of an equivalent layer — an example

Given the available refractive indices, what layer thicknesses are needed to synthesize an equivalent layer with a given n_e and β_e? Suppose that β_e is 270° and n_e is 0.88 at a wavelength $\lambda_0 = 1$ μm. Table 2-11 cites two designs that fulfill this requirement. Figure 2-75 illustrates that the sum of the optical thicknesses of the equivalent layer is nearly that of a three quarterwave homogeneous layer. The phase thicknesses of the equivalent layers vs wavelength (see Fig. 2-76) of both designs are virtually the same over a limited spectral region. The equivalent index curves in the same figure, however, are slightly different. The dispersion of the equivalent index is considerable, with n_e dropping to low values at wavelengths both longer and shorter than λ_0. Although there are two wavelengths at which n_e is 0.88 for each design, β_e is 270° only at λ_0. Also, there may or may not be two solutions, depending upon the β_e which is specified. When β_e is 90°, for example, there is only one design for the equivalent layer.

2.6.7.5 Equivalent index — its computation

This section reviews methods of computing the properties of the equivalent layer. An electrical engineer views the assembly of thin films as a symmetrical group of electrical transmission lines (see §2.2.5.3, Table 2-3) whose properties may be expressed in terms of image admittance — which is identical to equivalent index.

Image admittance was introduced in the 1920s to design electrical filters [Hund 28, Cauer 41] and was applied to microwave systems in the early 1940s [Fano 48]. Physicists [Herpin 47, Epstein 52] developed the concept independently and applied it to design bandpasses and antireflection coatings.

§5.2.3.2 shows that if a symmetrical nonabsorbing stack is surrounded by an incident medium and emergent medium of the same refractive index, the index of the equivalent layer may be computed from its amplitude reflection coefficient.

2.6.7.5.1 Use of the characteristic matrix

This method has the advantage that it is general and applies to any group of nonabsorbing symmetrical layers. Given the elements of the characteristic matrix (see §2.4.4.2) which, according to §2.6.2.4, has identical elements on its diagonal, the equivalent index is

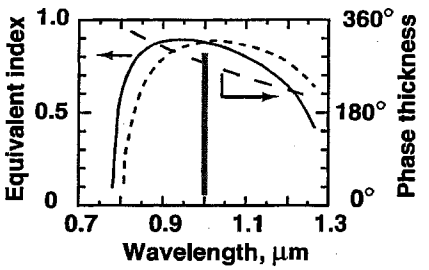


Fig. 2-76 Index (*left ordinate*) of the equivalent layers with the designs *A* (*solid*) and *B* (*short dash*) in Table 2-11. (The *long dash* curve (*right ordinate*) graphs the phase thicknesses of the equivalent layer for both designs.

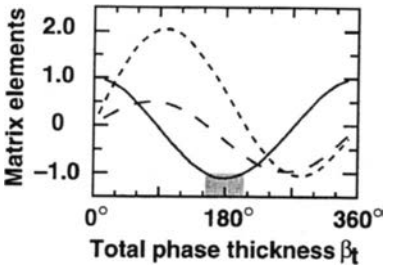


Fig. 2-77 Abscissa: total phase thickness of the group [see Eq. (2-46g)]. Ordinate: elements of the characteristic matrix [see Eq. (2-32g)] m_{11} (*solid*), m_{12} (*long dash*) and m_{21} (*short dash*) of the design $\frac{1}{2}\mathbf{L}\,\mathbf{H}\frac{1}{2}\mathbf{L}$ with $n_L = 1.38$ and $n_H = 2.30$. The stop band is shaded.

$$n_e = \sqrt{\frac{m_{21}}{m_{12}}} \tag{2-46h}$$

[Herpin 47, Kronig 48]. Its phase thickness may be found in several ways, as

$$\beta_e = \arccos(m_{11}) \tag{2-46i}$$

or from

$$\beta_e = \arctan\left(\frac{n_e\, m_{12}}{m_{11}}\right). \tag{2-47a}$$

The latter equation has the advantage that the numerator and denominator of the arctan may be inspected to determine β_e modulo 360°. As an example, Fig. 2-77 shows the matrix elements vs β_t for the design $\frac{1}{2}\mathbf{L}\,\mathbf{H}\frac{1}{2}\mathbf{L}$. Because the matrix is unimodular and according to §5.2.3.1 the magnitude of the trace [see Eq. (2-41e)] is 2.00 at the stop band edge, either m_{12} or m_{21} must be zero at this point. It is also clear that when β_t is 240°, both m_{11} and m_{12} are negative; hence, Eq. (2-47a) indicates that β_e must be in the third quadrant. This method of finding β_e and n_e applies even to inhomogeneous layers, provided the matrix elements are known. Thelen [66] provides an example.

2.6.7.5.2 Three-layer group

§2.11.2 contains equations for β_e and n_e of a three-layer group. Several ratios are used, including the thickness ratio Λ

$$\Lambda \equiv \frac{\text{total thickness of outer layers}}{\text{thickness of inner layer}} = \frac{2\beta_p}{\beta_q} \tag{2-47b}$$

and index ratio Ω

$$\Omega \equiv n_H / n_L\,. \tag{2-47c}$$

2.6.7.5.3 Limit of zero thickness

There are instances in which the equivalent index exists only as a limit. Consider, for example, the group $\frac{1}{2}\mathbf{L}\,\mathbf{H}\frac{1}{2}\mathbf{L}$ when the phase thickness of the **H** layer is 180° and is

therefore absentee (see §2.6.3.1). By the reasoning in §2.6.3.1, the two outer $\frac{1}{2}$**L** layers (each with 90° phase thickness) also coalesce into an absentee layer. Thus n_e exists only as a limit. Berning [62] shows that when $\beta_t \approx 0$

$$n_e \approx n_p \sqrt{\frac{\Lambda + n_q/n_p}{\Lambda + n_p/n_q}} \tag{2-47d}$$

and

$$\beta_e \approx \beta_t \sqrt{1 + \frac{(n_p - n_q)^2 \Lambda}{n_p n_q (1+\Lambda)^2}} \,. \tag{2-47e}$$

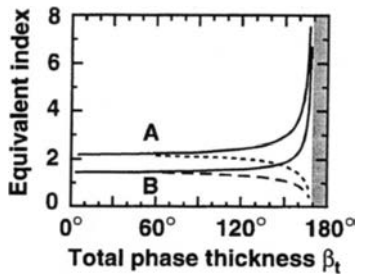


Fig. 2-78 Versus total phase thickness of the group [see Eq. (2-46g)], equivalent index of
(*solid A*) 0.1**L** 1.8**H** 0.1**L**
(*solid B*) 0.9**L** 0.2**H** 0.9**L**
(*short dash*) 0.9**H** 0.2**L** 0.9**H**
and (*long dash*) 0.1**H** 1.8**L** 0.1**H**
where the indices of **H** and **L** are 2.30 and 1.38, respectively. The stop band is shaded.

Consider the limit as $\Lambda << 1$, which means that the group of three layers is mostly composed of the inner layer. In that case the equivalent layer should, in most respects, function like the inner layer. This is illustrated in Fig. 2-78, which plots n_e vs β_t for both $\Lambda = 9$ and 1/9. Curve *A* obtains when the inner layer is *thick* zinc sulfide (n = 2.30). At $\beta_t = 0$ its n_e is close to the zinc sulfide index. The *B* curve occurs when the inner layer is *thin* zinc sulfide—the coating is composed mostly of magnesium fluoride (n = 1.38). Its n_e is close to 1.38 until β_t exceeds 120°.

The *dashed* curves show the behavior of n_e when the inner layers are magnesium fluoride. When the outer layers are thin, n_e is close to 1.38 until it drops abruptly at the edge of the stop band. The stop band (defined in §5.2.2) is narrow because Λ is either greater than or less than 1. Figure 5-7 shows the reflectance of a stack with mismatched layer thicknesses. Its basic period is not an equivalent layer, however.

2.6.7.5.4 Groups of equal optical thickness layers

Consider groups such as **L HL**, **HL H**, **L MHM L** and so on, in which the **L**, **M** and **H** layers have 90° phase thickness at a wavelength λ_0. If the layers are numbered from the outside to the center, the numbers run 121, 12321, 1234321 and so on. If the qth layer is situated at the center and the layers are ordered 1, 2 ... q–1, q, q–1 ... 2, 1, then at λ_0 when q is even

$$n_e = \frac{n_1^2 \; n_3^2 \ldots n_{q-1}^2}{n_2^2 \; n_4^2 \ldots n_{q-2}^2 \, n_q} \tag{2-47f}$$

and for q odd

$$n_e = \frac{n_1^2 \; n_3^2 \ldots n_{q-2}^2 \, n_q}{n_2^2 \; n_4^2 \ldots n_{q-1}^2} \,. \tag{2-47g}$$

Similar to Eqs. (2-48b) and (2-48c), the indices alternate between numerator and denominator. The difference is that the index of the center layer is not squared. The n_e cited above is utilized in Chap. 7 to design multicavity bandpass filters.

2.6.7.5.5 Performance at higher frequencies

In prior sections the equivalent index is analyzed in the frequency region where $0 < \beta_t < 180°$ — at frequencies below that of the first-order stop band. Figure 2-79 plots n_e vs β_t in regions up to the fourth-order stop band— at point *A*. The *dashed* curve resembles a cotangent function between stop bands of orders 1,2 and 2,3. It is parabolic at the (missing) fourth-order stop band. On the other hand, the *solid* curve in that Fig. is parabolic at all of the higher orders.

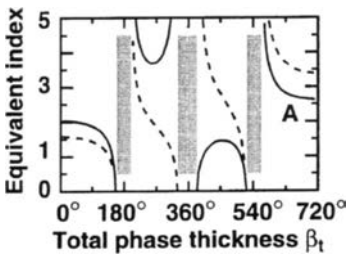


Fig. 2-79 Versus total phase thickness of the group [see Eq. (2-46g)], equivalent index of (*dashed*)
0.25H 1.5L 0.25H
and (*solid*)
0.75H 0.5L 0.75H
where the indices of H and L are 2.30 and 1.38, respectively. The stop bands are shaded.

The total phase thickness of the outer layers is the key to understanding this behavior. The design of the *dashed* curve in Fig. 2-79 is 0.25H 1.5L 0.25H. For the *solid* curve it is 0.75H 0.5L 0.75H. The cotangent shape of n_e predominates until the *total thickness* of the outer layers exceeds 180 °. It then becomes parabolic.

For the design with the *solid* curve, the foregoing condition occurs at the first-order stop band and the n_e curves resemble parabolas. In the case of the coating with the thinner outer layers (*dashed* in Fig. 2-79), the phase thickness of *each* L layer is 90° when the total phase thickness of the period is 720 °. Because 90° + 90° = 180°, the shape of the curve is parabolic.

In general, the parabolic shaped n_e curves are less dispersive and preferred in designs where the bandwidth must be relatively broad.

2.6.7.5.6 Synthesis of an equivalent layer

An appendix (§2.11.3.1) proves that:

> When $\beta_e = 90°$, $n_L < n_e < n_H$, where n_H and n_L refer to the high index and low index components of the equivalent layer.

The foregoing theorem is used frequently. Table 4-7 shows that a layer with index 2.56 is needed, which exceeds the highest possible index of 2.35. Because an equivalent layer with $\beta_e = 90°$ does not exist, the designer has no choice but to use an equivalent layer with a β_e of 270 °.

2.6.7.5.7 Producibility of equivalent layers

One of the problems in the production of equivalent layers is that they often contain films whose optical thicknesses are small compared to a wavelength. Such a layer should be avoided for two reasons: (1) The layer might be so thin that an island film (see §3.3.2.3 and Figs. 3-18, 3-32) forms which does not manifest the refractive index of a thicker layer; and (2) the optical thickness of a thin layer is often difficult to control when it is produced.

A few of the layers in the coating captioned in Fig. 4-45 are as thin as 11 nm in metric thickness. It is questionable whether such a film would manifest the same refractive index as that of a thicker coating.

2.7 Properties of a stack with equal optical thickness layers

2.7.1 Introduction

Equal thickness is intended to mean the following: The optical thickness of each nonabsorbing layer is an integral multiple of $\lambda_0/4$, where λ_0 is a reference wavelength. The stack in Table 2-12 serves as an example. Because the phase thickness of each layer is either 90° or multiple thereof, this is an "equal thickness" stack. There are many manifestations of such multilayers, such as quarterwave stack reflectors, multiple halfwave bandpass filters and so on.

2.7.2 Conventions for specifying thicknesses of layers

2.7.2.1 Definition of QWOT

QWOT is an acronym for *quarterwave optical thickness* and has been used since the early days of multilayer production. At first glance it may seem confusing. Nevertheless, it is widely used and accepted.

The definition of the QWOT is as follows. Suppose that at a particular wavelength λ_0 the optical thickness $n\,h$ is

$$n\,h = \lambda_0/4 \ . \tag{2-47h}$$

Then λ_0 is the QWOT. Solving Eq. (2-47h) for λ_0,

$$\text{QWOT} \equiv 4\,n\,h \tag{2-47i}$$

and thus should be called the *quadrupled optical thickness*. Nevertheless, QWOT is retained for historical reasons. If the n in Eq. (2-47i) is dispersive, the wavelength at which n is defined must also be known in order to find h.

Table 2-12 Metric thicknesses of *matched* layers

1	2	3	4	5	6	7	8	9
Layer	Index	h_i, $\theta_0 = 0$	β_i, at $\theta_0 = 0°$	θ_i	$\cos\theta_i$	β_i, $\theta_0 = 50°$	h_i, $\theta_0 = 50°$	Tkns ratio
1	1.38	344 nm	270°	33.7°	0.832	224.6°	414 nm	1.202
2	2.30	68.8 nm	90°	19.5°	0.943	84.9°	73 nm	1.061
3	1.63	194 nm	180°	28.0°	0.883	158.9°	220 nm	1.133

Metric thicknesses (*column* 3) of nonabsorbing layers whose phase thicknesses β_i are 27°, 90° and 180° at 633 nm, respectively. For an angle of incidence of 50° (in air) and a refractive index appearing in *column* 3, *column* 5 is the angle of refraction and *column* 6 its cosine. The phase thicknesses in *column* 7 occur at 50 °. *Column* 3 divided by *column* 6 equals *column* 8, which is the metric thickness needed to maintain the phase thickness in *column* 4 when the incidence angle is 50°. *Column* 9 is the reciprocal of *column* 6.

2.7.2.2 At a given angle of incidence, *matched thickness* layers

This concept is useful in the design of coatings that function at nonnormal incidence. Two or more nonabsorbing layers are said to be *matched* if, at a particular wavelength λ_0 and angle of incidence θ_0, the phase thicknesses [see Eq. (2-4d)] have a predetermined ratio.

A stack with three *matched thickness* layers serves as an example. The outer film (index 1.38) has a phase thickness of 270°—three quarters of a wave—at a reference wavelength λ_0 of 633 nm. The center layer of index 2.30 has a 90° phase thickness and

the bottom layer (index 1.63) is 180° in phase thickness at the same wavelength. Table 2-12 summarizes these data.

At normal incidence, the metric thickness of the 1.38 layer is obtained by substituting $3\pi/2$ on the left of Eq. (2-4d) and solving for h_1

$$h_1 = 3\,\lambda_0/(4\,n_1) = 3 \times 633/(4 \times 1.38) = 344 \text{ nm} \ . \qquad (2\text{-}47j)$$

Similar calculations are made for the other two layers, as shown in Table 2-12. If the production control is such that those metric thicknesses can be produced, then these layers are *matched* at normal incidence; the ratio of the phase thicknesses is 3:1:2.

What happens if this stack is used at an angle of incidence θ_0 of 50°? The angle of refraction θ_1 in the low index layer is obtained from Eq. (2-4b):

$$\theta_1 = \arcsin[(1.38)^{-1} \sin 50°] = 33.72° \ . \qquad (2\text{-}47k)$$

The "cosine theta" term appearing in Eq. (2-4d) is $\cos\theta_1 = \cos(33.72°) = 0.832$. A similar calculation is made for the other two layers. The results are shown in Table 2-12. Equation (2-4d) is used to find the phase thickness for each layer. Table 2-12 shows that, at $\lambda = \lambda_0$, the phase thicknesses are no longer 270°, 90° and 180°, but rather 225°, 85° and 159°. These data are shown in Fig. 2-80. The result is that the phase thicknesses are not matched at 50° or any other nonzero angle.

There is a simple remedy. The requisite phase thickness at $\theta_0 = 50°$ is obtained for the outer layer if its metric thickness is *increased* by a factor of $\sec\theta_1$. Solving for its metric thickness gives

$$h_1 = 3\,\lambda_0 \sec\theta_1/(4\,n_1) = 3 \times 633 \times \sec(33.72°)/(4 \times 1.38) = 414 \text{ nm} \ . \qquad (2\text{-}47l)$$

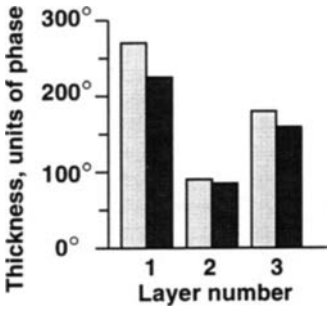


Fig. 2-80 Phase thicknesses of three layers at normal incidence (*left* of each pair, lighter shading) and at 50° incidence in air (*right* of each pair, darker shading). The refractive indices and thicknesses of the layers appear in Table 2-12.

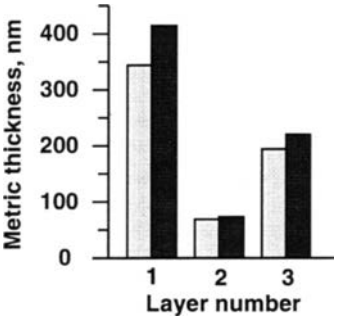


Fig. 2-81 Given the metric thickness of a layer at normal incidence (*left* bar of each pair, shaded *lightly*) metric thicknesses of that layer *matched* at 50° (*right* bar of each pair, *darker* shading) Please see Table 2-12 for details.

A comparison of Eq. (2-47j) and Eq. (2-47l) shows that this amounts to a 20% increase in thickness, as indicated in the last column of Table 2-12. Similar adjustments are made for the other two layers, although the thickness increases are less because they are of higher refractive index than the top layer. When the metric thicknesses shown in *column* 8 of Table 2-12 are used, then phase thicknesses of all layers are matched at 50° incidence angle. Figure 2-81 shows the increase in metric thickness for each layer needed to maintain the *matched* condition at 50°.

The foregoing procedure can be used to adjust the layer thicknesses in any stack so that the *matched* condition obtains. The problem is, of course, that the stack must be redesigned each time the match angle is changed.

2.7.3 Admittance of a stack with layers of equal optical thickness

2.7.3.1 Reduced admittance

The admittance $\hat{Y}'$ and normalized admittance $\hat{Y}$ are introduced in §2.2.5.2. The reduced admittance y appears in §2.3.2. From Eq. (2-16e) the amplitude reflection coefficient is

$$\hat{\rho} = \frac{1-\hat{y}}{1+\hat{y}} \ . \tag{2-48a}$$

Consider the case when $\hat{y}$ is pure real and is replaced by its reciprocal. This adds 180° to the phase shift upon reflection, but does not alter the radiant reflectance.

2.7.3.2 Admittance at the quarterwave point

Consider a stack of nonabsorbing layers in which the β of each layer is 90° or a odd multiple thereof. Equation (2-36c) may be used to show that the reduced admittance $\hat{y}$ is pure real and is

$$\hat{y} = \left(\frac{n_{\ell-1}}{n_\ell} \frac{n_{\ell-3}}{n_{\ell-2}} \cdots \frac{n_3}{n_4} \frac{n_1}{n_2} \right)^2 \frac{n_s}{n_0} + j0 \tag{2-48b}$$

when the layer count ℓ is even and when ℓ is odd:

$$\hat{y} = \left(\frac{n_\ell}{n_{\ell-1}} \frac{n_{\ell-2}}{n_{\ell-3}} \cdots \frac{n_3}{n_2} n_1 \right)^2 \frac{1}{n_s n_0} + j0 \ . \tag{2-48c}$$

Effective indices can be used in the above equations at nonnormal incidence.

2.7.3.3 Standing wave ratio at nonnormal incidence

At nonnormal incidence there is often a need to design a multilayer that is nonpolarizing or possibly nonreflecting. The concept of the standing wave ratio at angle is a useful tool.

Each layer of the stack is matched at angle (see §2.7.2.2) so that a phase thickness of 90° is maintained, regardless of the angle of incidence. Equation (2-48b) factors into the product of two functions [Baumeister 61]

$$U = \left(\frac{n_1}{n_2} \frac{n_3}{n_4} \cdots \frac{n_{\ell-3}}{n_{\ell-2}} \frac{n_{\ell-1}}{n_\ell} \right)^2 \frac{n_s}{n_0} \tag{2-48d}$$

$$W = \left(\frac{\cos\theta_1}{\cos\theta_2} \frac{\cos\theta_3}{\cos\theta_4} \cdots \frac{\cos\theta_{\ell-3}}{\cos\theta_{\ell-2}} \frac{\cos\theta_{\ell-1}}{\cos\theta_\ell} \right)^2 \frac{\cos\theta_s}{\cos\theta_0} \ . \tag{2-48e}$$

The U equation is merely a recast of Eq. (2-48b). The advantage of this formulation is that $\hat{\rho}_s$ and $\hat{\rho}_p$ (at the quarterwave point) simplify. Equation (2-48a) reduces to

$$\hat{\rho}_s = \frac{1-UW}{1+UW} = \frac{1-y_s}{1+y_s} + j0 \tag{2-48f}$$

and

$$\hat{\rho}_p = \frac{1-U/W}{1+U/W} = \frac{1-y_p}{1+y_p} + j0 \tag{2-48g}$$

for the amplitude reflection coefficients for s and p polarized flux, respectively. y_p and y_s, which are pure real, are the respective reduced admittances for p and s. The foregoing formulation provides useful insight. An example is that when either $U = 1$ or $W = 1$,

$$|\hat{\rho}_s| = |\hat{\rho}_p| \qquad (2\text{-}48h)$$

and the stack is nonpolarizing at λ_0. This is further investigated in §6.2.1.1.1.

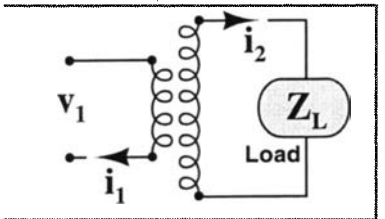


Fig. 2-82 An ac network consists of a transformer with a load impedance Z_L attached to its secondary windings.

2.7.3.4 Quarterwave films viewed as admittance transformers

A concept in circuit theory is applicable to coatings. A stack of equal thickness layers can be viewed as an *admittance transformer*. Figure 2-82 shows a lossless transformer with negligible leakage inductance with a load impedance Z_L connected to its secondary windings. The impedance Z_{input} at the input of the transformer is

$$Z_{input} = \eta^{-2} Z_L \qquad (2\text{-}49a)$$

where η is the turns ratio. It transforms the reciprocal of the Z (admittance Y) as

$$Y_{input} = \eta^2 Y_L \ . \qquad (2\text{-}49b)$$

As an illustration of this concept, the normalized admittance of a glass substrate is

$$\hat{Y} = n_s = 1.52 + j0 \qquad (2\text{-}49c)$$

and is viewed as a pure real load admittance. Suppose that two nonabsorbing quarterwave thickness films, represented by **H L**, are deposited on this substrate. The **H** is next to air. Equation (2-36c) provides the normalized admittance at λ_0

$$Y_{input} = \frac{n_H^2 n_s}{n_L^2} = \frac{(1.71)^2\, 1.52}{(1.38)^2} = 2.33 \ . \qquad (2\text{-}49d)$$

This transforms the load admittance upwards. There is clearly an admittance mismatch if air (normalized admittance = 1.00) is the incident medium. The substitution into Eqs. (2-16f) and (2-12c) shows that a radiant reflectance of 16% (at λ_0) results from this mismatch. In §4.2.3.1 the pair **L H** is used as a stepdown "transformer" to form an antireflection coating.

2.7.4 *R/T* polynomial

2.7.4.1 Introduction

The *R/T* polynomial for a nonabsorbing multilayer with equal optical thickness layers was developed separately in two disciplines: optics and microwave technology. Efforts have been made to meld them, since the techniques developed in one can apply directly to the other. §2.2.5.3 presents analogies that relate the two disciplines.

2.7.4.2 Development of the *R/T* polynomial

This section contains a brief overview. Computational details appear in comprehensive review articles [Delano 69; Pegis 61]. Consider a stack in which each layer has the same phase thickness β. Given the sines and cosines of that phase

$$S \equiv \cos\beta \qquad \text{and} \qquad C \equiv \cos\beta \ . \qquad (2\text{-}50a)$$

Pohlack [52] showed that for ℓ layers

$$R/T \equiv F = a_0 + a_2 S^2 + a_4 S^4 + \ldots + a_{2\ell} S^{2\ell} \quad (2\text{-}50b)$$

where each of the a coefficients is a function of the refractive indices. Trigonometric identities transform Eq. (2-50b) into alternate forms

$$F = b_0 + b_2 C^2 + b_4 C^4 + \ldots + b_{2\ell} C^{2\ell} \quad (2\text{-}50c)$$

$$F = c_0 + c_2 \cos(2\beta) + c_4 \cos(4\beta) + \ldots + c_{2\ell} \cos(2\ell\beta) \ . \quad (2\text{-}50d)$$

Delano [69]) shows that the R/T curve can be expanded into a Fourier series and the coefficients in Eq. (2-50d) thereby specified. Details are also given on methods of finding the indices, given the c coefficients in Eq. (2-50d). A few features of this method are:

(1) There are multiple solutions—2^ℓ of them—and the computational labor increases accordingly. With 6 layers, for example, $2^6 = 64$ sets of indices exist. There are instances in which the indices represent complex roots of a polynomial and are, therefore, unrealizable.

(2) There are no constraints on the indices. The solution may require an index of 1.105, which is unattainable if the incident medium is air. The procedure of Kard described in the next section allows the designer to linearize the equations and thereby impose boundary constraints on the indices.

2.7.4.3 Kard formulation of the R/T polynomial

The Kard matrix (see §2.4.3.4) is used to generate the R/T polynomial. For a single layer this is

$$R/T \equiv F = [\sinh(u_0)]^2 \ C^2 + [\sinh(u_0 - 2u_1)]^2 S^2 \ . \quad (2\text{-}51a)$$

The R/T polynomials for two layers and three layers are, respectively,

$$F = (a_0 C^2 - a_2 S^2)^2 + (a_1 CS)^2 \quad (2\text{-}51b)$$

and

$$F = (a_0 C^3 - a_2 C S^2)^2 + (a_1 C^2 S - a_3 S^3)^2 \quad (2\text{-}51c)$$

Table 2-13 Coefficients for one, two and three layer stacks

Coefficient	Number of layers 2	3
a_0	$\sinh(u_0)$	$\sinh(u_0)$
a_1	$\sinh(u_0 - 2u_1) + \sinh(u_0 - 2u_2)$	$\sinh(u_0 - 2u_1) + \sinh(u_0 - 2u_2) + \sinh(u_0 - 2u_3)$
a_2	$\sinh(u_0 - 2u_1 + 2u_2)$	$\sinh(u_0 - 2u_1 + 2u_2) + \sinh(u_0 - 2u_1 + 2u_3) + \sinh(u_0 - 2u_2 + 2u_3)$
a_3		$\sinh(u_0 - 2u_1 + 2u_2 - 2u_3)$

Coefficients of the R/T polynomial appearing in Eqs. (2-51b) and (2-51c) for two layers and three layers, respectively. The phase thicknesses of the layers are equal.

where the coefficients appear in Table 2-13. The general form is

$$F = B_1^2 + B_2^2 \quad (2\text{-}51d)$$

where

$$B_1 \equiv \sum_{i=0}^{\ell} a_i \; j^i \; S^i C^{\ell-i} \tag{2-51e}$$

for i even and

$$B_2 \equiv \sum_{i=1}^{\ell} a_i \, j^{i-1} \, S^i \; C^{\ell-i} \tag{2-51f}$$

for i odd, with $j \equiv \sqrt{-1}$.

An algorithm generates the coefficients a_i for a stack of ℓ layers, as illustrated in Table 2-14. The example assumes that the coefficients for the one, two and three layer stacks have already been copied into Table 2-14 from Table 2-13. The "sinh" is omitted and only the arguments of the hyperbolic functions appear in Table 2-14.

Given the coefficients of a three-layer stack, the algorithm produces the coefficients for the four-layer coating in two steps:

(1) Copy *each term* in the coefficients of the *three-layer* stack into the column for the *four-layer* stack. The result of this copying operation is shown in sans serif type in Table 2-14. A double subscript of $a_{i\ell}$ denotes the ith coefficient in a stack with ℓ layers. The copying operation means that for $i = 0$ to ℓ

$$a_{i,\ell} = a_{i,\ell-1} \; . \tag{2-51g}$$

(2) Copy *each term* of the $(i-1)^{th}$ coefficients of the three-layer stack to the ith box and add "$2\,(-1)^i\, u_4$" to that term. This $(i-1)^{th}$ term is in a diagonal position of Table 2-14, where boldface shows the u_4 terms added by this operation. For $i = 1$ to ℓ, this is

$$a_{i,\ell} = a_{i-1,\ell-1} + 2\,(-1)^i\, u_\ell \; . \tag{2-51h}$$

It order to verify that the operations have been executed properly, steps (1) and (2) above should generate

$$\frac{\ell!}{i!\,(\ell-i)!} \tag{2-51i}$$

terms for each $a_{i,\ell}$.

Kard shows that the hyperbolic functions in Table 2-13 can be approximated by their arguments when they are much less than one. Consider an expansion for a_0

$$a_0 = \sinh(u_0) \approx u_0 + \tfrac{1}{6} u_0^3 + \ldots \; . \tag{2-52a}$$

The magnitude of u_0 is the greatest when the substrate index is large, as, for example, germanium with an index of 4.00 in the IR. Then from Eq. (2-29c)

$$u_0 = 0.5 \ln(1/4) = -0.69 \; . \tag{2-52b}$$

The cubic term in Eq. (2-52a) is -0.056; an 8% error is incurred by neglecting it and approximating the hyperbolic function by its argument. If this is done, the coefficients a_m simplify considerably and are written

$$a_m \approx \sum_{i=0}^{\ell} b_{mi} \; u_i \quad . \tag{2-52c}$$

The coefficients b_{mi} are listed in Table 2-15. Applications of these equations are developed in Chap. 4.

Table 2-14 Coefficients for a four-layer stack

Layer count 0	1	2	3	4	
u_0	u_0	u_0	u_0	u_0	a_0
	$u_0 - 2\,u_1$	$u_0 - 2\,u_1$ $u_0 - 2\,u_2$	$u_0 - 2\,u_1$ $u_0 - 2\,u_2$ $u_0 - 2\,u_3$	$u_0 - 2\,u_1$ $u_0 - 2\,u_2$ $u_0 - 2\,u_3$ $u_0 - \mathbf{2\,u_4}$	a_1
		$u_0 - 2\,u_1 + 2\,u_2$	$u_0 - 2\,u_1 + 2\,u_2$ $u_0 - 2\,u_1 + 2\,u_3$ $u_0 - 2\,u_2 + 2\,u_3$	$u_0 - 2\,u_1 + 2\,u_2$ $u_0 - 2\,u_1 + 2\,u_3$ $u_0 - 2\,u_2 + 2\,u_3$ $u_0 - 2\,u_1 + \mathbf{2\,u_4}$ $u_0 - 2\,u_2 + \mathbf{2\,u_4}$ $u_0 - 2\,u_3 + \mathbf{2\,u_4}$	a_2
			$u_0 - 2\,u_1 + 2\,u_2 - 2\,u_3$	$u_0 - 2\,u_1 + 2\,u_2 - 2\,u_3$ $u_0 - 2\,u_1 + 2\,u_2 - \mathbf{2\,u_4}$ $u_0 - 2\,u_1 + 2\,u_3 - \mathbf{2\,u_4}$ $u_0 - 2\,u_2 + 2\,u_3 - \mathbf{2\,u_4}$	a_3
				$u_0 - 2\,u_1 + 2\,u_2 - 2\,u_3 + \mathbf{2\,u_4}$	a_4

Implementation of an algorithm that generates the coefficients a_0, a_1, a_2 and a_4 [see Eqs. (2-51d), (2-51e) and (2-51f)] of the Kard R/T polynomial for four layers of equal phase thickness. The significance of the **boldface** and sans serif type is explained in §2.7.4.3.

2.7.4.4 Effects of symmetry upon the R/T polynomial

Symmetry refers to a stack in which there are pairs of Fresnel coefficients equal in magnitude. This usually causes terms to drop out of the R/T polynomial. A stack of four nonabsorbing layers serves as an example. There are five Fresnel reflection coefficients at the interfaces. Suppose that two of them are constrained by symmetry

$$r_0 = r_4 \quad \text{and} \quad r_1 = r_3 \ . \tag{2-52d}$$

This, in turn, constrains the refractive indices

$$n_1\, n_4 = n_0\, n_s \quad \text{and} \quad n_2\, n_3 = n_0\, n_s \tag{2-52e}$$

or, in terms of the Kard u coefficients [see Eq. (2-29c)],

$$u_1 + u_4 = u_2 + u_3 = u_0 \ . \tag{2-52f}$$

Examination of Table 2-15 shows that the foregoing constraints cause a_1 and a_3 to be zero. The result is that B_2 in Eq. (2-51d) is zero. This simplifies the design of antireflection coatings and is discussed in Chap. 4.

In general, if the following conditions are satisfied for a stack with layer count ℓ even

$$n_0\, n_s = n_1\, n_\ell = n_2\, n_{\ell-1} = n_3\, n_{\ell-2} \tag{2-52g}$$

and so on, the B_2 term vanishes. When the layer count is odd, the foregoing equation must be satisfied and, in addition, the index of the center layer is

$$n_{\ell/2} = \sqrt{n_0\, n_s} \tag{2-52h}$$

which translates to

Table 2-15
Kard *R/T* polynomial — approximate values of coefficients

1	1	1	–2				
2	1	2	–2	–2			
2	2	1	–2	2			
3	1	3	–2	–2	–2		
3	2	3	–4	0	4		
3	3	1	–2	2	–2		
4	1	4	–2	–2	–2	–2	
4	2	6	–6	–2	2	6	
4	3	4	–6	2	2	–6	
4	4	1	–2	2	–2	2	
5	1	5	–2	–2	–2	–2	–2
5	2	10	–8	–4	0	4	8
5	3	10	–12	0	4	0	–12
5	4	5	–8	4	0	–4	8
5	5	1	–2	2	–2	2	–2
ℓ	m	0	1	2	3	4	5
		Coefficient number, i					

Coefficients b_{mi} appearing in Eq. (2-52c), where ℓ is the number of layers.

$$u_{\ell/2} = 0.5\, u_0 \ . \tag{2-53a}$$

A useful equation is the inversion of Eq. (2-29c)

$$n_i = n_s \exp(2\, u_i) \ . \tag{2-53b}$$

2.8 Graphical aids to multilayer design

"*The purpose of computing is insight, not numbers.*" Hamming [62].

2.8.1 Introduction

Two uses of graphics are considered: (1) graphical *computation* and (2) graphical *presentation.* Two scenarios help to distinguish between these uses. The task is to coat a metallic substrate with three nonabsorbing layers that reduce its reflectance to zero at λ_0.

2.8.1.1 Graphics as a computational aid

In the late 1940s a thin-film designer would often utilize a Smith chart to find layer thicknesses. The first step is to plot the reduced admittance of the metallic substrate on the chart. The optical thickness of the bottom layer is read from angular scale—the Smith chart is essentially an analog computer. The next interface is encountered and a low index layer is added to the stack. The reduced admittance is read from the chart and, after a multiplication on the slide rule, the reduced admittance at the center layer is plotted. A circle is drawn with the compass through that admittance point. The top layer (of high index) is added and a circle is drawn so that zero reflectance is attained after

crossing its outer interface into air. The thicknesses of the center and outer layer are read from the angular scale on the periphery of the chart.

The designer uses the graphical computational method for design—to find the layer thicknesses. The Smith chart could also find a numerical value for the amplitude reflection coefficient of the stack.

2.8.2 Graphical presentation

The scene is the same desk in the 1990s, which now is graced by either a personal computer or a terminal to a larger number-crunching monster. The equations in §2.6.5.1 have been programmed on that computer. Execution of a computer code provides a numerical value for the thickness of the bottom layer of the stack. The procedure in §4.3.4.2.2 has been programmed on a digital computer. This supplies the thicknesses of the center and outer layers. The design of the entire stack is copied into the computer, which presents to the display screen a graph of either the admittance or amplitude reflection coefficient as the thicknesses of the layers in the stack progress from zero (i.e., the naked substrate) to the final three-layer design. The designer knows that the design is successful if zero reflectance or, alternatively, a normalized admittance of 1.00 is attained.

Methods of graphical presentation are discussed in §10.1.

2.8.3 Vector addition of amplitudes

2.8.3.1 Introduction

This is an approximate method in which it is assumed that the amplitude reflection coefficient $\hat{\rho}$ of a stack is the sum of the components of the electric field reflected at each interface. Consider, for example, the vector approximation for a single layer

$$\hat{\rho}_0 \approx \hat{r}_0 + \hat{r}_1 \exp(-j\,2\beta_1) \ . \tag{2-54a}$$

A comparison of Eq. (2-54a) with the exact relationship [see Eq. (2-23a)] shows that the denominator of the latter has been neglected. Texts such as Stone [63] and Born & Wolf [75] show that the denominator in Eq. (2-23a) is due to multiple reflections within the layer. This is illustrated in Fig. 2-83, which shows the incident ray reflecting at the top interface *A*. The amplitude of this reflection is $\hat{r}_0$. The incident ray also penetrates the film, reflects at the bottom interface *B* and emerges at *b* The amplitude of this reflection is $\hat{r}_1$ and it undergoes a phase retardation of "$-2\beta_1$." A multiply reflected ray is *dashed* and emerges at *c*. The vector addition method neglects this and all other multiply reflected rays.

2.8.3.2 Methods of laying out the vectors

The vector diagram is a graphical solution of Eq. (2-54b) or similar equations. Each term is a complex number represented as a vector on an Argand diagram. Any discussion of this may seem trivial, but, trivial or not, there are rules that make its implementation easier. As an example, the amplitude reflection coefficient of a stack of three nonabsorbing layers is

$$\hat{\rho}_3 \approx r_0 + r_1 \exp[-j\,2\beta_1\,] + r_2 \exp[-j\,2(\beta_1+\beta_2)] + r_3 \exp[-j\,2(\beta_1+\beta_2+\beta_3)] \tag{2-54b}$$

and is used in the examples below. As shown in later sections, even though $|\hat{\rho}|$ may even exceed 1.00, the usage of Eq. (2-54b) still provides insight.

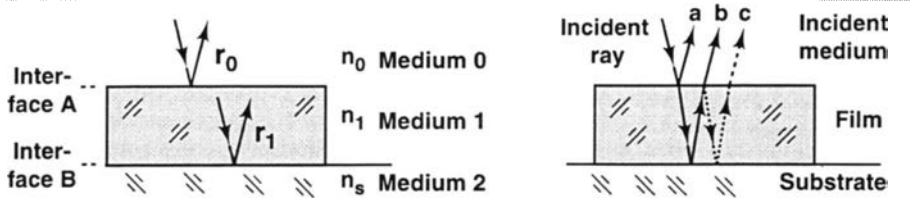


Fig. 2-83 (*Left*) Fresnel amplitude reflection coefficients r_0 and r_1 at the upper A and lower B interfaces of a single layer. (*Right*) A wave is incident upon the layer. Division of wave front causes the waves a and b to reflect from interfaces A and B, respectively. The *dashed* line represents the multiply reflected wave c.

2.8.3.2.1 Method one: Distributive multiplication of phase factors

The phase retardance of each vector is measured *relative to the previous one.* As an illustration, given the vector that represents "$r_1 \exp(-j\,2\,\beta_1)$," the direction of "$r_2 \exp(-j\,2\,\beta_2)$" is obtained by rotating clockwise by an angle "$-2\,\beta_2$" using the vector "$r_1 \exp(-j\,2\,\beta_1)$" *as a reference.* The advantage of this method is that no additional computations are necessary and the layout is done rapidly. The disadvantage is that if an error is made in the direction of any one vector, this error then propagates.

The following rules facilitate the use of this method:

(a) Compute the Fresnel reflection coefficient $\hat{r}_i$ at each interface and convert it to polar form

$$\hat{r}_i \equiv |r_i| \exp(j\,\delta_i) \tag{2-54c}$$

where $|r_i|$ is the magnitude of the vector and δ_i is its phase.

(b) Plot $|r_0|$ along the positive real axis. In so doing, ignore the phase δ_0 of that vector.

(c) Lay out $|r_1|$ at an angle (clockwise) of "$-2\,\beta_1$" relative to $|r_0|$. In so doing, ignore the phase δ_1 of the vector r_1.

(d) Apply step (c) to the next vector and continue in this way until all vectors have been plotted.

(e) Then, and only then, adjust the direction of the ith vector to account for its phase δ_i [see Eq. (2-54d)].

(f) Add the vectors. The resultant vector $\hat{\rho}$ is drawn from the origin to the tip of the final vector.

In mathematical terms, step (c) above is equivalent to applying the distributive rule of multiplication in Eq. (2-54b). As an illustration, step (c) is written for a two-layer stack

$$\hat{\rho} \approx |r_0| + [\exp(-j\,2\,\beta_1)]\,[|r_1| + |r_2| \exp(-j\,2\,\beta_2)]\ . \tag{2-54d}$$

In step (d) the phase factors of the Fresnel amplitude reflection coefficients are included

$$\hat{\rho} \approx |r_0| \exp(j\,\delta_0) + [\exp(-j\,2\,\beta_1)]\,[|r_1| \exp(j\,\delta_1) + |r_2| \exp(-j\,2\,\beta_2 + j\,\delta_2)] \tag{2-54e}$$

and thus the vectors are in their final form. If adjacent media are nonabsorbing, the phase factor of each vector is merely ± 1, since $\exp(j\,0) \equiv 1$ and $\exp(j\,\pi) \equiv -1$.

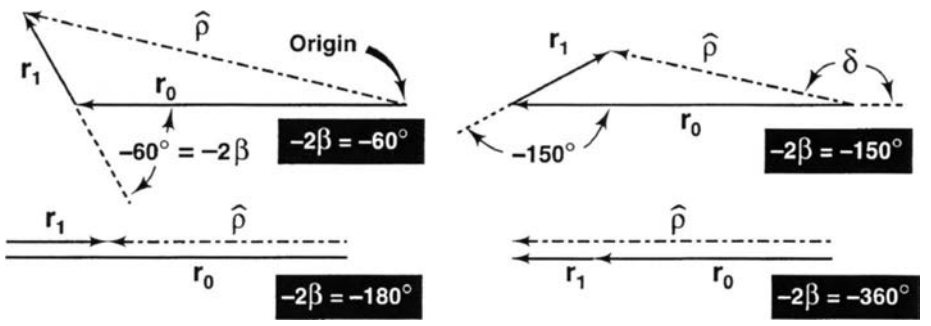


Fig. 2-84 In the addition of waves *a* and *b* (see Fig. 2-83), their amplitudes are represented as vectors in the Argand plane. β is the phase thickness of the layer of index 1.38 deposited on a substrate of index 1.52. $\hat{\rho}$ (*long-short dash*) is the net amplitude reflection coefficient and δ is the phase shift upon reflection. The incident medium is air. At $-2\beta = -60°$, the *short-dash* line represents the magnitude of $|r_1|$.

A final comment relates to the reference for zero phase. It has been tacitly assumed that it is at the top surface of the multilayer. There are several other possibilities, however. These are: (a) the substrate is used as the reference surface; or (b) an arbitrary layer in the stack is designated as "the reference point for zero phase."

2.8.3.2.2 Method two: Summation of the phase thicknesses

In this method, the user computes the Fresnel amplitude reflection coefficient in polar form as in step (a) of §2.8.3.2.1 and then adds the phase retardances and the phase of the Fresnel coefficient to find the total phase ϕ of each layer. An example is that the total phase ϕ_3 for layer three is

$$\phi_3 = -2\beta_1 - 2\beta_2 - 2\beta_3 + \delta_3 \ . \qquad (2\text{-}54f)$$

After this is done, the vectors are plotted using the positive real axis as the reference. $+\phi$ is plotted counterclockwise and $-\phi$, clockwise.

2.8.3.3 Examples of vector addition — antireflection

2.8.3.3.1 Single layer

§2.2.7 illustrates the optical interference between the waves reflecting from the two surfaces of a single layer (of index 1.38) deposited upon a glass (of index 1.65) surface. Each of those waves is now represented by a vector.

A lower index BK-7 substrate ($n_s = 1.52$) is used in this example of vector addition. The phase retardances are 30°, 75°, 90° and 180°, which correspond to layer thicknesses of 1/12, 5/24, 1/4 and 1/2 waves, respectively. The first step is to lay out the vectors without any regard for the sign of the Fresnel coefficient, as outlined in steps (b), (c) and (d) in §2.8.3.2.1. Figure 2-84 shows the vector "$|r_1|\exp(-j\,60°)$" as a *dashed* line. $|r_0|$ lies along the positive x axis. The next step is to account for the signs mentioned in the foregoing sentence. Because $r_0 = -0.16$ and $r_1 = -0.048$, both Fresnel coefficients are negative and, hence, the vectors must be inverted, as shown by the *solid* vectors in Fig. 2-84. $\hat{\rho}$ is represented by the *long-dash* vector. The phase shift upon reflection δ is also shown. The vectors are collinear when the layer thickness is a quarterwave ($\beta_1 = 90°$) or a halfwave ($\beta_1 = 180°$).

2.8.3.3.2 Triple-layer antireflection coating

A vector diagram is constructed for the design

air **L H Q** *glass*

where *air*, **L**, **H**, **Q** and *glass* have indices of 1.0, 1.38, 2.37, 1.70 and 1.52, respectively. The optical thickness of **L** and **Q** is $\lambda_0/4$. The optical thickness of **H** is double that of the other layers. Its $\hat{\rho}$ at $\lambda = 0.878\,\lambda_0$ is found. Table 2-16 shows that the phase retardance "-2β" of layers 1 and 3 is –205°; the "-2β" of layer 2 is –50° (i.e., –410° modulo 360°).

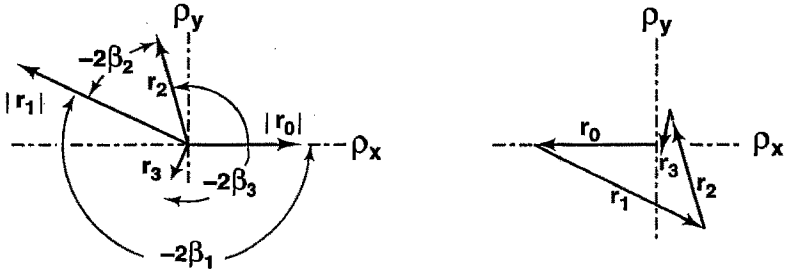


Fig. 2-85 Plotted in the amplitude reflection plane are the vectors representing the waves reflected from each of the interfaces of the coating described in Table 2-16 — at a wavelength $0.878\,\lambda_0$. (*Left*) The vectors represent $|r_0|$, $|r_1|\exp(-j2\beta_1)$, $|r_2|\exp(-j2\beta_1 - j2\beta_2)$, and so on. (*Right*) The vectors represent r_0, $r_1\exp(-j2\beta_1)$, $r_2\exp(-j2\beta_1 - j2\beta_2)$ and so on. Angles are modulo 360°.

Table 2-16 Vector diagram analysis of a three-layer antireflection coating

Medium number	Designation	Refractive index	Optical thickness	Phase thickness -2β	Fresnel coefficient	Total phase*
0	*air*	1.00	massive			
					−0.16	−180°
1	**L**	1.38	$\lambda_0/4$	−205°		
					−0.264	−25°
2	**H**	2.37	$\lambda_0/2$	−410°		
					+0.165	−255°
3	**Q**	1.70	$\lambda_0/4$	−205°		
					+0.056	−100°
4	*substrate*	1.52	massive			

Total phase (modulo 360°) — the sum of the phase thickness of the layer and phase shift upon reflection — [see Eq. (2-54f)] of three nonabsorbing layers at a wavelength $\lambda = 0.875\,\lambda_0$. The Fresnel amplitude reflection coefficients at the interfaces between the media are computed using Eq. (2-13f).

The next task is to follow steps (b), (c) and (d) in §2.8.3.2.1. Figure 2-85 (*left* panel) illustrates how the vectors are laid out according to this prescription. $|r|$ is plotted. The phase angle of three of the vectors is relative to the previous one.

Step (e) in §2.8.3.2.1 is followed next. Table 2-16 shows that the Fresnel coefficients r_0 and r_1 are negative. When these are inverted, the vector diagram is then

constructed, as displayed in the right panel of Fig. 2-85. The tip of the resultant vector is close to the origin. This is why §4.5.2 describes this design as *quasi-maximally flat.*

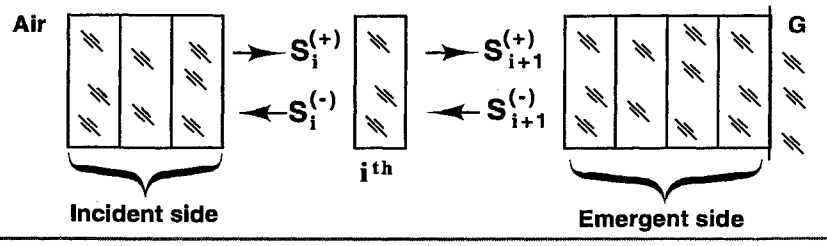


Fig. 2-86 Incident component of Poynting vector $S_i^{(+)}$ and reflected component $S_i^{(-)}$ on the incident side of the ith layer. The subscript "i+1" refers to the components on its emergent side. G is the substrate.

Table 2-16 illustrates the second method of laying out the vectors, as outlined in §2.8.3.2.2. The total phase in Eq. (2-54f) is computed for each layer. These angles are identical with those in the graphical layout in Fig. 2-85—adhering to the convention that negative angles are drawn clockwise.

Further examples of vector diagrams appear in §4.3.4.1, §4.5.2.2, §4.6.3.2, §4.9.2.2, §5.1.2.1, §5.1.3.1, §5.1.3.2, §5.1.4.1, §5.3.6.1 and §5.6.5.1.2.

2.9 Standing waves, net flux ratio and absorption

2.9.1 Introduction

This section deals with the absorption and transmittance of absorbing stacks. Suppose, for example, the multilayer in Fig. 2-66 is irradiated with an intense laser beam and the designer desires to compute the joule heating—the absorption—in layer c. As discussed below, several exact methods may be used; the *standing waves method* is approximate, but has the advantage that it offers insight. Regardless of which method is used, the absorption in any given layer depends upon the thickness, n and k of the layer itself, but also upon the optical properties of the stacks on both its incident side and emergent side.

2.9.2 Net flux ratio—definition and properties

The use of this parameter offers an exact method of finding the absorption in a stack or any layer. For reasons enumerated below, the *net flux* ratio Ψ is often termed *potential transmittance.* Figure 2-86 depicts the ith layer in a stack; its net flux ratio Ψ_i is

$$\Psi_{i,i+1} = \Psi_i \equiv \frac{S_{i+1}^{(+)} - S_{i+1}^{(-)}}{S_i^{(+)} - S_i^{(-)}} \tag{2-55a}$$

where $S_i^{(+)}$ and $S_i^{(-)}$ are the projections of the Poynting vector $\vec{S}$ on the positive z axis in the $+z$ and $-z$ directions, respectively. The subscript "i+1" identifies the Poynting vector at the emergent interface of the layer. The numerator of Eq. (2-55a) is the *net flux* incident and the denominator is the *net flux* emergent—hence the name *net flux ratio.*

The net flux ratio may also be established for a group of layers

$$\Psi_{ij} = \frac{S_{j+1}^{(+)} - S_{j+1}^{(-)}}{S_i^{(+)} - S_i^{(-)}} \tag{2-55b}$$

where $j \geq i$. $S^{(+)}$ and $S^{(-)}$ have the same meaning as in Eq. (2-55a); i refers to the incident side of the ith layer and j to the incident side of the jth layer (or substrate), as shown in Fig. 2-87.

The general properties of Ψ are recapitulated below and are illustrated with reference to the four-layer stack shown in Fig. 2-88. By definition, there is no reflected component $S_4^{(-)}$ of the Poynting vector in the emergent medium, which lies to the right of layer 4.

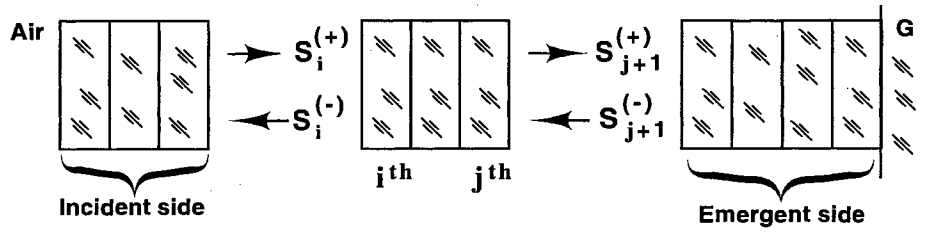


Fig. 2-87 The caption to Fig. 2-86 obtains, with the exception that the Poynting vector components are for a group of films between the ith layer and j^{th} layer.

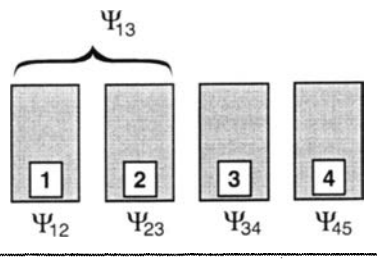


Fig. 2-88 Ψ_{13} is computed for the group of layers 1 and 2. Ψ_{34} and Ψ_{45} refer to the individual layers 3 and 4, respectively.

2.9.2.1 The net flux ratio is multiplicative

This property follows from its definition. Examples of this multiplicative property are (see Fig. 2-88)

$$\Psi_{13} = \Psi_{12}\,\Psi_{23} \qquad (2\text{-}55c)$$

or

$$\Psi_{15} = \Psi_{13}\,\Psi_{34}\,\Psi_{45} \qquad (2\text{-}55d)$$

where Ψ_{13} and Ψ_{15} represent the net flux ratios of groups of layers. A consequence is that if Ψ is known for a single layer or group of layers, it can be combined multiplicatively with other groups.

2.9.2.2 The net flux ratio is unity for an nonabsorbing stack

Conservation of energy requires that

$$0 \leq \Psi \leq 1 \ . \qquad (2\text{-}55e)$$

By definition $\Psi = 1.0$ if any layer or group of layers is nonabsorbing. *Ab intuito*, if there is no absorption, the net flux arriving must equal the net flux leaving any layer or group of layers.

2.9.2.3 The net flux ratio and potential transmittance

In a medium in which R and T can be defined, it follows from its definition

$$\Psi = \frac{T}{1-R} = \frac{1}{1+\frac{A}{T}} \ . \qquad (2\text{-}55f)$$

Ψ is also called the *potential transmittance.* It follows from Eq. (2-55f) that in any multilayer with zero reflectance, $T = \Psi$.

2.9.2.4 The net flux ratio is independent of incident medium

The Ψ of any layer does *not* depend upon the properties of what lies on its incident side. An example is that Ψ is independent of the incident medium index. For a single layer or stack of layers, Ψ depends upon *only* the optical properties of the multilayer and/or substrate on its *emergent* side. This is proven in §2.9.3.

As an example, suppose that Ψ of layer c in Fig. 2-66 is calculated. Ψ depends upon only the thickness, n and k of that layer, as well as the thicknesses and optical constants of layers d, e, f and g, including the index of the substrate. Layers a and b in that stack do not affect the Ψ of layer c. The refractive index of the incident medium does not affect the Ψ of that layer or any layer.

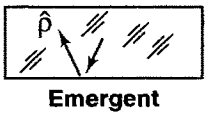


Fig. 2-89 An absorbing film is contiguous to a stack on its emergent side. The internal amplitude reflection coefficient $\hat{\rho}$ at the film-stack interface is specified.

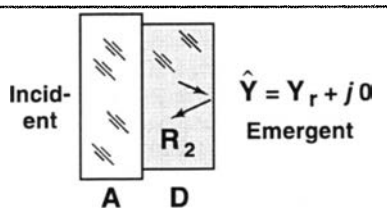


Fig. 2-90 An absorbing film is contiguous to a dielectric layer D interposed between it and a multilayer stack. R_2 is the radiant reflectance at the film-stack interface and $\hat{Y}$ is the normalized admittance of the stack on the emergent side of D.

2.9.3 Methods of computing the net flux ratio

2.9.3.1 Introduction

Figures 2-89, 2-90 and 2-91 depict an absorbing layer bounded on its emergent side by either a multilayer or a substrate. There are three methods of computing the Ψ of that layer, depending upon the convention used to express the properties of the substrate or stack contiguous to its emergent side. These are:

2.9.3.1 (a) The internal amplitude reflection coefficient $\hat{\rho}$ at the interface between the absorbing layer and the stack that abuts its emergent side is computed, as depicted in Fig. 2-89.

2.9.3.1 (b) A single nonabsorbing layer D (see Fig. 2-90) is inserted between the absorbing layer and the stack. The radiant reflectance R_2 at the interface is specified as well as the phase shift.

2.9.3.1 (c) The normalized admittance of the stack on the emergent side is found, as shown in Fig. 2-91.

In the last case, the Ψ of a group of layers, as well as a single layer, can be calculated. For a single layer, the methods in §2.9.3.1 (a) and §2.9.3.1 (c) are equivalent, since Eq. (2-16e) converts normalized admittance into an internal amplitude reflection coefficient.

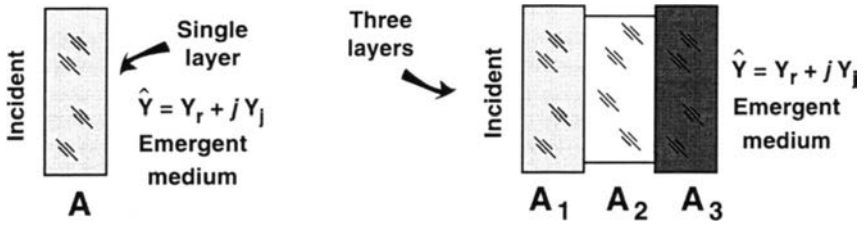


Fig. 2-91 The net flux ratio is computed for a single layer (*left* panel) and a three-layer absorbing stack (*right panel*). Each is contiguous to a stack on its emergent side, the normalized admittance $\hat{Y}$ of which is specified.

2.9.3.2 The net flux ratio in terms of the internal amplitude reflection coefficient

The procedure of Berning [57] is followed. Given the wavelength as well as the n, k and metric thickness h of an absorbing layer, the real part of the propagation constant β_r is found from Eq. (2-4d) and the imaginary part from Eq. (2-5c). Ψ reduces to

$$\Psi = \frac{1 - |\hat{\rho}|^2 - 2\,|k|\;n^{-1}\;|\hat{\rho}|\sin\delta}{\exp(-2\beta_j) - |\hat{\rho}|^2\exp(2\beta_j) - 2\,|k|\,n^{-1}\,|\hat{\rho}|\sin(\delta - 2\beta_r)} \tag{2-56a}$$

where $|\hat{\rho}|$ is the magnitude of the reflection coefficient and δ is the phase shift upon reflection at the interface between the absorbing film and whatever is contiguous to its emergent side. There is an alternative way of deriving Eq. (2-56a) [Holloway 69]. Substitution of Eqs. (2-31d) and Eq. (2-32f) into Eq. (2-58a) yields Eq. (2-56a).

2.9.3.3 Kard's formulation of the net flux ratio

Although this formulation seems at first glance to be unnecessarily complex, it is included because it offers useful insight. As depicted in Fig. 2-90, the absorbing film is bounded on its emergent side by a nonabsorbing layer of index n_m and phase thickness β_m. R_2 is the radiant reflectance at the interface between the nonabsorbing film and the medium (or multilayer) on its emergent side. The admittance of the latter is assumed to be pure real and, furthermore, that the phase shift upon reflection at the film-medium interface is zero. Although Kard does not impose this restriction, it is convenient to do so here to simplify the formulas. The following constants are established

$$B \equiv \tfrac{1}{2}\,[-n\,\sinh(2\beta_j) + |k|\,\sin(2\beta_r)]\;, \tag{2-56b}$$

$$C \equiv \tfrac{-1}{2}\,[n\,\sinh(2\beta_j) + |k|\,\sin(2\beta_r)]\;(n^2 + k^2)^{-1}\;, \tag{2-56c}$$

$$\kappa \equiv\; [\cosh(2\beta_j) + k^2\,n^{-2}\cos(2\beta_r)]\;(1 + k^2\,n^{-2})^{-1}\;, \tag{2-56d}$$

$$Q_1 \equiv B\,n_m^{-1} + C\,n_m\;, \tag{2-56e}$$

$$Q_2 \equiv B\,n_m^{-1} - C\,n_m \tag{2-56f}$$

and finally

$$\mu \equiv \arccos(Q_2 / U_0) \tag{2-56g}$$

where

$$U_0 = \left\{\, n^2\,k^2\,[\cosh(2\beta_j) - \cos(2\beta_r)]^2\,[n^2 + k^2]^{-2} + Q_2^2 \,\right\}^{1/2}. \tag{2-56h}$$

The angle in Eq. (2-56g) is restricted to quadrant I. The variables β_r, β_j, n and k have the same meaning as in §2.9.3.2. The Ψ is

$$\Psi = \left[\kappa + Q_1 \frac{1+R_2}{1-R_2} + \frac{2\,U_0 \sqrt{R_2}\cos(2\beta_m + \mu)}{1-R_2} \right]^{-1} . \tag{2-57a}$$

Of special interest is the case where $R_2 \approx 1$. In this instance, the standing wave ratio V (see §2.3.2) is substituted for "$1 - R_2$" and Eq. (2-57a) is approximated

$$\frac{A}{T} \approx \kappa - 1 + \tfrac{1}{4} V [Q_1 + 2\,U_0 \cos(2\beta_m + \mu)] \tag{2-57b}$$

where $V >> 2$ because R_2 is nearly unity. This shows that A/T depends linearly on V.

2.9.3.4 The net flux ratio derived from the characteristic matrix

This method is applicable to either a single layer or a group of layers [Baumeister 69]. As shown in Fig. 2-91, the characteristic matrix **M** of the absorbing multilayer is computed as in Eq. (2-33a). The normalized admittance $\hat{Y}$ [see Eq. (2-16b)] of the stack (or substrate) on the emergent side of the absorbing multilayer is given. Equation (2-33b) is written at normal incidence as

$$\begin{bmatrix} \hat{E}_0 \\ \mathcal{Z}_0 \hat{H}_0 \end{bmatrix} = \begin{bmatrix} \hat{m}_{11} & \hat{m}_{12} \\ \hat{m}_{21} & \hat{m}_{22} \end{bmatrix} \begin{bmatrix} \hat{E}_1 \\ \mathcal{Z}_0 \hat{H}_1 \end{bmatrix} \tag{2-57c}$$

where $\hat{E}_0$, $\hat{H}_0$, $\hat{E}_1$ and $\hat{H}_1$ are the total electric and magnetic fields (components parallel to the interface at nonnormal incidence) at the incident and emergent sides of the absorbing multilayer, respectively. The matrix elements $\hat{m}_{11}$, $\hat{m}_{12}$, $\hat{m}_{21}$ and $\hat{m}_{22}$ are those of the absorbing stack of which Ψ is to be found. The admittance [see Eq. (2-16a)] links the fields $\hat{E}_1$ and $\hat{H}_1$ on the emergent side of the layer

$$\hat{H}_1 = \hat{Y}\,\Upsilon_0\,\hat{E}_1 . \tag{2-57d}$$

Table 2-17 Net flux ratio in terms of matrix elements

Coefficient	Expression
D_0	$a_{11}\,a_{21} + b_{11}\,b_{21}$
D_{10}	$a_{21}\,a_{12} + a_{11}\,a_{22} + b_{11}\,b_{22} + b_{21}\,b_{12}$
D_{01}	$-a_{21}\,b_{12} - a_{11}\,b_{22} + a_{22}\,b_{11} + a_{12}\,b_{21}$
D_2	$a_{12}\,a_{22} + b_{12}\,b_{22}$

Coefficients in Eq. (2-58a) in terms of the matrix elements in Eq. (2-31c).

The irradiance emerging from the absorbing multilayer is found from Eq. (2-71i)

$$S_1 = \tfrac{1}{2}\,\Re\,(\hat{E}_1\,\hat{H}_1^*) = \tfrac{1}{2}\,\Re\,(\hat{E}_1\,Y^*\,\hat{E}_1^*)\,Y_0 = \tfrac{1}{2}|\hat{E}_1|^2\;Y_r\,\Upsilon_0 \tag{2-57e}$$

where Y_r appears in Eq. (2-16b). The net flux incident is obtained by substituting the $\hat{E}_0$ and $\hat{H}_0$ from Eq. (2-57c) into Eq. (2-71i)

$$S_0 = \tfrac{1}{2}\Re\,(\hat{E}_0\,\hat{H}_0^*) = \tfrac{1}{2}\Re\,[(\hat{m}_{11}\,\hat{E}_1 + \mathcal{Z}_0\hat{m}_{12}\hat{H}_1)(\Upsilon_0\,\hat{m}_{21}\,\hat{E}_1 + \hat{m}_{22}\hat{H}_1)^*] . \tag{2-57f}$$

Both S_0 and S_1 are independent of the properties of the incident medium. Substitution of Eq. (2-57e), Eq. (2-57f) and the matrix elements in Eq. (2-31c) into Eq. (2-55a) provides a denominator for Ψ which is quadratic in both Y_r and Y_j

$$\Psi = \frac{Y_r}{D_2 (Y_r^2 + Y_j^2) + D_{01} Y_j + D_{10} Y_r + D_0} \tag{2-58a}$$

where the coefficients cited in Table 2-17 are expressed in terms of the matrix elements.

2.9.3.5 Net flux ratio: its dependence upon external reflection coefficient

Several methods can be used to prove the following:

2.9.3.5 (a) Contours of constant Ψ are circles when plotted in the admittance plane. The proof follows by inspection of Eq. (2-58a). If Ψ is a constant and Y_r, Y_j are variables, this equation represents a conic — specifically, a circle. Because Ψ is a single-valued function of a given Y_r and Y_j, it follows that the circles must be nested.

2.9.3.5 (b) Contours of constant Ψ are circles when plotted in the amplitude reflection coefficient plane. As discussed in §2.4.2.5, Eq. (2-23b) maps the complex variable z into the w plane as *circles into circles*. Therefore, it follows from Eq. (2-16h) that circles in the admittance plane map into circles in the amplitude reflection coefficient plane.

Transformation of Eq. (2-58a) into the ρ plane yields the following equation for Ψ:

$$\Psi = \frac{1 - |\hat{\rho}|^2}{e_2 |\hat{\rho}|^2 + e_{10} \rho_x + e_{01} \rho_y + e_0} \tag{2-58b}$$

where ρ_x and ρ_y are the real and imaginary parts of $\hat{\rho}$, respectively. The coefficients are functions of the matrix elements, as cited in Table 2-18.

Lest there be confusion, it is emphasized that the $\hat{\rho}$ in Eq. (2-56a) is determined at the *interface* between the absorbing layer and the emergent medium as shown in Fig. 2-89. This is termed the *internal* $\hat{\rho}$. The $\hat{\rho}$ in Eq. (2-58b) is the *external* $\hat{\rho}$ — the amplitude reflection coefficient of the medium or multilayer on the emergent side of the absorbing layer (or multilayer). This $\hat{\rho}$ is measured *in air*, as shown in Fig. 2-91.

Table 2-18
Net flux ratio in terms of matrix elements — reflection coefficient

Coefficient	Expression
e_0	$a_{11}(a_{21} + a_{22}) + a_{12}(a_{21} + a_{22}) + b_{11}(b_{21} + b_{22}) + b_{12}(b_{21} + b_{22})$
e_{10}	$2(a_{21} a_{11} - a_{22} a_{12} + b_{21} b_{11} - b_{12} b_{22})$
e_{01}	$2(a_{11} b_{22} + a_{21} b_{12} - a_{22} b_{11} - a_{12} b_{21})$
e_2	$a_{11}(a_{21} - a_{22}) - a_{12}(a_{21} - a_{22}) + b_{11}(b_{21} - b_{22}) - b_{12}(b_{21} - b_{22})$

Coefficients in Eq. (2-58b) in terms of the matrix elements in Eq. (2-31c).

The radius and center of a circle of constant Ψ is found. Given the coefficients e_0, e_{10}, e_{01} and e_2 in Table 2-18, the following constant is defined

$$\eta \equiv e_2 + \Psi^{-1} \ . \tag{2-58c}$$

The center of a circle is at

$$\rho_x = -\tfrac{1}{2} e_{10} \eta^{-1} \tag{2-58d}$$

$$\rho_y = -\tfrac{1}{2}\, e_{01}\, \eta^{-1} \tag{2-58e}$$

and its radius Γ is

$$\Gamma = \sqrt{\frac{\Psi^{-1} - e_0}{\eta} + \left[\frac{e_{10}}{2\eta}\right]^2 + \left[\frac{e_{01}}{2\eta}\right]^2} \ . \tag{2-58f}$$

It follows from Eqs. (2-58d) and (2-58e) that the loci of the centers of the circles lie on a line intersecting the origin whose angle ξ is

$$\xi = \arctan(e_{01}\, e_{10}^{-1}) \ . \tag{2-58g}$$

This line is *dashed* in Fig. 2-92, which shows several iso-net-flux-ratio circles for a silver film in the visible part of the spectrum.

Table 2-19 Net flux ratio produced by three matching stacks

Theory cited in section	Parameter	Optical thickness of the H' layer			Refer to Fig.
		$0.20\,\lambda_0$	$0.25\,\lambda_0$	$0.30\,\lambda_0$	
2.9.3.4	Admittance*	$1.782 + j\,6.649$	$29.37 + j\,0$	$1.782 - j\,6.649$	2-89
2.9.3.5	$\hat{\rho}$ external*	$0.9288\exp(j196°)$	$0.9341\exp(j180°)$	$0.9288\exp(j164°)$	2-89
2.9.3.2	$\hat{\rho}$ internal¶	$3.495\exp(j210°)$	$0.9885\exp(j196°)$	$0.2662\exp(j203°)$	2-87
2.9.3.3	R_2 ‡	0.7306	0.7306	0.7306	2-88
2.9.3.3	μ	36.1°	36.1°	36.1°	2-88
2.9.3.3	$2\beta_m$	144°	180°	216°	2-88
all	Ψ	0.9204	0.6716	0.3942	

*Normalized admittance measured at the outermost layer of the stack whose design is shown below.

¶Measured at the emergent interface of the *absorbing layer*.

‡*Internal* radiant reflectance measured at the interface between the H' layer and the outermost L layer of the stack shown below.

At λ_0 of 650 nm, net flux ratio Ψ produced at $\lambda = \lambda_0$ by the design

air H' (L H)^2 *glass*

on the emergent side of an *absorbing layer* whose metric thickness is 30 nm and optical constant is $\hat{n} = 0.172 - j\,4.196$. H, H', L and *glass* represent media of indices 2.30, 2.30, 1.35 and 1.517, respectively. The optical thickness of H and L is $\lambda_0/4$. The parameters R_2, μ and β_m are explained in §2.9.3.3.

2.9.3.6 Numerical example

Three methods are used to compute the net flux ratio of an absorbing film of silver at a wavelength $\lambda_0 = 650$ nm. Its optical constant is $n - jk = 0.172 - j\,4.196$ and its thickness is 30 nm. The emergent side of the absorbing film is bounded by a dielectric stack

H' (L H)^2 *glass*

where **H** and **L** represent layers of optical thickness $\lambda_0/4$ and refractive indices 2.25 and 1.35, respectively. The optical thickness of the **H'** layer (whose index is 2.25) is varied to step the normalized admittance through three values. Table 2-19 summarizes the results. The magnitude of the *external* $\hat{\rho}$ appearing in Eq. (2-58b) is always less than 1.00. The magnitude of the *internal* $\hat{\rho}$, however, can exceed unity (see §2.3.5). The stack **L H L H** *glass* has a reduced admittance less than one and pure real—as required in the formulation of the Kard theory.

2.9.4 Maximum net flux ratio and its attainment

2.9.4.1 Introduction

From the fact that Ψ contours are circles in either the admittance plane or amplitude reflection coefficient plane, it follows *ab intuito* that a circle of zero radius exists. This is where Ψ attains a maximum. Although the three formalisms in §2.9.3.1 produce the same Ψ_{max}, they are included here for completeness. If all else fails, Ψ_{max} can be found by differentiating Ψ with respect to the variables—variables such as ρ_x, ρ_y, Y_r or Y_j.

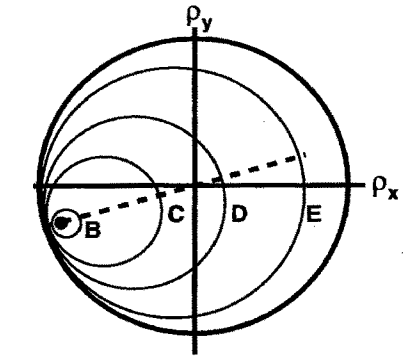


Fig. 2-92 Circles of constant Ψ in the complex amplitude reflection coefficient plane for Ψ values of (*B*) 0.9, (*C*) 0.7, (*D*) 0.5 and (*E*) 0.2 for the silver film captioned in Table 2-20. The solid circle is Ψ_{max}.

2.9.4.2 Maximum net flux ratio

Following Berning [57], the maximum that Ψ can attain in a single layer is

$$\Psi_{max} = \kappa - \sqrt{\kappa^2 - 1} \qquad (2\text{-}59a)$$

where κ appears in Eq. (2-56d). Hence Ψ_{max} depends upon only the *n, k and h*/λ of a given absorbing layer. Kard [60, 71] arrives at the same result by a different formalism. In the more general case, given the matrix elements (and, hence, the constants appearing in Table 2-17), Ψ_{max} is [Baumeister 69]

$$\Psi_{max} = (D_{10} + 2\Omega)^{-1} \qquad (2\text{-}59b)$$

where

$$\Omega \equiv \sqrt{D_2 D_0 - \tfrac{1}{4} D_{01}^2} \; . \qquad (2\text{-}59c)$$

In the case where β_j is sufficiently small that the *cosh* in Eq. (2-56d) can be expanded in a series and also assuming that $k >> n$, Kard [60] shows that

$$\Psi_{max} \approx \left[1 + 2n\,\beta_j^2 / (|k|\sqrt{3})\right]^{-1} \; . \qquad (2\text{-}59d)$$

The next question is: What optical properties must the emergent stack have in order that Ψ_{max} is achieved? Three methods of computing Ψ are developed in §2.9.3.2, §2.9.3.3 and §2.9.3.4. There also are three methods of expressing the properties of the emergent medium needed to attain Ψ_{max}.

2.9.4.3 Internal |ρ| and δ at the metal interface

Following the formalism of §2.9.3.2, the magnitude of the internal reflection coefficient $|\hat{\rho}|$ and phase shift upon reflection δ are computed as follows [Holloway 69, Berning 80]. The first step is to find δ as

$$\delta = \arctan\left[\frac{\Psi_{max}\cos(2\beta_r) - 1}{-\Psi_{max}\sin(2\beta_r)}\right] \tag{2-59e}$$

and the magnitude of the reflection coefficient as

$$|\hat{\rho}| = \sqrt{\frac{\Psi_{max}\exp(-2\beta_j) - 1}{1 - \Psi_{max}\exp(2\beta_j)}} \; . \tag{2-59f}$$

β_r and β_j appear in Eq. (2-5b). The negative sign in the denominator of Eq. (2-59e) forces δ into quadrant III. The next step is to translate the foregoing *internal* complex reflection coefficient $\hat{\rho}$ into the *external* $\hat{\rho}_1$ that a multilayer must attain, in order that Ψ_{max} is achieved. One approach is to use detailed equations [Holloway 69]. Another method is easy to implement if a calculator contains polar-to-Cartesian conversion routines. The first step is to find the Fresnel coefficient $\hat{r}$ at the metal-air interface — the metal (of complex optical constant $\hat{n}$) is on the left and the emergent medium of air is on the right

$$\hat{r} = \frac{\hat{n} - 1}{\hat{n} + 1} \; . \tag{2-59g}$$

Inversion of Eq. (2-36d) (with zero phase thickness) yields

$$\hat{\rho}_1 = \frac{\hat{\rho} - \hat{r}}{1 - \hat{r}\,\hat{\rho}} \; . \tag{2-59h}$$

The $\hat{\rho}$ is obtained from Eqs. (2-59e) and (2-59f). In other words, if Ψ_{max} is to be attained for this metallic film, the multilayer on its emergent side should have the $\hat{\rho}_1$ cited in Eq. (2-59h), as measured *in air*.

2.9.4.4 Formulation of Kard

Kard [60, 71] first published the conditions for the attainment of Ψ_{max}. Inspection of Eq. (2-57a) shows that for a given R_2, Ψ attains a maximum when the argument of the cosine is 180°. This leads to

$$2\beta_m + \mu = \pi \; . \tag{2-60a}$$

The foregoing is solved for the phase thickness β_m of the layer adjacent to the absorbing layer

$$\beta_m = 0.5\,(\pi - \mu) \tag{2-60b}$$

and is less than a quarterwave optical thickness. μ appears in Eq. (2-56g). The derivation of the value of R_2 is a bit more tedious and involves finding an extremum of the function

$$\left[Q_1\,(1 + R_2) - 2\,U_0\sqrt{R_2}\right](1 - R_2)^{-1} \tag{2-60c}$$

which consists of the last two terms of Eq. (2-57a). Solving for R_2 and converting it into the standing wave ratio V_2 at the interface results in

$$V_2 = \sqrt{\frac{Q_1 + U_0}{Q_1 - U_0}} \; . \tag{2-60d}$$

2.9.4.5 Admittance at the maximum net flux ratio Ψ_{max}

When Ψ_{max} is attained, the normalized admittance of the medium or multilayer on the emergent side of a stack of absorbing films is

$$Y_r = \Omega D_2^{-1} \tag{2-60e}$$

and

$$Y_j = -\tfrac{1}{2} D_{01} D_2^{-1} \tag{2-60f}$$

where Ω appears in Eq. (2-59c) [Baumeister 69]. D_{01} and D_2 appear in Table 2-17. The matrix elements for a single layer are obtained from §2.4.4.2. This admittance is transformed into an external amplitude reflection coefficient using Eq. (2-16h).

2.9.4.6 Transformation of admittance

If $\hat{Y}_m$ is the normalized admittance at the *emergent* interface that produces Ψ_{max} for a single metal layer, then the admittance at the *incidence* interface is the complex conjugate Y_m^*. Proof of this theorem appears in §8.9.1.

2.9.4.7 Numerical example: maximum net flux ratio Ψ_{max} of a single layer

Table 2-20 shows the Ψ_{max} of the silver film listed in the caption to Table 2-19. For the multilayer on the emergent side of that layer, Table 2-20 lists the required normalized admittance, as well as the internal $\hat{\rho}$ and external $\hat{\rho}$ needed to achieve Ψ_{max}. For the Kard formulation, it is necessary to choose a refractive index for the dielectric layer contiguous to the silver—2.25 is used. The procedures of Chap. 6 are used to design the stack that produces the requisite standing wave ratio V_2 at the film-multilayer interface.

Table 2-20 Maximum net flux ratio of a silver layer

Parameter	Equation number or comments	Value
Ψ_{max}	(2-59a) and (2-59b)	0.9255
$\hat{Y}$ at Ψ_{max}	(2-60e) and (2-60f)	$2.461 + j\,6.176$
$\hat{\rho}$ at Ψ_{max}	(2-59e) and (2-59f), $\hat{\rho}$ *is internal*	$3.224 \exp(j220.6°)$
$\hat{\rho}$ at Ψ_{max}	(2-59h) and (2-59e), $\hat{\rho}$ *is external* (2-59h) and (2-59f)	$0.8964 \exp(j196°)$
μ	(2-56g)	35.4°
β_m	(2-60b)	72.3°
R_2	(2-60d) and (2-19e)	0.6329
V_2	(2-60d)	8.782

At $\lambda_0 = 650$ nm, maximum net flux ratio Ψ_{max} of a silver film whose optical constant is $\hat{n} = 0.172 - j4.196$. It is 30 nm in metric thickness. Listed are the normalized admittance $\hat{Y}$, internal amplitude reflection coefficient $\hat{\rho}$ and external amplitude reflection coefficient $\hat{\rho}$ needed to achieve Ψ_{max}. β_m is the phase thickness of a dielectric film of index 2.25 on its emergent side; μ, R_2 and V_2 are defined in §2.9.3.3.

2.9.4.8 Minimum value of the net flux ratio

Much attention is given to methods of maximizing Ψ. Ψ values close to zero may also be produced. How close? The A/T appearing in Eq. (2-57b) is proportional to the standing wave ratio V of the stack on the emergent side. A/T can be large (and, hence, Ψ small) by making V as large as 10^4.

2.9.4.9 Brief summary

The net flux ratio Ψ of an absorbing layer depends upon the n, k and h/λ of that layer, as well as the normalized admittance of the medium or multilayer on its emergent side. Ψ can be either maximized or minimized by adjusting this admittance.

2.9.5 Computation of absorption using net flux ratio

2.9.5.1 Basic equations

The absorption A [see Eq. (2-12b)] is related to the reflectance and transmittance of a stack. The absorption of a single layer or a group of layers is

$$A \equiv \frac{\text{net flux incident} - \text{net flux emergent}}{\text{flux incident on the stack}} \quad . \tag{2-61a}$$

In terms of the components of the Poynting vector, the absorption in layer 1 is

$$A_1 = [S_0^{(+)} - S_0^{(-)} - (S_1^{(+)} - S_1^{(-)})] / S_0^{(+)} \quad . \tag{2-61b}$$

In terms of Ψ [Berning 63], the absorption in layers 1, 2 and 3 is

$$A_1 = (1 - R)(1 - \Psi_1) \tag{2-61c}$$

$$A_2 = (1 - R)\Psi_1 (1 - \Psi_2) \tag{2-61d}$$

$$A_3 = (1 - R)\Psi_1 \Psi_2 (1 - \Psi_3) \quad . \tag{2-61e}$$

In summary, the absorption on the incident side of the ith layer is proportional to $1-R$ of the stack. Given Ψ_i of the layer itself, it is proportional to $1 - \Psi_i$ and also the product of the Ψ values of all layers on its incident side. In the general case,

$$A_i = (1-R)(1 - \Psi_i) \prod_{m=1}^{i-1} \Psi_m \quad . \tag{2-61f}$$

The total absorption in a stack (containing ℓ layers) is

$$A = \sum_{i=1}^{\ell} A_i \quad . \tag{2-62a}$$

Often equations are expressed more easily in A/T. It follows that

$$1 - \Psi_i \approx \frac{A_i}{T_i} \quad . \tag{2-62b}$$

when $A/T << 1$.

2.9.5.2 Example: computation of the absorption in each layer

Shown in Fig. 2-93 is a stack of three dissimilar metals. Each has a 15 nm metric thickness. Practical problems such as the diffusion of the gold and aluminum are ignored. Table 2-21 lists the absorption at $\lambda = 500$ nm, as found by calculating the admittance from the elements of the characteristic matrix (see §2.9.3.4).

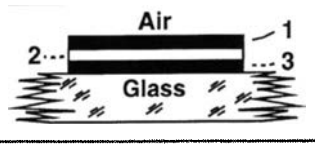


Fig. 2-93 Three absorbing layers are deposited upon a substrate. Design data appear in Table 2-21.

The procedure is the following. The normalized admittance of the substrate is $1.52 + j\,0$. The Ψ_3 of layer 3 is then determined by computing the elements of the characteristic matrix for that layer and substituting them into Eq. (2-58a). Because the matrix elements now exist for layer 3, the normalized admittance $\hat{Y}_3$ of that layer deposited on the glass substrate is found using Eq. (2-33d) — using $\hat{\gamma}_s = 1 + j\,0$, since the angle of incidence

is zero. Then, given $\hat{Y}_3$ and the matrix elements for layer 2, Ψ_2 is found. The foregoing procedure is repeated for layer 1, which provides Ψ_1. The characteristic matrix of the entire stack is computed from the product of the matrices for each of the three layers and the reflectance R is found from Eq. (2-34f). Equations (2-61c), (2-61d) and (2-61e) provide the absorption.

The column "$\prod\Psi_j$" in Table 2-21 contains the product of the Ψ values for the films *above* a given layer. It is 1.00 for layer one—the outermost layer of the stack. The column "$\prod\Psi_j(1-\Psi_i)$" is the *relative* absorption for that layer. According to Eq. (2-61f), it is the *actual* absorption after it is multiplied by $1-R$. Because the radiant reflectance is 0.862, $1-R = 0.138$. The column "$\prod\Psi_j(1-\Psi_i)$" is multiplied by 0.138 to obtain the absorption A_i' listed in Table 2-21.

2.9.5.3 Example — metal layers with dielectric overcoat

The theorems concerning the net flux ratio are illustrated by overcoating the stack cited in §2.9.5.2 with five nonabsorbing layers that reduce its reflectance almost to zero at $\lambda_0 = 500$ nm. The design of such a stack is straightforward, since Table 2-21 shows that the normalized admittance of the three metal layers is $0.312 - j\,2.70$. Given this admittance, §4.4 describes several methods that can be used to design the antireflection coating. The final design is

air **H L H L H'** $\mathbf{M_1}$ $\mathbf{M_2}$ $\mathbf{M_3}$ *glass*

where $\mathbf{M_1}$, $\mathbf{M_2}$ and $\mathbf{M_3}$ and *glass* represent the metal layers and substrate cited in §2.9.5.2. The respective optical thicknesses of the layers **H**, **H'** and **L** are 125 nm, 68.7 nm and 125 nm. The indices of the layers **H** and **L** are 2.35 and 1.38, respectively. Because the reflectance at $\lambda_0 = 500$ nm is reduced to 0.5% by the overcoat, $1-R$ is ≈ 1.0. It follows that the absorption is essentially the same as "$\prod\Psi_j(1-\Psi_j)$" in Table 2-21.

Little additional computation in Table 2-21 is necessary to find the absorption in this overcoated stack. The reason is that the reflectance is the only parameter in Eq. (2-61f) altered by the addition of the five-layer overcoat. The justification of the first sentence of this paragraph is in §2.9.2.4, which asserts that the Ψ depends upon the optical properties of the films on the *emergent* side of any given layer. Because Ψ is unaltered in the overcoated stack, the Ψ values are unchanged and hence the *relative* absorption in any given layer is the same. It is also important that the five-layer overcoat is composed of *dielectric* layers. The theorem of §2.9.2.2 is utilized; the Ψ of any dielectric stack is 1.0. This means that the Ψ values of the three metal layers are unaltered by the dielectric overcoat and thus the relative absorption of the stack is unperturbed.

The results of Table 2-21 would no longer be valid if the medium on the emergent side of the three metal layers were altered in any way, as, for example, changing the substrate index from 1.52 to 1.62. The Ψ for each layer would have to be recomputed.

This example illustrates the concept of potential transmittance (see §2.9.2.3). The product of the Ψ values in Table 2-21 is 17.2%. This is, indeed, the transmittance into glass substrate after the dielectric overcoat—the antireflection coating—has been applied.

Table 2-21 Absorption in metallic layers

i	ID	$\hat{n}$	$\hat{Y}$	Ψ_i	$1-\Psi_i$	$\prod\Psi_j$	$\prod\Psi_j\,(1-\Psi_i)$	A_i'	A_i''
0*	*air*	$1.00+j\,0$							
			$0.312-j\,2.70$						
1	M_1	$0.11-j2.6$		0.784	0.216	1.0	0.216	2.98%	21.5%
			$0.692-j\,2.82$						
2	M_2	$0.84-j\,1.85$		0.505	0.495	0.784	0.388	5.36%	38.6%
			$1.30-j\,4.49$						
3	M_3	$0.64-j5.5$		0.435	0.565	0.396	0.224	3.10%	22.3%
			$1.52+j\,0$						
4§	*sub*	$1.52+j\,0$							
			Product Ψ	0.172			**Sums**	11.4%	82.4%

*Incident medium *air* §Substrate *sub*

Net flux ratio and absorption A_i' in the ith layer at a wavelength λ_0 of 500 nm of

air $\mathbf{M_1\ M_2\ M_3}$ *substrate*

where each layer is 15 nm in metric thickness. The normalized admittance $\hat{Y}$ is at the upper interface of the indicated layer or medium. $\prod\Psi_j$ is the final term in Eq. (2-61f). The absorption A_i'' is that of the design

air **H L H L H' M₁ M₂ M₃** *substrate*

where the indices of *substrate*, **H**, **H'** and **L** are 1.52, 2.30, 2.30 and 1.38, respectively. The optical thicknesses of **H**, **H'** and **L** are 0.25 λ_0, 0.137 λ_0 and 0.25 λ_0, respectively.

2.9.6 Loss in weakly absorbing layers

The level of absorption of the layers in §2.9.5.2 is large because they are metallic. There are many instances, however, in which dielectric films are weakly absorbing — where $k << 1$. Approximations are made for A/T or Ψ. These serve several purposes:

(1) Simplified equations can be programmed on small calculators.

(2) The designer gains insight. This hints at modifications of the stack that may enhance or diminish the absorption of a layer.

2.9.6.1 Absorption in terms of the square electric field

Kossel [69] used this concept to design interference photocathodes. Other authors [Newnam 73, Hemmingway 73, Apfel 76, Lissberger 78, Arnon 77, Sparks 79, Bennett 80, Carniglia 80, Lissberger 81] have also exploited this technique. The treatment here follows Arnon [78a] and applied to normal incidence.

Starting with Eq. (2-13h), the infinitesimal attenuation of the irradiance dS of a thin lamina of thickness dz in the ith layer is

$$dS = -S(z)\ \alpha_i\ dz \tag{2-62c}$$

where α_i appears in Eq. (2-14b). Combining Eqs. (2-15f) and (2-71i), Eq. (2-62c) is rewritten in terms of the total electric field $\hat{E}_i$ in the ith layer

$$dS = n_i\ |\hat{E}_i(z)|^2\ \alpha_i\ \Upsilon_0\ dz\,. \tag{2-62d}$$

As evident from the derivation of Eq. (2-14a) in §2.2.4.4, the S is the magnitude of the total Poynting vector. It is not the component $\perp$ to the interface appearing in Eq. (2-12b). At nonnormal incidence a "sec θ_i" factor should be included to account for the slant distance that the ray travels though the absorbing medium. The total electric field in Eq. (2-62d) is expressed as the sum of the incident and reflected components that differ in phase by an amount δ

$$|\hat{E}_i(z)|^2 = |\hat{E}_i^{(+)}(z) + \hat{E}_i^{(-)}(z)|^2 \tag{2-62e}$$

or

$$|\hat{E}_i(z)|^2 = |\hat{E}_i^{(+)}|^2 [\exp(-j\beta") + |\hat{\rho}_i| \exp(j\beta" + j\delta)] [\exp(j\beta") + |\hat{\rho}_i| \exp(-j\beta" - j\delta)] \tag{2-62f}$$

where

$$\beta" \equiv 2\pi n_i \lambda_0^{-1} z \ . \tag{2-62g}$$

Equation (2-62f) simplifies to

$$|\hat{E}_i(z)|^2 = |\hat{E}_i^{(+)}|^2 [1 + |\hat{\rho}_i|^2 + 2|\hat{\rho}_i| \cos(2\beta" + \delta)] \ . \tag{2-62h}$$

Integration provides the total relative absorption in a layer of metric thickness h_i

$$\int_0^{h_i} |E_i(z)|^2 \, dz \propto (1 + |\hat{\rho}_i|^2) h_i + 2|\hat{\rho}_i| \int_0^{h_i} \cos(4\pi n_i \lambda_0^{-1} z + \delta) \, dz \ . \tag{2-63a}$$

Of special interest is the situation where the film has a matched quarterwave optical thickness (see §2.7.1). In this case the last term in Eq. (2-63a) is zero and from Eq. (2-61a) the total absorption in the ith layer is approximated

$$A_i \approx \frac{n_i h_i (1 + |\hat{\rho}_i|^2) \alpha_i |\hat{E}_i^{(+)}|^2}{n_0 |\hat{E}_0^{(+)}|^2} \ . \tag{2-63b}$$

Young [61a] observed that when $k << 1$ and the absorption is sufficiently low, the net irradiance amongst the various layers of the stack is conserved. This leads to the approximate equality

$$n_0 |\hat{E}_0^{(+)}|^2 (1 - |\hat{\rho}_0|^2) \approx n_i |\hat{E}_i^{(+)}|^2 (1 - |\hat{\rho}_i|^2) \tag{2-63c}$$

where $\hat{\rho}_0$ is the amplitude reflection coefficient of the entire stack and $\hat{E}_0^{(+)}$ is the incident component of the electric field. $\hat{\rho}_i$ is the amplitude reflection coefficient on the emergent side of the ith layer. Equation (2-63b) becomes

$$A_i \approx (1 - R) \alpha_i h_i \frac{1 + |\hat{\rho}_i|^2}{1 - |\hat{\rho}_i|^2} \tag{2-63d}$$

or

$$A_i \approx (1 - R) \alpha_i h_i G_i \tag{2-63e}$$

where

$$G_i \equiv \tfrac{1}{2} (V_i + V_i^{-1}) \tag{2-63f}$$

and where R is the radiant reflectance of the stack in which the absorbing film is embedded. V_i is the standing wave ratio on the emergent side of the ith layer. When $V_i >> 1$, then $G_i \approx V_i / 2$. It is emphasized that Eq. (2-63d) predicts the absorption in layers whose optical thicknesses are quarterwaves or multiples thereof.

The foregoing equation shows that for a multilayer with a given reflectance and a layer with a specified α, the absorption is proportional to: (1) layer thickness; and (2) the standing wave ratio on its emergent side. Chapters 5 and 8 discuss how these concepts

are applied to control the absorption. The use of Eq. (2-63d) is illustrated in Table 7-21, which lists the absorption in each layer of a bandpass filter.

A final observation is that the approximate A_i in Eq. (2-63d) relates to the more exact formulation of Kard. Equation (2-57a) shows that A/T is proportional to the parameter "$(1+R)/(1-R)$" and includes an *interference* term that vanishes when the optical thickness is a quarterwave or a multiple thereof.

2.9.6.2 Square electric field (TE wave)

Several authors, as referenced in §2.9.6.1, have developed methods of computing the square electric field using the $E+E^-$ matrix (see §2.4.2.3). Both this and the characteristic matrix have been used. The parameter which is computed is the *normalized* square electric field $|\hat{E}_{i,n}|^2$ appearing in Eq. (2-63g)

$$|\hat{E}_{i,n}|^2 \equiv \frac{|\hat{E}_i|^2}{|\hat{E}_0^{(+)}|^2} \tag{2-63g}$$

where $\hat{E}_i$ is the total electric field at an arbitrary point within the ith layer and $\hat{E}_0^{(+)}$ is the electrical field component in the incident medium flowing in the +z direction.

The first task is to compute the characteristic matrix for the stack and thence the denominator of Eq. (2-63g). As indicated on the left of Eqs. (2-7a) and (2-7b), the total E or H field in the incident medium is the sum of the "+" and "–" components. Combining Eqs. (2-7a), (2-7c) and (2-33b) for the characteristic matrix results in

$$\hat{E}_0^{(+)} + \hat{E}_0^{(-)} = (m_{11} + j m_{12}\, \hat{\gamma}_s\, n_s)\ \hat{E}_s^{(+)} \tag{2-63h}$$

and

$$n_0 \cos\theta_0\, (\hat{E}_0^{(+)} - \hat{E}_0^{(-)}) = (j m_{21} + m_{22}\, \hat{\gamma}_s\, n_s)\, \hat{E}_s^{(+)}\ . \tag{2-64a}$$

The matrix elements m_{11}, m_{12}, m_{21} and m_{22} are complex when the films are highly absorbing. A consequence is that the algebra to find $\hat{E}_0^{(+)}$ is more complicated. The matrix elements simplify when the absorption is sufficiently small that they are treated as nonabsorbing. In Eqs. (2-63h) and (2-64a), m_{11}, m_{22} are pure real and $j m_{12}$, $j m_{21}$ are pure imaginary—these are the matrix elements of the *entire stack.* Solving those equations for the square of the incident electric field component provides

$$|\hat{E}_0^{(+)}|^2 = \tfrac{1}{4}\left[(m_{11} + m_{22}\, N_s/N_0)^2 + (\, m_{21}\, N_0^{-1} + m_{12}\, N_s)^2\right] |\hat{E}_s^{(+)}|^2 \tag{2-64b}$$

where N_0 and N_s are the effective indices [see Eq. (2-38a)] for the incident medium and substrate, respectively. At normal incidence these reduce to n_0 and n_s, respectively.

If m'_{11}, m'_{12}, m'_{21} and m'_{22} represent the matrix elements at a point in the interior of the stack, the total field at that point is

$$|\hat{E}_i|^2 = \left[(m'_{11})^2 + (m'_{12}\, N_s)^2\right] |\hat{E}_s^{(+)}|^2 \tag{2-64c}$$

where N_s has the same meaning as in Eq. (2-64b). Substitution of Eqs. (2-64b) and (2-64c) into Eq. (2-63g) provides the normalized square electric field

$$|\hat{E}_{i,n}|^2 = \frac{4\left[(m'_{11})^2 + (m'_{12}\, N_s)^2\right]}{(m_{11} + m_{22}\, N_s/N_0)^2 + (\, m_{21}\, N_0^{-1} + m_{12}\, N_s)^2}\ . \tag{2-64d}$$

2.9.6.3 Square electric field (TM wave)

Following Apfel [76], the total electric field for the TM wave is composed of two components: E_y (|| to the interface) and E_z ($\perp$ to the interface). The total square field in the ith medium is

$$|\hat{E}_i|^2 = |\hat{E}_{y,i}|^2 + |\hat{E}_{z,i}|^2 \ . \tag{2-64e}$$

The first step is to find the y component in the foregoing equation. The continuity of the electric and magnetic fields (see Fig. 2-12) at the surface of the outermost layer of the stack requires that

$$\hat{E}_{y,0}^{(+)} + \hat{E}_{y,0}^{(-)} = \hat{E}_{y,0} \tag{2-64f}$$

$$\hat{H}_{x,0}^{(+)} + \hat{H}_{x,0}^{(-)} = \hat{H}_{x,0} \tag{2-64g}$$

where the subscript "0" refers to the incident medium and x, y to the respective components. From Eq. (2-69c) and the wave functions for the E and H fields [see Eq. (2-3a)], Eq. (2-64g) reduces to

$$N_0 \left(\hat{E}_{y,0}^{(+)} - \hat{E}_{y,0}^{(-)} \right) = \hat{H}_{x,0} \mathcal{Z}_0 \tag{2-64h}$$

where N_0 is the effective index [see Eq. (2-38b)] of the incident medium for p polarization and $\mathcal{Z}_0$ is the impedance of free space. Solving Eqs. (2-64f) and (2-64h) for the component of the E field traveling in the +z direction results in

$$\hat{E}_{y,0}^{(+)} = \tfrac{1}{2} \left(\hat{E}_{y,0} + \hat{H}_{x,0} N_0^{-1} \mathcal{Z}_0 \right) \ . \tag{2-64i}$$

The characteristic matrix (see §2.4.4) relates the fields $\hat{E}_{y,0}$ and $\hat{H}_{x,0}$ in the incident medium to those in the emergent medium (i.e., substrate) $E_{y,s}$ and $H_{x,s}$. The magnitude of the vector field $\hat{E}_0^{(+)}$ propagating in the +z direction in the incident medium is related to its y component by

$$\hat{E}_0^{(+)} = \hat{E}_{y,0}^{(+)} \sec\theta_0 \ . \tag{2-65a}$$

Combining Eqs. (2-64i) and (2-65a)—with substitution of the matrix elements from Eq. (2-33c) for $\hat{E}_{y,0}$ and $\hat{H}_{x,0}$—the normalizing factor appearing in Eq. (2-63g) is

$$|\hat{E}_0^{(+)}|^2 = \tfrac{1}{4} \sec^2\theta_0 \left[(m_{11} + m_{22} N_s N_0^{-1})^2 + (m_{21} N_0^{-1} + m_{12} N_s)^2 \right] |\hat{E}_s^{(+)}|^2 \tag{2-65b}$$

where the matrix elements are those of the *entire stack* for the p polarization. $N_0 \equiv n_0 \sec\theta_0$ and $N_s \equiv n_s \sec\theta_s$. θ_0 and θ_s are the angles of incidence in the incident medium and substrate, respectively. It is not valid merely to replace the effective indices with those for p polarization in Eq. (2-64b) to obtain Eq. (2-65b). Although the "effective index prescription" works well when the radiant reflectance or radiant transmittance is computed, it fails to provide the correct electric field.

The next task is to compute the total electric field in the ith layer. $E_{y,i}$ in Eq. (2-64e) is obtained from the characteristic matrix—from Eq. (2-64c)—with the exception that N_s is the effective index for p polarization and the matrix elements are for the TM wave. $E_{z,i}$ in the ith layer is obtained using Eq. (2-69c). For the time-harmonic wave,

$$j\omega\varepsilon_0 \varepsilon_i \hat{E}_{z,i} = -\frac{\partial \hat{H}_{x,i}}{\partial y} \tag{2-65c}$$

where ε_i is the dielectric constant of the ith medium and ω is the angular frequency.

Equations (2-3a) and (2-3b) contain the wave functions for the electric field. Although the wave functions for H_x are not written down, they are essentially the same as those for the E field. Substitution of these wave functions into Eq. (2-65c) leads to

$$|\hat{E}_{z,i}|^2 = L^2 |\hat{H}_{x,i}|^2 n_i^{-4} \mathcal{Z}_0^2 \tag{2-65d}$$

where L appears in Eq. (2-4a). $\hat{H}_{x,i}$ is obtained from the elements of the characteristic matrix appearing in Eq. (2-33c). Substitution of Eq. (2-65d) into Eq. (2-64e) provides the normalized square electric field in the ith layer

$$|\hat{E}_{i,n}|^2 = 4\cos^2\theta_0 \frac{(m'_{11})^2 + (m'_{12} N_s)^2 + L^2 n_i^{-4} [(m'_{21})^2 + (m'_{22} N_s)^2]}{(m_{11} + m_{22} N_s N_0^{-1})^2 + (m_{21} N_0^{-1} + m_{12} N_s)^2} \tag{2-65e}$$

where m'_{11}, m'_{12}, m'_{21} and m'_{22} represent the matrix elements at a point in the interior of the stack. The matrix elements m_{11}, m_{12}, m_{21} and m_{22} are those of the *entire* stack. At nonnormal incidence $L = 0$ and $\cos\theta_0 = 1.00$; Eqs. (2-64d) and (2-65e) are identical.

2.9.6.4 Methods of computing the electric field

Two procedures are used to find the matrix elements at the interior of a multilayer. These elements are substituted into Eq. (2-64d) or Eq. (2-65e). The first method is to start at the substrate and find the fields at points spaced at a distance Δh closer to the surface of the stack. The other approach is to start at the incident medium and to progress towards the substrate. Regardless of which method is used, a necessary preliminary step is to compute the matrix elements of the *entire* stack so that the normalizing constants in Eq. (2-64b) or Eq. (2-65b) are known.

The task is to find the field at a distance Δh_i below the surface of the outermost layer. The phase thickness $\Delta\beta_i$ of the ith incremental section is calculated

$$\Delta\beta_i \equiv 2\pi \Delta h_i n_i / \lambda \tag{2-65f}$$

and thence the *inverse* matrix [see Eq. (2-41g)] for the ith section is generated

$$\begin{bmatrix} m''_{11} & jm''_{12} \\ jm''_{21} & m''_{22} \end{bmatrix} = \begin{bmatrix} \cos(\Delta\beta_i) & -jn_i^{-1}\sin(\Delta\beta_i) \\ -jn_i \sin(\Delta\beta_i) & \cos(\Delta\beta_i) \end{bmatrix} \tag{2-65g}$$

which, when multiplied by the matrix elements m_{11}, m_{12}, m_{21} and m_{22} for the entire stack, yields the matrix elements appearing in the numerators of Eqs. (2-64d) or (2-65e)

$$\begin{bmatrix} m''_{11} & jm''_{12} \\ jm''_{21} & m''_{22} \end{bmatrix} \begin{bmatrix} m_{11} & jm_{12} \\ jm_{21} & m_{22} \end{bmatrix} = \begin{bmatrix} m'_{11} & jm'_{12} \\ jm'_{21} & m'_{22} \end{bmatrix}. \tag{2-65h}$$

It is preferable to use the foregoing procedure when computing the fields in a total reflector, as a quarterwave stack, for example. In this instance the fields are large near the incident medium and diminish rapidly in the interior of the stack. The numerical values of the matrix elements near the top of the stack are more accurate because few arithmetic operations have been made to produce them. Hence, they are almost free of truncation errors. In this approach, the computations start at the incident medium and progress towards the substrate. The other approach starts at the substrate and progresses towards the incident medium.

2.9.6.5 Use of the square electric field to compute the absorption

If the objective is to find the absorption in a given layer, then the formalism developed in §2.9.5.1 provides numerical data, but often little insight. A graph of the square electric field (see §2.9.6.1) enables the designer to visualize its relative strength and, hence, how the absorption is partitioned amongst the various layers.

Starting with Eq. (2-62d), the absorption in the ith layer of metric thickness h_i is proportional to

$$A_i \propto n_i \int_0^{h_i} |E_i(z)|^2 \, dz \tag{2-66a}$$

The integral in the foregoing equation is proportional to the area under the curve of the square electric field. Two methods may be used to account for the n_i in Eq. (2-66a): (1) the ordinate is scaled by a factor of n_i; or (2) $|E_i(z)|^2$ is plotted vs *optical thickness*.

2.10 Appendices—Propagation of electromagnetic waves

§2.10 contains tedious derivations that should be available, but not necessarily as part of the text in §2.1 and §2.2. The sections below present a more rigorous derivation, as in the first three steps in Fig. 2-2.

2.10.1 Assumptions about waves and the medium in which they propagate

The ultimate goal is to find the radiant reflectance, radiant transmittance and phase shift upon reflection of one or more thin layers—the multilayers whose properties are described in Chap. 1. This is accomplished using electromagnetic theory.

The physical model establishes the assumptions about the incident radiant flux and the medium in which it propagates. The flux is described as a linearly polarized monochromatic plane wave with angular frequency ω. This is called a *time harmonic wave* and represents the steady state propagation of the electric and magnetic fields. The solution to these fields is of the form

$$E(x,y,z,t) = E(x,y,z)\exp(j\omega t) \tag{2-69a}$$

where $E(x,y,z)$ is the solution in terms of the spatial coordinates, j is $\sqrt{-1}$ and t is the time.

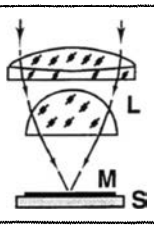


Fig. 2-94 A lens L irradiates a multilayer M with spatially coherent light.

The assumption of monochromatic flux is not restrictive. Given the spectral distribution of the source, the flux emerging from a multilayer can be reconstructed from either an incoherent or coherent superposition of the emerging waves.

The assumption of plane waves is also not restrictive. Suppose, for example, that a highly convergent beam is incident upon the multilayer shown in Fig. 2-94. The beam can be decomposed into an *angular spectrum* of plane waves [Goodman 68]. Each plane wave is propagated through the multilayer with an attenuation of its amplitude and a change in its phase. Diffraction theory reconstructs the emergent wave in the far field as a superposition of these plane waves.

The assumption that the wave is linearly polarized is not restrictive. Physical optics texts [Born & Wolf 75] show that other polarization states may be represented by linear combinations of these linearly polarized waves. An example is that a circular polarized wave is the superposition of two orthogonal linearly polarized waves with a 90° phase difference.

The "exp($j\,\omega\,t$)" time dependence means that the theory is inadequate to predict the transient response of an interference device, as, for example when a picosecond length pulse propagates in a cavity a few centimeters in length [Roychoudhuri 75, Daussy 78]. Another restriction is that the theory does not cope with diffraction losses. The effects of diffraction in optical cavities were first encountered in the early 1960s when the gas laser was analyzed [Fox 63]. Although the behavior of nonlinear materials in optical cavities is a subject of intense interest (for switching devices), this is beyond the scope of this book. The dielectric constant is assumed to be constant with field strength.

Chapter 1 describes multilayer filters containing many layers of different materials. It is assumed that the boundaries between these media are plane and can be represented as a discontinuity in the dielectric constant. This model of *plane boundaries* does not preclude the modeling of scattering in the bulk of the layers or at their surfaces as a perturbation.

Finally, the plane boundaries between successive layers are assumed to be parallel. Any nonparallelism introduces a *walk off* in the beams and a resultant distortion of the emergent wave fronts. Although Kinosita [53] was able to analyze this effect for a single lamina, it would be frightfully complicated in any multilayer structure.

2.10.2 Propagation of EM waves in isotropic homogeneous media

The propagation of plane electromagnetic EM waves in a homogeneous conducting medium is treated in physical optics texts [Stone 63, Born & Wolf 75]; many mathematical details are omitted here. Maxwell's equations are

$$curl\ \vec{E} = -\frac{\partial(\mu\,\mu_0\,\vec{H})}{\partial t} \tag{2-69b}$$

$$curl\ \vec{H} = \frac{\partial(\varepsilon\,\varepsilon_0\,\vec{E})}{\partial t} + \vec{J} \tag{2-69c}$$

where the vectors $\vec{H}$, $\vec{E}$ and $\vec{J}$ represent the magnetic intensity, electric field strength and current density, respectively. ε_0 , μ_0 , ε and μ are, respectively, the dielectric permitivity of free space, the magnetic permeability of free space, the dielectric constant of the material and the magnetic permeability of the material. It is tacitly assumed that there is no permanent electrical charge that can interact with the steady state wave. $\mu = 1$ at frequencies in excess of 1 THz. No experiment has yet been performed that contradicts this fact. The propagation of electromagnetic waves in a magnetic material is treated in terms of its effect on ε, rather than μ [Pershan 67].

$\vec{J}$ is related to $\vec{E}$ in an isotropic medium by

$$\vec{J} = \boldsymbol{\sigma}\,\vec{E} \tag{2-69d}$$

where $\boldsymbol{\sigma}$ (boldface "sigma") is the electrical conductivity at optical frequencies — the reciprocal of the specific resistivity. Although $\boldsymbol{\sigma}$ is calculated in studies that relate the structure of a metal to its optical properties (see, for example, Johnson [74, 75]), little use is made of it in this book. $\boldsymbol{\sigma}$ (boldface) appears only in §2.10. $\sigma \equiv 1/\lambda$ is used elsewhere to represent reciprocal wavelength — or wave number.

At low temperatures (for example, near 4 K) and in a metal with high electron density, the electric field penetration distance is small compared to a wavelength. Under

these conditions the diffusion length of the electrons often exceeds the penetration distance and hence Eq. (2-69d) is invalid. In other words, the current density $\vec{J}$ at a point beneath the surface of the metal is due to both the electric field at *that* point and also *neighboring* points. This is called the *anomalous skin effect* [Sokolov 67]. The theory developed below is inadequate to model such a phenomenon.

A manipulation of the auxiliary equations

$$\vec{\nabla} \bullet (\varepsilon_0 \varepsilon \vec{E}) = 0 \tag{2-69e}$$

and

$$\vec{\nabla} \bullet \vec{H} = 0 \tag{2-69f}$$

(where the symbol • denotes a scalar product) and Eq. (2-69d) results in

$$\nabla^2 \vec{E} = \mu_0 \varepsilon_0 \varepsilon \frac{\partial^2 \vec{E}}{\partial t^2} + \mu_0 \sigma \frac{\partial \vec{E}}{\partial t} . \tag{2-69g}$$

Substitution of the wave function for the electric field in Eq. (2-69a) provides

$$\nabla^2 \vec{E}(x,y,z) = (-\mu_0 \varepsilon_0 \varepsilon \omega^2 + j\mu_0 \sigma \omega)\vec{E} \tag{2-69h}$$

$$= \pm \frac{\omega^2}{c^2} \left(\varepsilon - \frac{j\sigma}{\varepsilon_0 \omega} \right) \vec{E} . \tag{2-69i}$$

The identity $\mu_0 \varepsilon_0 = c^{-2}$ is used, where c is the velocity of light in a vacuum. The right of the foregoing equation introduces the complex dielectric constant $\hat{\varepsilon}$

$$\hat{\varepsilon} \equiv \varepsilon_r - j\varepsilon_j \equiv \varepsilon - \frac{j\sigma}{\varepsilon_0 \omega} . \tag{2-69j}$$

The next step is to analyze a component of the vector electric field $\vec{E}$ represented by the symbol E. The components of the E and H vectors along the x and y axes are shown in Fig. 2-19. A wave function that satisfies Eq. (2-69h) is

$$E(x,y,z) = \hat{E}_0 \exp[-j(k_x x + k_y y + k_z z)] . \tag{2-70a}$$

Substitution of Eq. (2-70a) into Eq. (2-69h) produces the eigenvalue relationship

$$k_x^2 + k_y^2 + k_z^2 = \pm \hat{\varepsilon} \omega^2 c^{-2} . \tag{2-70b}$$

The ± in the foregoing equation ensures that there are two linearly independent solutions to a second-order differential equation.

2.10.2.1 Wave propagation in nonabsorbing media

The conductivity σ of a nonabsorbing medium is zero. For a wave propagating in the z direction, Eq. (2-70b) simplifies to

$$k = k_z = 2\pi \upsilon n / c = 2\pi n / \lambda_0 = 2\pi n \sigma . \tag{2-70c}$$

This introduces the vacuum wavelength λ_0 and the wave number σ. It also incorporates the refractive index n, which is defined as

$$n \equiv \frac{\text{velocity of light in a vacuum}}{\text{velocity of light in the medium}} \tag{2-70d}$$

or, from Eq. (2-70c), as

$$n \equiv \sqrt{\varepsilon} . \tag{2-70e}$$

2.10.2.2 Wave propagation in absorbing media, normal incidence

The conductivity σ is nonzero in an absorbing medium and hence $\hat{\varepsilon}$ in Eq. (2-69j) is complex. The optical constant $\hat{n}$ appears in Eq. (2-1h). It is related to the (complex) dielectric constant by

$$\hat{n} = n - j\,k = \sqrt{\hat{\varepsilon}} \tag{2-70f}$$

where n is the refractive index and k is the extinction coefficient, not to be confused with the propagation constant $\acute{k}$. Equations (2-69j) and (2-70f) provide the identity

$$\varepsilon_r = n^2 - k^2 \tag{2-70g}$$

and

$$\varepsilon_j = 2\,n\,k = \frac{\sigma}{\varepsilon_0\,\omega}\ . \tag{2-70h}$$

The y component of the electric field of the wave propagating in the $+z$ direction in this medium is described by

$$E_y(z,t) = \hat{E}_0 \exp[j(\omega t - \acute{k} z)]\ \exp(-2\pi k z/\lambda_0)\ , \tag{2-70i}$$

which is similar to Eq. (2-2a).

2.10.2.3 Sign of the "time dependence" ωt

A wave propagating in the positive z direction is described mathematically by Eq. (2-70i)

$$E = \hat{E}_0 \exp[j(\omega t - \acute{k} z)]\ , \tag{2-70j}$$

which is consistent with most texts on electrical techniques and microwave theory. Many optics books use

$$E = \hat{E}_0 \exp[i(\acute{k} z - \omega t)] \tag{2-70k}$$

to represent the same wave. This causes no difficulty as long as the optical constant $\hat{n}$ for an absorbing medium is defined as $\hat{n} - j\,k$ in Eq. (2-70j) and as $\hat{n} + i\,k$ in Eq. (2-70k). In either case $i \equiv \sqrt{-1}$ and $j \equiv \sqrt{-1}$. It is useful to maintain the two symbols and to follow the rule: "To convert, change i to $-j$, and vice versa." Muller [69] summarizes the sign conventions used by at least twenty-four different writers.

2.10.2.4 Irradiance of a plane wave

The power (i.e., flux) per unit area (irradiance) transferred by a plane wave is described by the Poynting vector $\vec{S}$

$$\vec{S} = \vec{E} \times \vec{H} = |E|\ |H|\ \vec{s} = S\,\vec{s} \tag{2-71a}$$

where $\times$ denotes a vector product. In a nonabsorbing medium $\vec{S}$ is normal to the wave front, as shown in Fig. 2-19. The vectors $\vec{E}$ and $\vec{H}$ are $\perp$ to a unit vector $\vec{s}$ collinear with $\vec{S}$ and in the direction of propagation.

A time average of the Poynting vector appearing in Eq. (2-71a) is computed. The magnitude of the vector $\vec{S}$ is expressed in terms of the $\hat{E}_0$ from Eq. (2-70i) and $\hat{H}_0$ from Eq. (2-15b). At microwave frequencies it is possible to sample E instantaneously, but at higher frequencies an average of S over a time interval $\mathcal{T}$ is measured. $\mathcal{T}$ is long compared to the period of oscillation $2\pi/\omega$. At $z = 0$, Eqs. (2-70i) and (2-15b) are written

$$E(t) = \Re\,[\,|\hat{E}_0|\,\exp(j\,\xi_1)\,\exp(j\,\omega t)] \tag{2-71b}$$

$$H(t) = \Re\,[\,|\hat{H}_0|\,\exp(j\,\xi_2)\,\exp(j\,\omega t)] \tag{2-71c}$$

where $\Re$ is the "real part of" and ξ_1, ξ_2 are the phases at the origin. Using the identity

$$\Re\,[\exp(j\xi)] = \tfrac{1}{2}\,[\exp(j\,\xi) + \exp(-j\,\xi)]\ , \tag{2-71d}$$

it may be shown that S consists of two parts

$$S = \Re\,[\hat{E}(t,\xi_1)\,\hat{H}(t,\xi_2)] = S_1 + S_2(t) \tag{2-71e}$$

where $S_2(t)$ is time dependent

$$S_2(t) = \tfrac{1}{2}\,|\hat{E}_0|\,|\hat{H}_0|\cos(2\,\omega t + \xi_1 + \xi_2) \tag{2-71f}$$

and S_1 is time independent

$$S_1 = \tfrac{1}{2}\,|\hat{E}_0|\,|\hat{H}_0|\cos(\xi_1 - \xi_2)\ . \tag{2-71g}$$

The time average is denoted by $<S>$

$$<S> = (\mathcal{T})^{-1}\int_0^{\mathcal{T}} S_2(t)\,dt\ . \tag{2-71h}$$

Evaluation of the integral shows that S_2 may be neglected in comparison to S_1 provided $\omega\mathcal{T} >> 1$. The brackets are discarded in the remainder of this book. S is intended to be the time average. The Poynting vector is identical with the irradiance, which is represented by the SI unit $\mathcal{E}$. Another form of Eq. (2-71g) is

$$\mathcal{E} = S = \tfrac{1}{2}\,\Re\,(\hat{E}_0\,\hat{H}_0^*) \tag{2-71i}$$

where the asterisk denotes the complex conjugate.

2.10.3 Propagation in an inhomogeneous and anisotropic medium

2.10.3.1 Introduction

A detailed analysis of optical interference in inhomogeneous and anisotropic layers would constitute a volume in itself. There are thorough treatments in other books and review articles. The discussion here is abbreviated.

Table 2-22 Wave functions for optically inhomogeneous films

Dielectric constant ε vs distance z	Definitions	Wave functions
$\varepsilon(z) = \varepsilon_1\,(1 + a\,z)^m$	$h \equiv h_0 + \int_0^z \sqrt{\varepsilon(z)}\,dz$ where $h_0 \equiv \dfrac{2\sqrt{\varepsilon_1}}{a\,(m+2)}$ and $p \equiv \dfrac{m+1}{m+2}$	$J_{1-p}(h\,k)$ $N_{1-p}(h\,k)$
$\varepsilon(z) = \varepsilon_1\,\exp(a\,z)$	h = same as above, with $h_0 \equiv 2\,a^{-1}\sqrt{\varepsilon_1}$	$J_0(h\,k)$ $N_0(h\,k)$

Wave functions for films in which the functional dependence of the dielectric constant ε vs the distance z is shown in *column* 1. J_q and N_q are Bessel functions (of order q) of the first and second kinds, respectively. a, m and ε_1 are constants. $k \equiv 2\pi/\lambda$.
After Monaco [61], Rajappan [62] and Jacobsson [75].

2.10.3.2 Inhomogeneous films

In previous sections, it is assumed that the dielectric constant ε in Eq. (2-69j) is uniform throughout the medium. Either accidentally or intentionally, however, inhomogeneities

may be introduced into a film and thus ε depends upon the spatial coordinates. Inhomogeneous films are the subject of several review articles [Jacobsson 66a, 66b, 75].

In the most general case, $\varepsilon(x,y,z)$ may vary simultaneously in several directions, as occurs in a gradient index lens [Moore 80]. In optical coatings, however, the gradient of the index is usually in a direction normal to the substrate. The mathematics simplifies considerably when ε varies only in the z direction. Jacobsson [66a, 66b] shows that in a nonabsorbing medium, Eq. (2-69h) is replaced by

$$\nabla^2 \vec{E}(x,y,z) + \nabla(\vec{E} \cdot \frac{\nabla\varepsilon}{\varepsilon}) = \bar{\mu}_0 \varepsilon_0 \varepsilon \omega^2 \vec{E} \quad . \tag{2-71j}$$

The wave functions that satisfy Eq. (2-71j) are not simple exponentials [as in Eq. (2-70a)] but are usually found by numerical integration. There are, however, several index variations that have closed form solutions. These are listed, along with the wave functions, in Table 2-22.

2.10.3.3 Anisotropic and gyrotropic media

The tradition in physical optics is that the modifier *anisotropic* describes a medium in which the refractive index depends upon the orientation of the E vector. The asymmetrical electron charge distribution in certain crystals causes the polarizability, and, hence, the refractive index, to vary as the direction of the E vector changes with respect to the crystal lattice. This is treated fully in Chap. 14 of Born & Wolf [75] and also in the monumental tome of Svivessy [28].

If the coordinate system is chosen properly, the properties of this medium can be described mathematically in terms of a dielectric constant ε which is a tensor (designated as $\bar{\varepsilon}$), rather than a scalar

$$\bar{\varepsilon} = \begin{bmatrix} \varepsilon_1 & a & 0 \\ b & \varepsilon_2 & 0 \\ 0 & 0 & \varepsilon_3 \end{bmatrix} . \tag{2-71k}$$

The constants a and b are nonzero if the medium is gyrotropic and biaxial when $\varepsilon_1 \neq \varepsilon_2 \neq \varepsilon_3$. The simple expressions such as Eq. (2-70b) or Eq. (2-70d) are no longer valid when $\bar{\varepsilon}$ is represented by Eq. (2-71k). The formalism developed by Svivessy [28] or Berreman [72] is used to describe the wave propagation. Azzam [77] briefly reviews this subject.

This book emphasizes interference phenomenon. Examples of optical interference in anisotropic media are rarely encountered. Such interference would be observed, for example, if a plane parallel slab of crystalline quartz were coated with partial reflectors. Another example is a thin sheet of mica (which is optically anisotropic) used as the spacer in a single-cavity bandpass filter [Dobrowolski 59, 62]. Layers have been evaporated that are uniaxial in their optical properties [Barr 48, Blout 48, Hyman 48, Woodward 48]. Billings [50] describes a bandpass filter incorporating an evaporated birefringent film.

Gyrotropic media have received considerable attention because magneto-optical layers are used in data storage devices. An extensive treatment of this subject appears in Sokolov [67]. Lissberger [70, 71a, 71b] reviews applications.

2.10.4 Numerical evaluation of the radiant reflectance and transmittance

Equations (2-33f) and (2-34d) include the parameter *DEN* for the TE and TM polarization, respectively. It is evaluated numerically as follows:

2.10.4.1 TE polarization

Make the following substitutions:

$n_0 \cos\theta_0 \rightarrow e$ (2-72a)

real part of $(\hat{n}_s\, \hat{\gamma}_s) \rightarrow c$ (2-72b)

imaginary part of $(\hat{n}_s\, \hat{\gamma}_s) \rightarrow d$ (2-72c)

2.10.4.2 TM polarization

Make the following substitutions:

$n_0 \sec\theta_0 \rightarrow e$ (2-72d)

real part of $(\hat{n}_s\, \hat{\gamma}_s^{-1}) \rightarrow c$ (2-72e)

imaginary part of $(\hat{n}_s\, \hat{\gamma}_s^{-1}) \rightarrow d$ (2-72f)

2.10.4.3 Evaluation

The matrix elements for a given polarization are in Eq. (2-31c).

$$|DEN|^2 = B_1 + B_2 \quad (2\text{-}72g)$$

where

$$B_1 \equiv [(b_{12}\, c + b_{11} + a_{12}\, d)\, e + a_{22}\, d + b_{22}\, c + b_{21}]^2 \quad \text{and} \quad (2\text{-}72h)$$

$$B_2 \equiv [(-b_{12}\, d + a_{11} + a_{12}\, c)\, e + a_{22}\, c - b_{22}\, d + a_{21}]^2 \quad (2\text{-}72i)$$

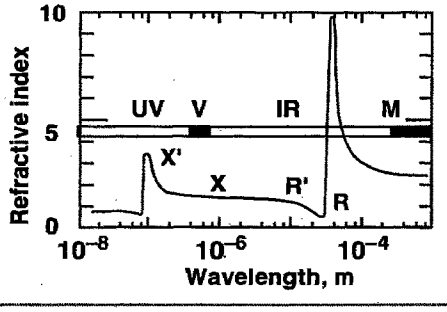


Fig. 2-95 Refractive index vs wavelength of a typical crystalline solid, such as calcium fluoride. UV, V, IR and M identify the ultraviolet, visible, infrared and microwave regions of the electromagnetic spectrum, respectively.

2.11 Appendices

2.11.1 Dispersion of the medium

The Drude-Sellmeier oscillator model is usually effective in modeling the optical properties of a material which is transparent in at least some portions of the spectrum. It is not useful in modeling the properties of a metal.

Treatises on solid state physics [Seitz 40, Moss 59] or physical optics [Milonni 88] develop models for dispersion. The simplest is attributed to Drude [1900] and Sellmeier [1871]. It models a refracting medium as a collection of groups of bound electrons. For the incident flux at an angular frequency ω, each group is excited and oscillates at a different amplitude and phase. The optical properties of the ith group are described in terms of three constants: f_i, $\omega_{0,i}$ and g_i. $\omega_{0,i}$ is the resonant frequency of the electrons and g_i is a measure of their dissipative loss. f_i [which has physical dimensions of $(\text{time})^{-2}$] is proportional to both the oscillator strength and density of the electrons in that group.

The incident time-varying E field — as described by Eq. (2-69a) — polarizes the electrons described in the foregoing paragraph. ε is obtained by computing the induced dipole moment of those electrons. At an angular frequency ω, the complex dielectric constant is

$$\hat{\varepsilon} \equiv \varepsilon_r - j\,\varepsilon_j = n_\infty^2 + \sum_{i=1}^{m} f_i \frac{\omega^2 - \omega_{0,i}^2 - j\,\omega\, g_i}{(\omega^2 - \omega_{0,i}^2)^2 + (\omega g_i)^2} \qquad (2\text{-}78a)$$

where m is the number of oscillators. n_∞^2 is the dielectric constant at high frequencies.

Figure 2-95 shows the refractive index of a typical crystalline solid over a broad frequency range — from wavelengths in the UV to the far IR. The rise of n at X and its increase at X' are produced by the resonance of the UV oscillator. The minimum of n at R and the decline of n at R' are due to a vibration of the crystalline lattice. The optical effects of this vibration may be represented by one or more oscillators.

Although studies exist that relate the ω_0, g and f in Eq. (2-78a) to the physical properties of a solid in the IR [Turner 65, Martin 66, Mitra 85] and at shorter wavelengths [Verleur 68], the chief utility of the foregoing model is the fitting of the index to a polynomial. Suppose, for example, that there is a single dominant oscillator in the UV at a frequency ω_0. It is also assumed that $\omega << \omega_0$ and $\omega << g$. Equation (2-78a) is expanded in a Taylor series

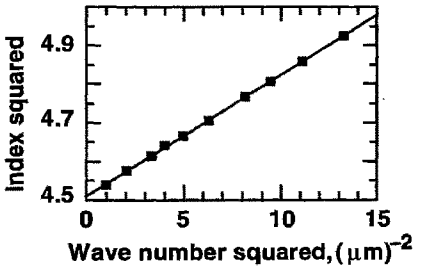


Fig. 2-96 Refractive index squared vs wave number squared [units $(\mu m)^{-2}$] of a thin layer of hafnia deposited by ion plating [see Eq. (2-78c)]. $A = 4.51$ and $B = 0.0312$. Courtesy of Coherent, Inc.

$$n^2 \approx A' + B'\,\omega^2\,\omega_0^{-2} + \ldots \qquad (2\text{-}78b)$$

or in terms of wave number σ

$$n^2 \approx A + B\sigma^2 \ . \qquad (2\text{-}78c)$$

which is variant of Eq. (2-2c). In practice, A and B are determined from a plot of n^2 vs σ^2. Figure 2-96 depicts such a plot for a layer of hafnia. The indices were obtained by depositing a single layer and measuring its transmittance on a spectrophotometer.

A similar expansion may be made in the IR, but in this case ω_0 is at lower frequencies than ω. If it is true that $\omega >> \omega_0$, the contribution of the IR oscillators adds a term to Eq. (2-78c)

$$n^2 \approx A + B\sigma^2 - \frac{C}{\sigma^2 - D} \ . \qquad (2\text{-}78d)$$

2.11.2 Equivalent layers in terms of dimensionless parameters

The starting point is the definition of Λ and Ω in Eqs. (2-47b) and (2-47c). Related parameters are

$$\Upsilon \equiv \frac{1}{2}\left[\frac{n_p}{n_q} + \frac{n_q}{n_p}\right] = \frac{1}{2}(\Omega + \Omega^{-1}) \qquad (2\text{-}79a)$$

$$\Gamma \equiv \frac{1}{2}\left[\frac{n_p}{n_q} - \frac{n_q}{n_p}\right] \qquad (2\text{-}79b)$$

and for the thickness,

$$\zeta \equiv \frac{\Lambda - 1}{\Lambda + 1} \ . \qquad (2\text{-}79c)$$

In terms of the phase thicknesses of the inner and outer layers [Epstein 52], β_e is found from the arccos of

$$\cos\beta_e = (-\Upsilon \sin\beta_q)\, \mathcal{S} + (\cos\beta_q)\, C \tag{2-79d}$$

whence $\sin\beta_e$ may be found and substituted into

$$n_e = n_p\ [(\cos\beta_q)\, \mathcal{S} + (\Upsilon \sin\beta_q)\, C - \Gamma \sin\beta_q]\, (\sin\beta_e)^{-1} \tag{2-79e}$$

where

$$C \equiv \cos(2\beta_p) \qquad \text{and} \qquad \mathcal{S} \equiv \sin(2\beta_p)\ . \tag{2-79f}$$

Another form is

$$\cos\beta_e = 0.5\ [(1 + \Upsilon)\cos\beta_t + (1 - \Upsilon)\cos(\zeta\, \beta_t)] \tag{2-79g}$$

and

$$n_e^2 = n_p^2\, \frac{A - \Gamma \sin\left[\dfrac{\beta_t}{1+\Lambda}\right]}{A + \Gamma \sin\left[\dfrac{\beta_t}{1+\Lambda}\right]} \tag{2-79h}$$

where

$$A \equiv 0.5\ [(1 + \Upsilon)\sin\beta_t + (1 - \Upsilon)\sin(\zeta\, \beta_t)]\ . \tag{2-79i}$$

2.11.2.1 General properties of the equivalent layer

There are general properties of n_e and β_e that may be used as computational shortcuts. Reference is made to β_t that appears in Eq. (2-46g). By inspection of Eqs. (2-79g) and (2-79h), it follows that [Thelen 66b]

$$\beta_e = F_1\ (\Lambda\ ,\ \Omega\ ,\ \beta_t\ ) \tag{2-80a}$$

$$\frac{n_e}{n_p} = F_2\ (\Lambda\ ,\ \Omega\ ,\ \beta_t\ ) \tag{2-80b}$$

where F_1 and F_2 are functions. It follows that β_e is a function of only the total phase thickness, the thickness ratio and index ratio. The same is true of the ratio of n_e to the index of the outer layer. It is concluded that:

(1) Given the thickness ratio Λ, if n_e is computed for an *LHL* type design, then the equivalent index n_e' for an *HLH* design is

$$n_e' = \frac{n_L\ n_H}{n_e}\ . \tag{2-80c}$$

(2) Given the index ratio Ω, there is no simple way to compute n_e if the thickness ratio Λ is changed or replaced by its reciprocal.

(3) For a given value of β_t, β_e is unaltered if either Λ or Ω are replaced by their reciprocals. In mathematical terms

$$\beta_e(\Lambda, \Omega) = \beta_e\ (\Lambda^{-1}, \Omega) = \beta_e(\Lambda, \Omega^{-1}) = \beta_e(\Lambda^{-1}, \Omega^{-1})\ . \tag{2-80d}$$

The equivalent layer whose design is captioned in Fig. 2-74 illustrates the foregoing theorems. When β_t is 130°, its n_e is 4.03 and from Eq. (2-80c) n_e of 0.667**H** 0.667**L** 0.667**H** is

$$n_e' = \frac{1.85 \times 3.66}{4.03} = 1.68\ . \tag{2-80e}$$

β_e (which is 144°) is the same for both designs. Table 5-1 lists the permutations of the 2:1 and 1:2 stacks that satisfy Eq. (2-80d): 0.67**H** 0.667**L** 0.67**H**, 0.67**L** 0.67**H** 0.67**L**, 0.33**L** 1.33**H** 0.33**L** and 0.33**H** 1.33**L** 0.33**H**. For a given β_t , each of the four foregoing designs manifests the same β_e.

2.11.3 Synthesis of an equivalent layer

An equivalent layer with a prescribed index and phase thickness is synthesized [Ohmer 78]. Eq. (2-79e) is obtained from the matrix element m_{21}. Rewriting it and its companion for m_{12} produces

$$n_e\, n_p^{-1} \sin\beta_e + \Gamma \sin\beta_q = \cos\beta_q\, \mathcal{S} + \Upsilon \sin\beta_q\, C \qquad (2\text{-}81a)$$

$$n_e^{-1}\, n_p \sin\beta_e - \Gamma \sin\beta_q = \cos\beta_q\, \mathcal{S} + \Upsilon \sin\beta_q\, C \qquad (2\text{-}81b)$$

where C and $\mathcal{S}$ are defined in Eq. (2-79f). The common right of Eq. 2-81a and Eq. (2-81b) is eliminated, which results in an expression for $\sin\beta_q$

$$\sin\beta_q = w \sin\beta_e \qquad (2\text{-}81c)$$

where

$$w \equiv \left[\frac{n_p}{n_e} - \frac{n_e}{n_p}\right]\left[\frac{n_p}{n_q} - \frac{n_q}{n_p}\right]^{-1} . \qquad (2\text{-}81d)$$

2.11.3.1 Procedure for finding the layer thicknesses

The following steps may be programmed on hand-held calculators or larger computers:

(1) Compute the right of Eq. (2-81c).

(2) If it exceeds 1.0, the equivalent index does not exist.

(3) If β_q is positive, try the solutions: β_q and $180° - \beta_q$. These are the possible solutions: *solution one* and *solution two.*

(4) If β_q is negative, try the solutions $360° - |\beta_q|$ and $180° + |\beta_q|$. These are the possible solutions: *solution one* and *solution two.*

(5) Given β_q and $\sin\beta_q$, compute $\cos\beta_q$. Evaluate the left of Eq. (2-81b) and Eq. (2-79d) as well as the coefficients of C and $\mathcal{S}$.

(6) Solve Eqs. (2-81b) and (2-79d) simultaneously for C and $\mathcal{S}$. Obtain β_p from

$$\beta_p = 0.5 \arctan(\mathcal{S}/C) . \qquad (2\text{-}81e)$$

β_p should always be located in the correct quadrant.

(7) The equivalent layer does not exist if the linear equations in step (6) are singular or if either $|C|$ or $|\mathcal{S}|$ exceeds 1.0.

(8) At a design wavelength λ_0, convert the phase thicknesses into quarterwave optical thickness QWOT or optical thicknesses OT by

$$\text{QWOT}]_{outer} = \beta_p\, \lambda_0/90 \qquad \text{OT}]_{outer} = \beta_p\, \lambda_0/360 \qquad (2\text{-}81f)$$

and

$$\text{QWOT}]_{inner} = \beta_q\, \lambda_0/90 \qquad \text{OT}]_{inner} = \beta_q\, \lambda_0/360 . \qquad (2\text{-}81g)$$

where the QWOT and OT have the physical units as λ_0.When $\beta_e = 90°$, $n_L < n_e < n_H$, where n_H and n_L refer to the high index and low index components of the equivalent layer. The foregoing obtains regardless of whether the design is *LHL* or *HLH.* The proof follows from Eq. (2-81c). Because $sin\beta_q$ is less than 1.0 in magnitude, the same must be true of w when β_e is 90°. An illustration is that when the denominator of Eq. (2-81c) is positive, the design is *HLH.* For w to be less than 1.00, then

$$\frac{n_p}{n_e} < \frac{n_p}{n_q} \qquad (2\text{-}81h)$$

and it follows that n_e must exceed n_q, which is the low index component. A similar argument leads to the conclusion that n_H must exceed n_e.

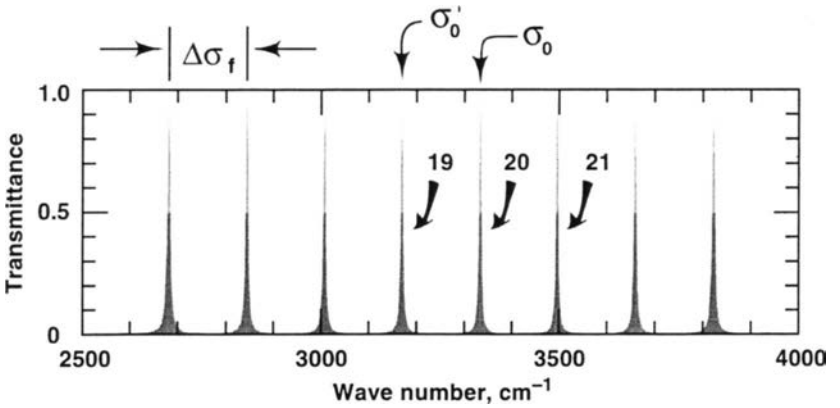


Fig. 2-97 Transmittance vs wave number of the single-cavity bandpass: *air* **H L H' L H M** *Ge*, where the refractive indices of *Ge*, **L**, **M** and **H** are 4.00, 1.35, 1.90 and 4.1, respectively. The optical thickness of **L**, **M** and **H** is $\lambda_0/4$ at $\lambda_0 = 3.00$ μm. The optical thickness of **H'** is 30 μm. $\Delta\sigma_f$ is the separation between orders. The 19, 20 and 21 are the order numbers of the respective passbands.

2.11.4 Spectral bandwidth of a single-cavity bandpass

The transmittance of a generic bandpass filter is depicted in Fig. 1-85. Figures 2-97 and 7-8 show the transmittances of single-cavity (i.e., Fabry-Perot) filters. This filter is the focus of this section, in which equations are derived for: (1) the wavelength of the passband center; (2) passband spectral width; and (3) span of wavelength between transmission peaks.

2.11.4.1 Spectral location of the passband center

A maximum of a given passband occurs at a wavelength λ_0, which is equivalent to a wave number σ_0. Maxima of T in Eq. (2-45b) occur when

$$m\,\pi = \beta'' \equiv 2\,\pi\,\sigma_0\,n_c\,h_c\cos\theta_m - \delta_m(\sigma_0) \qquad (2\text{-}82a)$$

where the integer m is the *order number*. The foregoing is usually a transcendental equation because the dependence of the phase shift δ upon σ is a polynomial in σ.

2.11.4.1.1 Order number

Equation (2-82a) is rewritten in terms of wavelength at normal incidence for a wavelength λ_m, which corresponds to the mth order resonance

$$m\,\pi = 2\,\pi\,n_c\,h_c/\lambda_m - \delta_m \qquad (2\text{-}82b)$$

In the special case where the δ_m is either zero or 180°, this simplifies to

$$n_c\,h_c = m\,\tfrac{1}{2}\,\lambda_m\ . \qquad (2\text{-}82c)$$

In other words, an integral number of halfwaves "fit" into the optical thickness of the spacer, as illustrated in Fig. 2-98.

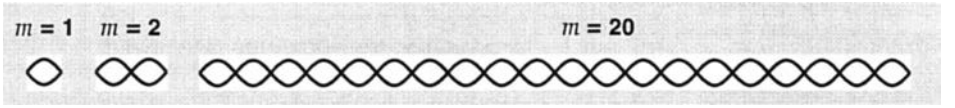


Fig. 2-98 Illustration of the order number m as the number of halfwaves that "fit" into the optical thickness of each of the cavities (*nonshaded*).

2.11.4.1.2 Phase shift upon reflection and order number

§2.3.2 discusses the phase shift upon reflection in terms of the node of the standing wave at a reflecting surface. The same concepts apply to resonance in a cavity. Consider the situation in which Eq. (2-82b) is rewritten for order number 1 and λ_a.

$$1 = 2\,n_c\,h_c\,/\lambda_a - [\delta(\lambda)]/\pi \qquad (2\text{-}82d)$$

where δ is the mean phase shift upon reflection of the reflectors that surround the cavity. Consider first the case when δ (= δ_a) is 360° (or 0°). A wavelength λ_a "fits" into the cavity, as shown in the left panel of Fig. 2-99.

Equation (2-82d) is rewritten when the phase is 325°, which is equivalent to –35°.

$$1 = 2\,n_c\,h_c\,/\lambda_b - (-70°)/180° = 2\,n_c h_c\,/\lambda_b + 0.194 \qquad (2\text{-}82e)$$

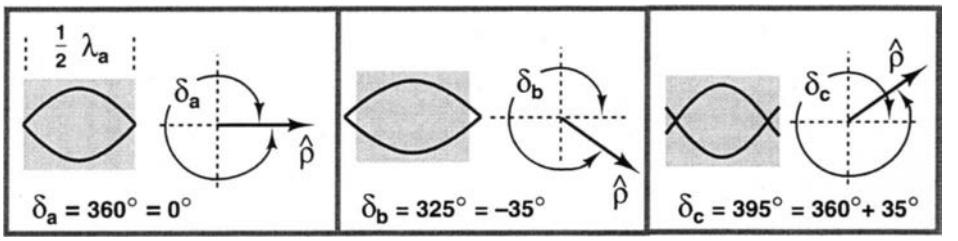


Fig. 2-99 For first-order interference in a cavity of fixed length, the standing wave electric fields when the phase shift upon reflection of the reflectors is 0°, (*left* panel), 325° (*center* panel) and 35° (*right* panel).

Because the left of the foregoing equation is a constant, the term "$2\,n_c\,h_c\,/\lambda_b$" must be smaller because of the addition of the positive 0.194. The conclusion is that λ_b is larger than λ_a. As shown in the center panel of Fig. 2-99, the half wavelength exceeds the width of the cavity. This case occurs when the cavity wall are semitransparent metal films, as depicted in Fig. 8-14.

Consider the case when δ_c is 395°. which is equivalent to +35°. By arguments similar to those in the foregoing paragraph, the wavelength λ_c is smaller than λ_a, as drawn in the right panel of Fig. 2-99.

2.11.4.1.3 Computation of the order number

An example is to compute the order number of the transmittance peak at $\sigma_0 = 3333$ cm^{-1} ($\lambda_0 = 3$ μm) in Fig. 2-97 (*solid* curve). The caption to Fig. 2-97 lists the spacer optical thickness as 30 μm = 30×10^{-6} m. The order number m is

$$m = 2\,n_c\,h_c\,/\lambda_0 = 2 \times (30\times10^{-6}\text{ m})/(3\ 10^{-6}\text{ m}) = 20\ , \qquad (2\text{-}82f)$$

as depicted in the right panel of Fig. 2-98. The wavelength of the transmittance spike at λ_0' (which is equivalent to a wave number ϕ_0 in Fig. 2-97) is at order 19. Neglecting the influence of the phase shift upon reflection, that wavelength is [see Eq. (2-82c)]

$$\lambda_0' = 2\, n_c h_c / 19 = 2 \times (30\ \mu m)/19 = 3.158\ \mu m \qquad (2\text{-}83a)$$

which translates to $\sigma_0' = 3167\ cm^{-1}$. In actuality, the foregoing wavelength is altered slightly by the phase shift upon reflection (mentioned in §2.11.4.1.2).

2.11.4.2 Spectral bandwidth

The goal is to find $d\sigma$, which is the full width of the passband at the $T = 50\%$ level—the semiwidth. The starting point is Eq. (2-45e). As depicted in Fig. 2-100, the transmittance is 50% when $\mathcal{F} \sin^2\beta''$ in Eq. (2-45b) is 1.00, at which point the phase thickness on the left of Eq. (2-82a) increases to $\beta'' + \Delta\beta$ at a wave number of $\sigma_0 + d\sigma$ (where $d\sigma$ is one-half of the width at which T = 50%)

$$\beta'' + \Delta\beta \equiv 2\pi\,(\sigma_0 + d\sigma)\, n_c h_c \cos\theta_m - \delta_m(\sigma) \qquad (2\text{-}83b)$$

where $\Delta\beta$ is one-half of the semiwidth in radian units. It follows that at $T = 50\%$

$$\sin(m\pi + \Delta\beta) = \sin\Delta\beta = (\mathcal{F})^{-1/2}\ . \qquad (2\text{-}83c)$$

When $\mathcal{F}$ is large—30 or greater—the arcsin may be approximated by its argument

$$\Delta\beta \approx (\mathcal{F})^{-1/2}\ . \qquad (2\text{-}83d)$$

The phase shift upon reflection $\delta_m(\sigma)$ may be expanded in a Taylor series about σ_0

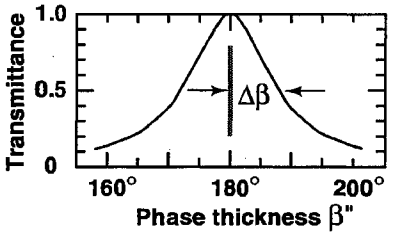


Fig. 2-100 Transmittance vs phase thickness (in degrees) β'' of the spacer layer of a single-cavity bandpass. $\Delta\beta$ is one-half of the width at which $T = 50\%$.

$$\delta_m(\sigma) \approx \delta_m(\sigma_0) + \frac{d\delta_m}{d\sigma} d\sigma + \tfrac{1}{2} \frac{d^2\delta_m}{d\sigma^2} (d\sigma)^2 + \ldots\ . \qquad (2\text{-}83e)$$

The second-order term is neglected and $\delta_m(\sigma)$ is substituted into Eq. (2-83b), along with $\Delta\beta$ from Eq. (2-83d)

$$m\pi + (\mathcal{F})^{-1/2} \approx 2\pi\,(\sigma_0 + d\sigma)\, n_c h_c \cos\theta_m - \delta_m(\sigma_0) - \frac{d\delta_m}{d\sigma} d\sigma\ . \qquad (2\text{-}83f)$$

$d\sigma$ is factored from the foregoing

$$d\sigma\,(2\pi n_c h_c \cos\theta_m - \frac{d\delta_m}{d\sigma}) \approx (\mathcal{F})^{-1/2}\ . \qquad (2\text{-}83g)$$

The fractional bandwidth [see Eq. (1-9)] is $(2d\sigma)/\sigma_0$. From Eq. (2-45d) the reciprocal of the fractional bandwidth Q [see Eq. (1-10)] is

$$Q \approx (2\pi n_c h_c - \frac{d\delta_m}{d\sigma})\, \sigma_0 \sqrt{R_m}\, (1 - R_m)^{-1}\ . \qquad (2\text{-}83h)$$

at normal incidence. Q may be increased by: (1) a thickening of the spacer—which is equivalent to using a larger order number; (2) an increase of the magnitude of the phase dispersion (provided its derivative with respect to wave number is negative); and (3) enhancement of the reflectance of the reflectors.

2.11.4.3 Separation between orders

The separation between orders $\Delta\sigma_f$ (see Fig. 2-97) is found by differentiating Eq. (2-82a) with respect to m. The phase shift is assumed to be constant—to a first approximation. An example is the Fabry-Perot captioned in Fig. 2-97—its optical thickness is 30 μm, which is 0.003 cm.

$$\Delta\sigma_f = (2\, n_c\, h_c)^{-1} = (2 \times 0.003)^{-1} = 166.7 \text{ cm}^{-1} \ . \qquad (2\text{-}83i)$$

The advantage of using wave number is that the transmission peaks are equally spaced in wave number, as shown in Fig. 2-97.

Equations (2-83h) and (2-83i) show that the achievement of a narrow bandwidth of a filter by using a relatively thick (i.e., high order) spacer may introduce other problems, such as the unwanted passbands that crowd in closely to it.

2.11.5 Reflection coefficients in terms of sheet resistance

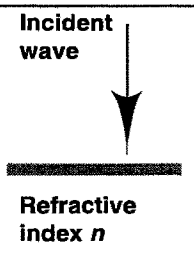


Fig. 2-101 A wave is incident upon a thin, conducting lamina (*shaded*). The refractive index of the emergent medium is n.

There are useful equations in terms of sheet resistance for the radiant reflectance and transmittance at an interface of a thin conductive layer [Hoffer 75]. A caveat is that they apply only to the far infrared spectral region. As shown in Fig. 2-101, the conductive lamina is bounded on its upper surface by an incident medium of refractive index $n_0 = 1.00$ and on its lower surface by an emergent medium of refractive index n. It is assumed that the metric thickness of the lamina is smaller than the skin depth of the radiation and also is much smaller than the wavelength of the radiation. The transmittance T is

$$T \approx \frac{4\, n}{(n + 1 + \mathcal{Z}_0/\mathcal{R})^2} \qquad (2\text{-}84a)$$

$$R \approx \frac{(n - 1 + \mathcal{Z}_0/\mathcal{R})^2}{(n + 1 + \mathcal{Z}_0/\mathcal{R})^2} \qquad (2\text{-}84b)$$

where $\mathcal{R}$ is the sheet resistance of the lamina [Eq. (1-23), units of ohm/square] and $\mathcal{Z}_0$ (377 ohm) is the impedance of free space [Eq. (2-15e)].

As an example of the questionable use of this equation, suppose that an indium-tin oxide conductive sheet is deposited upon a substrate of refractive index 1.50. Its sheet resistance is 20 ohm/square. Equation (2-84b) shows that its approximate reflectance is

$$R \approx \frac{(1.5 - 1 + 377/20)^2}{(1.5 + 1 + 377/20)^2} = 0.82 \qquad (2\text{-}84c)$$

which, according to the author's experience, is within one percent of the correct value.

Chapter 3
Thin films, the building blocks of multilayers

3.0 Purpose of this chapter

The basic theory of multilayer design appears in Chap. 2. Not a few questions remain unanswered, such as: (1) What range of refractive indices may be used in multilayers? (2) Are there lower and upper limits on the thicknesses of coatings? (3) What is the influence of the dispersion of the optical constants on the performance of the coatings? These and other questions are addressed in this chapter.

The range of refractive indices depends upon the method used to deposit the films and, consequently, this chapter begins with a survey of deposition methods. The focus, however, is on the optical and other properties of films per se. Although mention is made of factors that might influence these properties, such as rate of deposition or substrate temperature, a discussion of the control of these parameters is deferred until Chap. 9.

Table 3-1 Methods of depositing optical thin films

Name	Class	Brief description of process	Comment
Evaporation in a vacuum	PVD	A liquid or solid evaporates and its vapor condenses upon the substrate.	The most widely used method of depositing thin films for non-architectural purposes.
Sputtering	PVD	Sputtered material (in the form of a vapor) leaves the cathode and condenses upon the substrate.	DC magnetron sputtering is used extensively to deposit thin films on glass for architectural purposes.
CVD*	Chem	A vaporous chemical compound reacts chemically to produce a solid film upon a substrate.	Curved surfaces and the interior of cylinders are easily coated uniformly.
PCVD§	PVD, Chem	The CVD process (above) is enhanced by a plasma.	Since 1993, extensively used to produce cold mirrors.
Chemical solution	Chem	A solution containing a chemical compound coats a surface — the compound remains on the surface after the solvent evaporates.	Used to deposit antireflection coatings on face plates for VDTs.
Electro-deposition	Chem	An electrostatic force causes an ion to migrate through a liquid and to impinge upon a substrate.	The substrate must be electrically conducting.

*Chemical Vapor Deposition §Plasma Chemical Vapor Deposition
The *Class* column classifies the coating process as Physical Vapor Deposition (PVD) and/or chemical deposition (Chem).

3.1 Methods of depositing a thin film — a conceptual view

Two sections are devoted to thin film deposition. §3.1 is conceptional and relatively short. §3.2 reviews deposition methods in vogue as this is written in the year 2000.

3.1.1 Transport and attachment

Each of the deposition methods appearing in Table 3-1 involves at least two important steps: (1) Material *transport*; and (2) *attachment* of the material to the substrate. The *transport* refers to the mechanism that conveys the thin film material to the receiving surface. The *attachment* is the physical and/or chemical process that causes the film material to remain on that surface.

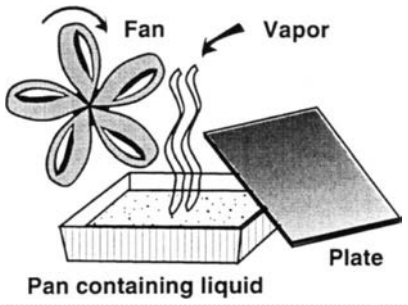


Fig. 3-1 In an analogy to the CVD process, a fan blows a stream of air over a pan of liquid. The vapor from that liquid is entrained in the air and is chem-adsorbed on the plate.

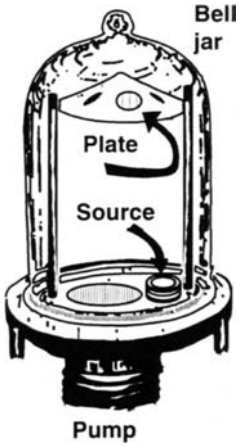


Fig. 3-2 The *plate* (in the rack at the top of the chamber) is coated with a solid layer by physical vapor deposition as the vapor emitted by the electrically heated evaporation *source* condenses upon it. The *bell jar* is evacuated by the *pump*.

By analogy, suppose that there is a field of exposed dirt that is loose and devoid of plant life. Straw is spread upon the soil in an effort to stabilize it. The straw is *transported* to the site and then dumped upon the soil. But, alas, the straw may be blown away by winds that buffet the site. This is prevented by anchoring the straw in place, as, for example, by covering it with netting. The former step involves the *transport* and the latter is the *attachment*.

3.1.1.1 Example [analogous to physical vapor deposition (PVD)]

Physical vapor deposition PVD is illustrated by a simple example in which solid water (i.e., ice) is deposited upon a plate. A refrigerated plate is positioned above a pot of boiling water. The *transport* occurs as the water vapor rises (convectively) from the pot. It moves to the plate, where it condenses.

But why should the ice remain on the plate? The *attachment* (mentioned in §3.1.1) of the ice is related to the heat of sublimation that the water vapor surrenders when it turns to ice. As long as the plate is cold and can absorb this heat, the ice will remain on the plate.

3.1.1.2 Example [analogous to chemical vapor deposition (CVD)]

Figure 3-1 illustrates chemical vapor deposition CVD. A fan blows air over a pan of liquid. The vapor from the liquid is entrained in the moving air and the wind *transports* the vapor to the plate. The plate is composed of a material that induces the vapor to adhere to it by chem-adsorption. The latter is the *attachment* process.

3.2 Deposition methods — hardware and procedures

Reference is made to Table 3-1, which classes the deposition method as Physical Vapor Deposition (PVD) or Chemical (Chem). This survey of deposition methods is abbreviated in the extreme. Further details appear in reviews such as those of Maissel [70] and Bunshah [94]. Recommended is the text of Ohring [92], which presents a scholarly and balanced overview of most of the processes cited below.

3.2.1 Physical Vapor Deposition

Evaporation in a vacuum and sputtering are the two widely used deposition methods of the class.

3.2.1.1 Evaporation

§3.1.1.1 makes the analogy between boiling water and physical vapor deposition PVD. The "boiling" analogy is complete, except that the process takes place in a vacuum. Figure 3-2 depicts a glass bell jar in which a layer is deposited on a plate by evaporation in a vacuum.

The procedure is: The plate is placed in the rack at the top of the bell jar, which is then sealed. After the pump removes the ambient air in the jar, the evaporation source is heated — usually electrically. Regardless of whether the heated evaporant melts or sublimes, the vapor that emerges from it impinges upon the plate — subliming to form a solid film. The aforementioned PVD process is the focus of Chap. 9.

At the time of this writing in the year 2000, the chief advantages of using vacuum evaporation are that the equipment, evaporants and personnel trained to use the equipment, are readily available. Also, a relatively small amount of material may be evaporated — this is important when costly materials, such as rhodium, gold or scandia, are deposited. Fluoride compounds, as well as oxides, are easily evaporated.



Fig. 3-3 The momentum transfer in the sputtering process is illustrated by the impact of a mallet upon a pivoted beam. The momentum imparted to the block of ice resting upon the beam causes it to be ejected upwards.

A disadvantage of vacuum evaporation is that the substrates are usually heated to induce mechanical hardness in the deposited films. Substrates, such as plastics, simply will not withstand this harsh environment. Although possible, it is expensive and difficult to evaporate downwards or sideways. This upwards evaporation requires that expensive tooling must be manufactured to hold the parts to be coated.

3.2.1.2 Sputtering

Sputtering is a physical process in which energetic ions impact a solid surface (called the target, or cathode) — thus dislodging molecules and/or atoms from that surface. Although the sputtered molecules/atoms are in the vapor state, sputtering is not an evaporation process, as described in §3.2.1.1.

The important fact is that a bombarding ion imparts momentum to an atom in the target. Thus sputtering is viewed as a momentum transfer process. Egregious liberties are taken in illustrating this transfer in Fig. 3-3. The impact of a mallet transfers

momentum to a pivoted beam and thence to a block of ice resting upon the beam. The result is that the ice bounds upwards into the air.

By analogy, the aforementioned momentum is imparted by the impact of an energetic ion upon the target, which in turn may dislodge one (or possibly more) molecule(s)/atom(s) from the solid. That sputtered material subsequently condenses upon the substrate, as in §3.2.2.1.

Consider next the *attachment* process. Similar to §3.1.2.1, the sputtered material surrenders its energy to the surface upon which it impinges — and, for that reason, remains upon the surface. There are instances in which the sputtered atom may chemically react at the surface, as, for example, when silicon is sputtered and then exposed to oxygen to form silica. Chapter 9 discusses in more detail ion beam sputtering and reactive planar magnetron sputtering.

One of the chief advantages of planar magnetron sputtering is that high deposition rates may be obtained — especially of metals. For this reason, an area of many square kilometers of architectural glass is sputter coated annually. This glass coating is aided by the fact that it is easy to sputter downwards.

Another important advantage is the energy with which the sputtered material arrives at the substrate. The sputtered molecule/atom arrives at the substrate with an energy at least 10 times larger than that of an atom/molecule produced by the vacuum evaporation. As mentioned in §3.4.1, this energetic arrival engenders films that are physically more dense and adhere more tenaciously to the substrate. Also, it is relatively easy to sputter mixtures of coating materials, either by sputtering a target composed of a mixture or by co-sputtering two materials.

As compared to the vacuum evaporation process, the sputtering process is quite stable. That is, the deposition in a sputtering system usually proceeds quite well without human intervention. In contrast, in most instances electron beam evaporation must be constantly monitored by an operator to ensure that events such as "hole drilling" of the electron beam into a solid evaporant do not occur.

Layers of the metal salts of fluorine are useful components of multilayer stacks because of their relatively low refractive index and their transparency in the UV. A disadvantage of sputtering is that it is difficult to sputter fluoride compounds. Another problem with sputtering arises when a new material is to be sputtered. A solid target of that material must be prepared, which is usually quite expensive.

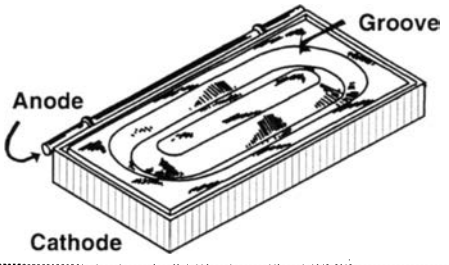


Fig. 3-4 Showing the anode and cathode of a planar magnetron sputter source. The sputtering has etched a groove in the cathode.

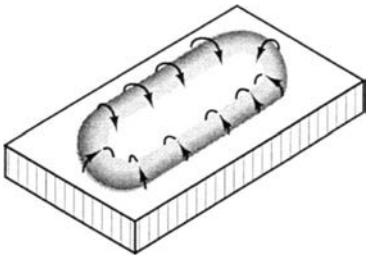


Fig. 3-5 Planar magnetron source (cathode), showing the plasma (*shaded*) and the directions of the magnetic field lines.

3.2.1.2.1 Planar magnetron sputtering

Figures 3-4 and 3-5 illustrate elements of a planar magnetron sputter source [Herrmann 95]. When the ambient gas pressure is sufficiently low (typically 1 mTorr or 0.13 Pa), the anode (see Fig. 3-4) is energized at 600 V. A plasma — a mixture of nearly equal numbers of ions and electrons — forms at the cathode surface. Magnets beneath the cathode produce the magnetic field lines shown in Fig. 3-5. The magnetic field confines this plasma in that region, which is where the sputtering takes place. Sputtering occurs when the cathode is bombarded by ions from the plasma. Figure 3-4 illustrates that the erosion of the sputtered material from the cathode eventually forms a groove in it.

As compared to other types of sputtering, a principal advantage of planar magnetron is the high rate at which most metals may be sputtered. It is a "cold" process because the high thermal conductivity of the metal cathode enables it to be water cooled. The pressure in the vacuum chamber is usually several orders of magnitude lower than the pressure during conventional dc diode or rf diode sputtering.

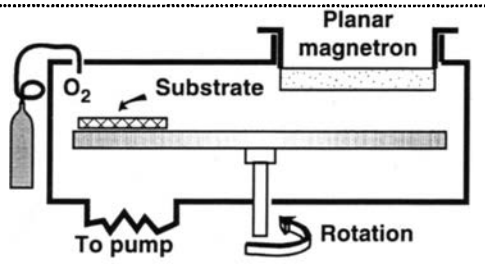


Fig. 3-6 Hartsough [83] deposited an alumina film upon a substrate residing upon a rotating table. As the table rotates, a thin layer of aluminum is magnetron sputtered and then oxidized.

3.2.1.2.2 Reactive sputtering

The apparatus of Hartsough [83] illustrates that a combination of the CVD and PVD processes may be used to deposit an oxide layer, such as alumina. As depicted in Fig. 3-6, a substrate is placed upon a rotary table. The substrate moves under a planar magnetron sputter source during part of its rotation, where a thin layer of aluminum is deposited upon it. That substrate then moves to another physical part of the chamber where there is a relatively high concentration of oxygen. That thin aluminum layer then completely oxidizes to form alumina.

In this instance, the *transport* involves the transfer of aluminum from the cathode to the substrate by sputtering. There is also a flow of the oxygen from the tank to the substrate. Because the chemical reaction of aluminum to aluminum oxide is exothermic, the *attachment* of the alumina layer to the substrate is guaranteed.

As another example, Pawlewicz [79] deposited optical films of titania and tantala by rf sputtering targets of the pure metals in an atmosphere of argon and oxygen. By adjusting the rf power level, he controlled the grain size of the crystallites, which was related to the laser damage threshold.

3.2.2 Chemical processes

The focus of this section is on chemical deposition.

3.2.2.1 Chemical Vapor Deposition

Although Chemical Vapor Deposition CVD is widely used in the semiconductor industry, its use in depositing optical coatings is in its infancy. Several publications [Rogers 89, Fuller 96] extol its effectiveness in coating a dome-shaped substrate. Review articles such as Feist [69] provide an idea of the great diversity of films that may be deposited by this method.

3.2.2.1.1 CVD deposition of tin oxide

Figure 3-7 depicts a belt furnace that coats glass slabs with a layer of tin oxide. The process occurs at atmospheric pressure. After the substrates are placed on a moving belt and enter the machine, they are heated by a bank of furnace elements — these are not shown in Fig. 3-7. In the center of the machine, there is a reactor where the tin oxide is deposited by the reaction

$$SnCl_4 + 2\,H_2O \rightarrow SnO_2 + 4\,HCl\ .$$

The *transport* occurs when the dry nitrogen carrier gas is bubbled through a liquid bath of tin tetrachloride. The vapor of that compound is entrained in the nitrogen and reaches the reactor. The tin tetrachloride and the water vapor enter the reactor, mix and react to form a solid tin oxide layer on the heated glass. The *attachment* of the tin oxide occurs because it is a solid and the HCl vapor is removed from the system. In actual practice, the process is more complex than described above. Methanol is added to control the speed of the reaction. In some instances fluorine dopants are included to increase the electrical conductivity of the tin oxide.

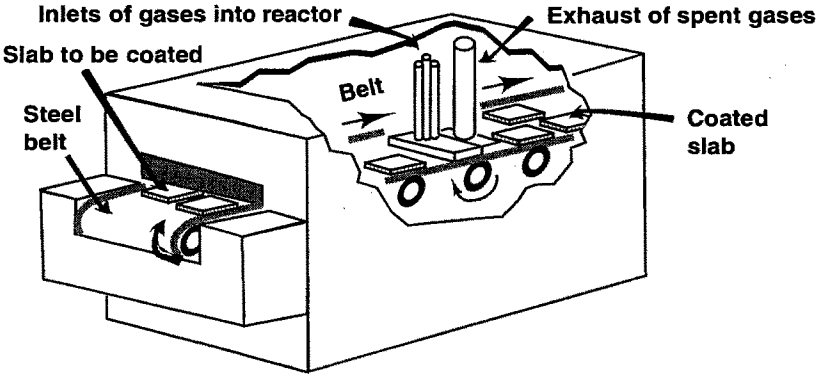


Fig. 3-7 Glass slabs on a moving mesh steel belt are radiantly heated by furnace elements (which are not shown) and subsequently enter the reactor, where they are coated with a layer of tin oxide. After SierraTherm [96].

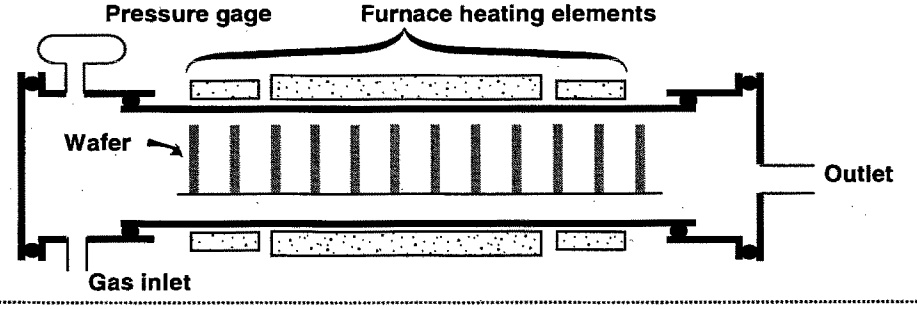


Fig. 3-8 A vessel (shown in cross section) in which relatively thin slabs (i.e., the wafers) are coated by CVD. After Desu [89] and Rogers [89].

3.2.2.1.2 LPCVD deposition

The prefix "LP" denotes "Low Pressure" CVD. The acronym MOCVD (Metal Oxide CVD) is also used. Unlike the atmospheric pressure coating process in §3.2.2.1.1,

LPCVD typically takes place at a pressure in the range 0.1 Torr (13 Pa) to 1 Torr (133 Pa). As shown in Fig. 3-8, the substrates are stacked on edge in a sealed reaction chamber, which is surrounded by a furnace.

The "plumbing" of the system is not shown in Fig. 3-8. Flow controllers meter the reactants into the chamber. Connected to the outlet are a pump and cold trap that remove the unreacted chemicals, as well as the organic and inorganic byproducts from the reaction. There are also additional plumbing elements (such as the bypass line) that have been omitted, for sake of clarity.

In the following example, a silica film is to be first deposited, then followed by a tantala layer. The substrates are placed in the reaction chamber, which is then sealed. Pumps then reduce the pressure to below 10^{-5} Torr (1.33 mPa). The furnace elements heat the substrates to a temperature near 700 °C. The precursor is an organometallic such as tetraethoxysilane, alias tetraethylorthosilicate, alias TEOS (see Fig. 3-9). It is entrained in an inert gas (such as nitrogen) and flows into the heated reactor. There, it decomposes to form a silica layer on the hot surfaces. The use of the precursor di-t-butoxydiacetoxysilane (alias diacetoxydi-t-butoxysilane, alias DADBS) has the advantage that it reacts at a lower temperature of 460 °C [Desu 89].

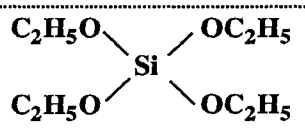


Fig. 3-9 Chemical structure of tetraethoxysilane. After Gelest [96].

The aforementioned process is terminated when the silica layer has reached its prescribed thickness. The reactor vessel and feed lines are then purged of the silica precursor. A precursor such as "tantalum V ethoxide" enters the chamber and a layer of tantala is deposited. It is important that the temperature of the furnace remain stable. A consequence is that the precursors are selected so that they react at approximately the same rate at a given temperature. Pulker [84] lists more than 25 oxide materials that have been produced by the CVD.

3.2.2.1.3 Discussion of CVD

CVD has the following advantages:

- All surfaces of the substrate are coated simultaneously. A coating of uniform thickness is attained if the temperature in the reactor is relatively uniform.
- Films of greater packing density are obtained (see §3.3.2.6.1), which results in layers of higher refractive index than produced by vacuum evaporation.

Disadvantages of CVD are:

- The layers are usually mechanically stressed tensively. In the production of heat reflectors (see Fig. 1-68 and §5.3.5), the stress limits the total coating thickness to the range 4 µm to 5 µm.
- The materials that may be deposited are limited to the available precursors. If a scandia layer is to be deposited, for example, then an organometallic compound containing scandium must be synthesized.
- Due to the organic reactants, carbon may be included in the metal oxide layer.

3.2.2.2 Plasma CVD

The acronyms PECVD (Plasma Enhanced CVD) or PICVD (Plasma Impulse CVD) are often used to identify these processes. Figure 3-10 shows the layout of a system. Figure

3-11 depicts a cross section of the reactors in which the interior of a 50 mm ϕ dome or a flat substrate are coated.

Suppose that alternate layers of titania and silica are to be deposited. The substrate (at a temperature of 90 °C) is inserted into the reaction vessel in Fig. 3-11 and pumps remove the air. Flow controllers (see Fig. 3-10) admit a gaseous mixture of titanium tetrachloride and oxygen (in a ratio of 1:67) at a pressure of approximately 1 Torr (133 Pa). A microwave pulse of 600 μs duration (at a frequency of 2.45 GHz) enters the reactor and forms a plasma. The titanium tetrachloride reacts with the oxygen to form a solid titania layer. The pumps remove the chlorine and other waste gases after the pulse subsides. The physical volume near the substrate fills with a fresh gas mixture and the system is pulsed again. The pulsing (at a repetition rate of 50 Hz) continues for 30 seconds until a QWOT layer 660 nm is laid down.

The precursor for the deposition of a silica layer is hexamethyldisiloxane, the chemical structure of which appears in Fig. 3-12. With each microwave pulse, the precursor is decomposed and combines chemically with the oxygen to form a silica layer. The deposition time is comparable to that of the titania layer. For both materials, the layer thickness is controlled by the number of pulses.

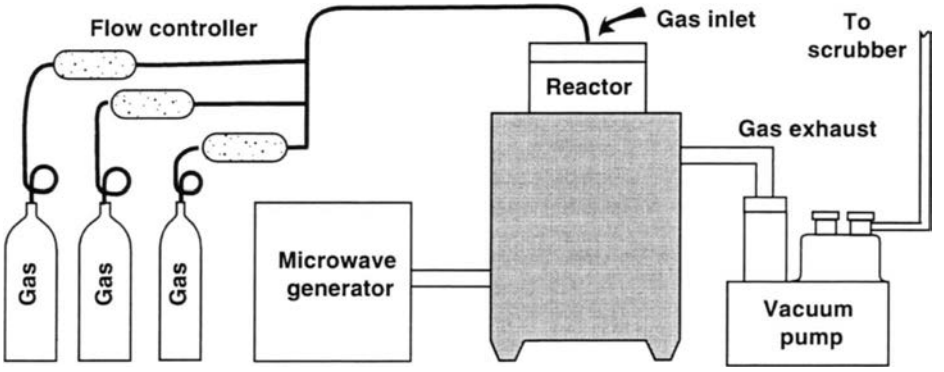


Fig. 3-10 Optical coatings are produced in the reactor by pulsed microwave plasma CVD deposition apparatus. After Heming [92] and Etzkorn [91].

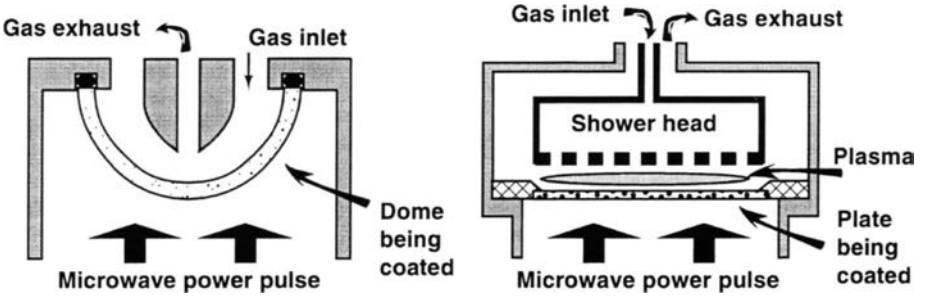


Fig. 3-11 Cross section of a pulsed microwave plasma CVD reactor. Optical coatings are produced on the interior of a glass dome (*left* panel) and on the upper surface of a flat ceramic plate (*right* panel). After Etzkorn [91], Heming [92], and Segner [95].

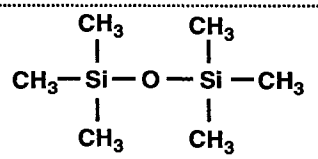


Fig. 3-12 Chemical structure of hexamethyldi-siloxane. After Gelest [96].

At the time of this writing in 2000, tens of millions of cold mirrors (see Fig. 1-60) are coated yearly by the Schott Company, using this process.

3.2.2.2.1 Selection of the frequency

The process is not necessarily limited to the microwave frequency cited above. Both Etzkorn [91] and Bulkin [96] mention the possibility of using a frequency of 13.546 MHz in their processes. George [94] presents a useful survey of both CVD and PCVD processes.

3.2.2.2.2 Overview

One of the advantages of this plasma CVD process is that mechanically hard, dense oxide layers are deposited. Heming [92] documents that there is a quite low shift of the refractive index of the films as they are exposed to water vapor. Another advantage is that the deposition rate is quite fast. A 25 layer cold mirror may be deposited in approximately 10 minutes, as compared with hours in the conventional PVD processes.

As far as the disadvantages of this process, the apparatus in Fig. 3-11 could not be used to coat a metallic substrate, since the substrate must transmit the microwaves. This does not imply, however, that such problems could not be solved by an alternate mechanical design. As in the case of the CVD deposition, the materials that may be coated are limited to the available precursors.

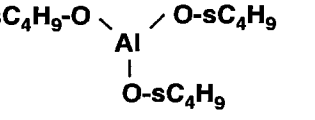


Fig. 3-13 Chemical structure of aluminum III s-butoxide. After Gelest [96].

3.2.2.3 Deposition from a liquid

The slang term describing this method is "wet chemical" to distinguish it from the vapor process CVD mentioned above. Chemical deposition in a liquid is historically the oldest, dating to the 1800s when metallic silver layers were deposited by the reduction of a silver salt. A review article [Schwartz 94] calls this process "electrodeless deposition." In the following two examples, the goal is to deposit oxide layers.

3.2.2.3.1 Deposition by spinning

Suppose that a layer of alumina is to be deposited upon a slab of glass. The first step is to select an organometallic compound containing aluminum, such as aluminum III s-butoxide, whose chemical structure appears in Fig. 3-13. A solution is created by dissolving the compound in an alcohol, such as n-butanol.



Fig. 3-14 The centrifugal force imparted to the liquid on the spinning disk causes the liquid to spread into a thin sheet.

The next step is to apply that solution to a glass surface. Figure 3-14 shows a substrate mounted on a rotating disk. A few drops of solution splatter upon its surface. Centrifugal force causes the liquid to spread and form a thin

layer. The alcohol then evaporates. Baking the glass at temperatures near 400 °C removes the organic components — an alumina layer remains.

Spinning is a way of depositing layers on relatively small surfaces that are nearly flat. The dipping method, described below, allows sheets of glass — as large as 2 m × 3 m — to be coated.

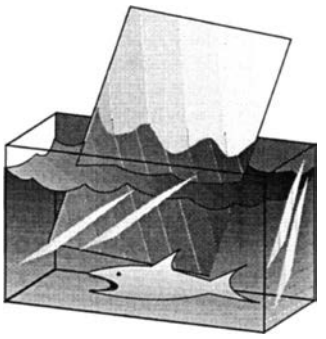


Fig. 3-15 In the first step of coating a glass sheet with an oxide layer, it is immersed in a tank containing a solution of an organometallic compound. A liquid layer of the solution adheres to the surface of the sheet as it is slowly withdrawn.
The surface waves on the liquid and the marine creature are purely figments of the artist's imagination, which, admittedly, is quite fertile.

3.2.2.3.2 Deposition by dipping

This coating method is described by Schröder [69]. Suppose that a glass sheet is to be coated with a layer of titania. The sheet is immersed in a dipping tank filled with alcohol (see Fig. 3-15). Dissolved in that liquid is the organometallic compound titanium dichloride diethoxide, whose chemical structure appears in Fig. 3-16. The alcohol evaporates as the sheet is withdrawn from the solution. The humidity causes a partly hydrolyzed layer of the organometallic to remain on the surface. The thickness of the titania layer is determined by the rate of withdrawal and the viscosity of the liquid.

After a "hard bake" at temperatures approaching 450 °C, a titania layer forms on the surface. The Schott Company (Mainz, Germany) has used this process for several decades. Under license from Schott, the Denton Vacuum Company in the United States has coated panels for computer displays.

3.2.2.3.3 Comparison of processes

It is tempting to believe that the apparatus in which the "wet chemical" deposition occurs is much simpler than the PVD apparatus described in §3.2.1 and, therefore, easier to use. In fact, the humidity in the room where the dipping process takes place must be controlled with the utmost precision. There is also the possibility that particles of the partially hydrolyzed organometallic will form in the tank and will emerge as cosmetic defects on the coated sheet. Another disadvantage of this method is evident in the development process — when multiple layers are to be deposited. Due to the time required to bake the coating — before more layers may be deposited — the development cycle of a new coating may extend into weeks.

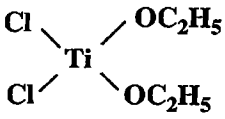


Fig. 3-16 Chemical structure of titanium dichloride diethoxide. After Gelest [96].

An advantage of the dipping deposition method is that both sides of a sheet are coated simultaneously. It is also possible to coat cylinders. Another advantage is that it is easy to deposit layers of refractory materials, such as thoria or tungsten. Layers of intermediate refractive index are obtained by dissolving mixtures of organometallic compounds in the dipping tank.

3.2.2.3.4 Sol-gel coatings

The alcohol solutions of chemicals similar to the tetraethoxysilane mentioned above are also used to deposit sol-gel antireflection coatings for high power laser applications. Lowdermilk [83] describes a chemical process in which partially hydrolyzed tetramethoxysilane was deposited on the glass by dipping.

3.2.2.3.5 Deposition of fluoride compounds by spraying

The procedures in §3.2.2.3.1 and §3.2.2.3.2 emphasized the baking of coatings in air to produce oxides that are mechanically hard. Haisma [85] deposited layers of magnesium fluoride by a spray process. The organometallic compound magnesium-trifluoro-acetylacetonate was dissolved in 2-ethoxyethanol. That solution was sprayed on a glass surface at 500 °C. An analysis by X-rays of the magnesium fluoride films thus produced showed that their properties were identical to those produced by vacuum evaporation.

3.2.2.4 Electroplating

Schwartz [94] reviews electroplating. Electroplated gold films [Epner 80] have a substantial IR reflectance and can be deposited on highly curved surfaces that would be quite difficult to coat by PVD. Reactive processes are used to deposit compounds such as oxides, sulfides and phosphides in plating solutions. Black chrome — a porous film composed of mixtures of chromium oxides — is produced in electroplating baths and is widely used as a photothermal energy absorber. An example of "electrodeless plating" is nickel, which is deposited by decomposing nickel sulfate in a sodium phosphate solution [ASM 64]. Schwartz [94] lists at least 10 metals that may be deposited in a similar manner.

3.2.3 Miscellaneous processes

3.2.3.1 Photochemical

A sophisticated method of depositing a metal is to photodecompose an organometallic compound with a UV, visible or IR radiation source. The advantage of using a laser is that it may be focused to a relatively small area. The review of Ehrlich [83] mentions the laser deposition of both metallic and dielectric materials. Eden [92] lists many methods of photochemical deposition.

3.2.3.2 Plasma polymerization

A plasma is created in a chamber by maintaining a high voltage between electrodes at an appropriate gas pressure. The introduction of a monomer, such as tetrafluoromethane, into the plasma produces polymers that condense on substrates adjacent to the plasma.

Several monographs [Shen 76, Hollahan 74a] review this technology. One of the problems is that the polymer contains many unsaturated chemical bonds, which, in turn, cause it to be more chemically reactive.

3.3 Overview of physical vapor deposition and film formation

Many properties of thin films, such as their grain structure and physical density, are related to the processes by which they are formed. These processes are reviewed briefly. Books such as Lewis [78] and Matthews [75] discuss thin film formation in greater detail.

3.3.1 Conditions for thin film deposition

Figure 3-2 depicts a rudimentary apparatus in which a solid film of a chemical element, as, for example, zinc may be condensed on a substrate by evaporation — i.e., physical vapor deposition. The essential components are: (a) a hermetically sealed chamber from which the gases can be evacuated; (b) an oven containing the zinc; and (c) the substrate, which is the plate in Fig. 3-2.

Two conditions must be met for a film to form:
(1) The residual gas pressure in the chamber must be sufficiently low;
(2) supersaturation of the vapor occurs.

Each is discussed below.

3.3.1.1 Residual gas pressure

The residual gas pressure must be sufficiently low that the zinc atoms emerging from the oven make few collisions, if any, with the residual gases before they condense on the substrate. If many collisions occur, the film is porous and mechanically quite soft. The control of the residual gases is discussed in §9.4.8.

3.3.1.2 Supersaturation of the vapor

Supersaturation relates to the relative rate at which the atoms from the oven impinge on the substrate and the rate at which they re-evaporate from the substrate. If $\mathcal{N}_q$ is the arrival rate (in units of areal density per unit time) and $\mathcal{N}_s$ the re-evaporation rate, then the net rate of film formation $\mathcal{N}$ is

$$\mathcal{N} = \mathcal{N}_q - \mathcal{N}_s \ . \tag{3-1}$$

Supersaturation occurs if $\mathcal{N}$ is positive. In other words, the atoms arrive faster than they re-evaporate. The arrival rate depends upon parameters such as the oven temperature and its physical proximity to the substrate. The re-evaporation rate is a more complex phenomenon and depends upon many parameters, such as the substrate temperature as well as the chemical and physical properties of the surface on which the atoms are condensing. This is related to the sticking coefficient (see §9.5.1.1).

Given that the evaporant is impinging upon the surface, the next step is to examine the mechanisms of film formation.

3.3.2 Steps in film formation

3.3.2.1 Introduction

Often the optical properties of a film such as its refractive index and scatter are directly related to its physical properties, such as its crystalline structure, orientation and size of the crystals. The properties are greatly influenced by the processes, such as nucleation, that occur during its formation. They are briefly reviewed.

3.3.2.2 Condensation and the formation of nuclei

Assume that the conditions in §3.3.1 are met and that vapor is condensing to form a solid film. The impinging atoms lose most — although not all — of their kinetic energy when they land on the surface of the substrate. These energetic atoms — called *adatoms* — have sufficient energy to migrate around the surface, behaving like gas atoms that move in two dimensions, as shown in Fig. 3-17. Several events are possible:

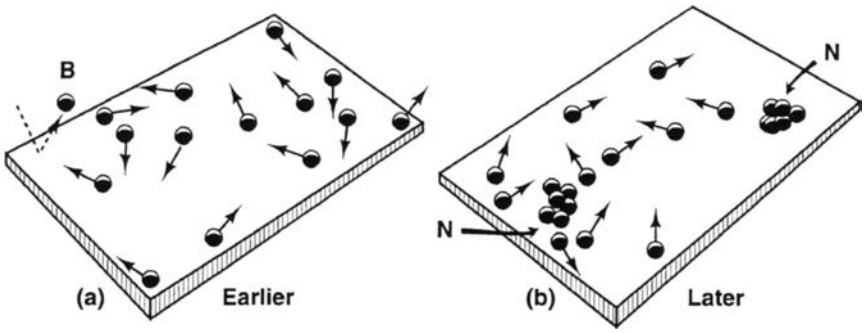


Fig. 3-17 In the formation of a film, atoms of the condensing evaporant and adatoms are represented by spheres. (*a*) Adatoms on the surface move in random directions shown by arrows. At "B," an impinging atom collides elastically with the substrate. (*b*) At a later time, adatoms coalesce to form nuclei *N*.

(1) An adatom may acquire additional energy due to a collision and re-evaporate from the surface.

(2) An adatom may find a site of lower potential energy on the surface, at which point it surrenders its kinetic energy to the substrate (in the form of heat) and remains immobile at the site. This site can take many forms, as, for example, a physical defect or a mote of contamination left from the cleaning process.

(3) A more probable event is that a critical nucleus composed of n atoms will form. Various theories set n in the range from 3 to 6.

3.3.2.3 Formation of island films

At the early stages of growth, the film consists of a myriad of critical nuclei that continue to grow as they attract more atoms. These nuclei are clearly resolvable in electron micrography when they exceed 1 nm in size. These agglomerations are termed *islands*. This progression from nuclei to islands is shown in Figs. 3-18 and 3-19.

As discussed below, the islands usually coalesce into a grain structure if the deposition is continued. But suppose the deposition is terminated when a well-developed island structure exists. Such a film is of interest for several reasons. First, its optical constant is usually quite different from that of a thicker film. An example is that an island film of gold is highly absorbing in the visible portion of the spectrum, in contrast to a thicker film which is quite reflective. A further example is cited in §3.3.2.6.3.

3.3.2.4 Growth of islands and the formation of grains

As shown in Fig. 3-19, the islands continue to grow in both the lateral and vertical directions as the deposition progresses. Eventually, the islands coalesce to form a quasi-continuous film.

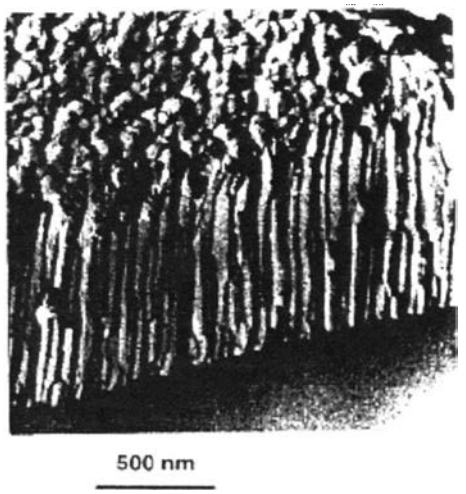
The *quasi* is used in the foregoing sentence because there are voids between the islands. Quite often a structure is observed [Pearson 70, Günther 76] that shows a growth of elongated grains or columns normal to the substrate, as shown in Fig. 3-20.
In fact, Nieuwenhuizen [70] observed that the grains are slanted toward the direction of the impinging vapor. Figure 3-21 depicts an electron micrograph of a cross section of an aluminum film in which the vapor was incident at +45°.



Fig. 3-18 Simulation of the early growth of a gold film, as the nuclei grow to form islands. Its thickness is only a few nm. Based upon the work of



Fig. 3-19 Simulation of the growth of a gold film thicker than that in Fig. 3-18. The islands grow to form the grain structure. Based upon the work of



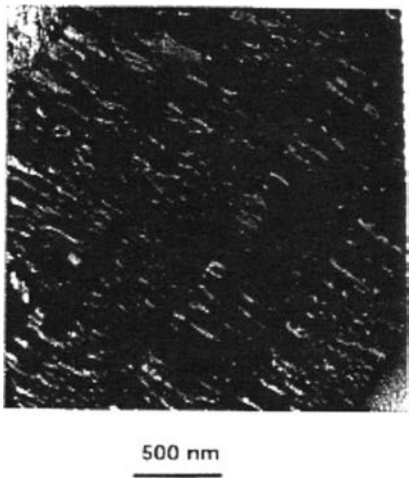


Fig. 3-20 (*Left* panel) Showing the growth of grains in a layer of lead fluoride. After Duparré [96]. (*Right* panel) A cross section of a multilayer with alternate layers of germanium and zinc sulfide. Each photograph is a transmission electron micrograph of a collodion replica of the surface after it was shadowed with platinum.

After exposing the film to air, more aluminum was deposited with the vapor incident upon the substrate at –45°. The change in the direction of the grains is quite evident.

3.3.2.4.1 Modeling of grain growth

Günther [82] modeled grain growth with a "hard sphere" model. Each impinging atom (represented by a sphere) sticks and remains immobile when it hits either the substrate or other frozen atoms. From the illustration in Fig. 3-22, it is evident that the columnar growth — including the voids between the columns — is modeled. Several of the voids in Fig. 3-22 have been highlighted by shading them. Müller [87] allowed the atoms to migrate after contact. The resulting film has a more realistic physical density.

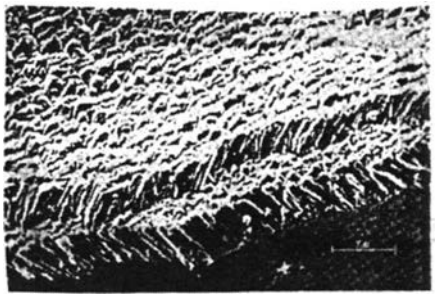


Fig. 3-21 Electron micrograph of a replica of a fractured evaporated aluminum film, where the vapor was incident at an angle +45°. Air was admitted to the chamber and another film was deposited at –45°. After Nieuwenhuizen [70].

Figure 3-23 shows simulated film growth when the vapor is incident at an angle of –60°. The phenomenon of *self-shadowing* causes the directions of the grains to point towards the evaporation source. Self-shadowing is explained as follows: Imagine that atoms have landed on the substrate, been "frozen" into place and have piled on top of each other, thus forming a small protrusion above the substrate. The vapor stream may be regarded as a "wind" and because of the higher angle of vapor impingement, there is a "windward" and "leeward" side of the protrusion. The arriving atoms simply pile up on the "windward" side, leaving the "leeward" side "shadowed."

The influence of the columnar structure on the anisotropy of the refractive index is explored further in §3.4.2.

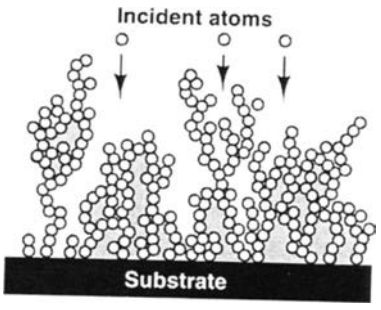


Fig. 3-22 Computer simulation [Günther 82] of grain growth in an evaporated film, in which hard spheres (actually, circles) impinge and remain immobile. The incident atoms are shown by the arrows.

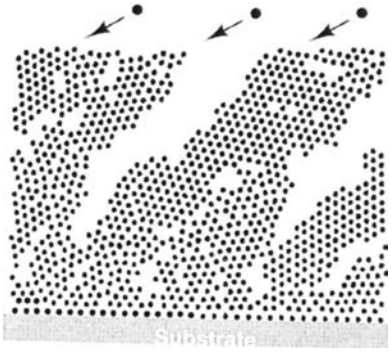


Fig. 3-23 The caption to Fig. 3-22 obtains, with the exception that the incidence angle μ of the vapor (atoms of which are shown by the arrows) is – 60°. After Smith [96].

3.3.2.5 Deposition of compounds

The evaporation of zinc atoms is cited as an example in §3.3.1 because of the simplicity of the process — zinc vapor (i.e., atoms of zinc) condenses into a solid zinc film. A complication occurs when selenium or sulfur is evaporated. The vapor contains molecular species such as S_4 or the dimer S_2.

Glang [70] provides an overview of the deposition of compounds, which may or may not decompose during evaporation. This is illustrated by the experiments of Pulker [69], who examined the species in the vapor as films were deposited from a thermal evaporation source. He found that magnesium fluoride evolved with virtually no decomposition, whereas zinc sulfide decomposed completely into elemental zinc and sulfur.

3.3.2.6 Macroscopic properties of films

§3.3.2.4 mentions the grain growth of films. This section reviews the macroscopic properties of films and §3.3.3.7 considers the influence of grain growth on the properties of films.

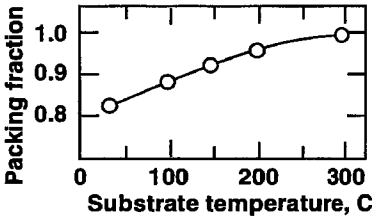


Fig. 3-24 Packing fraction [see Eq. (3-2)] of a magnesium fluoride film vs its substrate temperature during deposition. After Ogura [75].

3.3.2.6.1 Packing density

This parameter, which is also called the packing fraction, relates to the physical density of a film, which is often 10% below that of the bulk material. Insight into the phenomena that cause this lower density is provided by the "hard circle" model in Fig. 3-22. The voids in the columnar growth are evident.

The packing fraction f_p is explained in reference to a layer of magnesium fluoride. This parameter is an indicator of the percentage of the volume of the film filled with atoms of magnesium or fluorine, as compared with the volume of the "voids." A working definition of f_p is to ratio the film density to the density of a bulk single crystal

$$f_p \equiv \frac{\text{density of film}}{\text{density of bulk}} \cdot \tag{3-2}$$

The packing fraction (alias packing density) is illustrated by the work of Pulker [71], who showed that the f_p of thermally evaporated layers of magnesium fluoride layers was 0.9. The data of Ogura [75] in Fig. 3-24 illustrates that the packing fraction of magnesium fluoride increases with increasing substrate temperature during deposition. At 300 °C, the film physical density approached that of the bulk.

There is further discussion of the packing fraction in §3.4.3.2.

3.3.2.6.2 Birefringence

Birefringence means literally "two refractive indices." The birefringence of a layer on a slab would be observed when the incident light (incident normally upon the slab) is linearly polarized, as depicted in Fig. 3-25. A refractive index n_y is observed when the direction of the electric vector (i.e., plane of polarization) of the light is along the y axis. The electric vector is then rotated 90° and a different index n_x is measured.

Birefringence is produced by several methods: (1) The film material itself is birefringent (see Barr [48]); (2) the birefringence is induced by mechanical stress (see Koppelman [67]); and (3) "form birefringence" occurs because of an orientation of the crystalline grains of the film — usually caused by the angle of impingement of the evaporant vapor (see §3.4.2).

3.3.2.7 Factors influencing nucleation and grain growth

A few factors are mentioned that influence the nucleation and subsequent grain growth. These, in turn, affect the mechanical hardness, refractive index and other properties.

3.3.2.7.1 Properties of the interface

Here the *interface* refers to the medium upon which the film is being deposited — is it a bare substrate, a film of the same composition or even a film of another material? Often nucleation layers are deliberately added. Jeppesen [63] found, for example, that if a substrate was coated with a nucleation layer consisting of a 1 nm thick gold layer, the 30 nm thick lead layer deposited on top if it was virtually continuous and manifested fine grain structure.

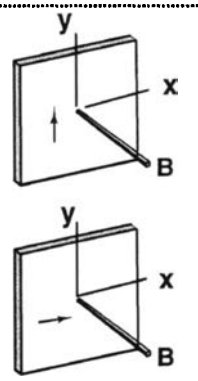


Fig. 3-25 A beam of linearly polarized light *B* is incident normally upon a slab. (*Upper*) The direction of the electric vector is along the *y* axis. (*Lower*) The direction of the electric vector is along the *x* axis.

3.3.2.7.2 Temperature

An increase of the substrate temperature increases the mobility of the adatoms during the early stages of film formation. Fewer nucleation sites are produced. Films of larger grain size result, especially when the deposition rate is slow. This is illustrated by Segmuller [63], who observed that nickel films deposited on a 250 °C substrate had a grain size of 30 nm; at 300 °C the size was 150 nm. The size of the crystallites of rutile sputtered by Pawlewicz [79] increased from 20 nm to 60 nm as the substrate temperature was raised by boosting the rf power.

3.3.2.7.3 Deposition rate

In general, a slow deposition rate provides an opportunity for quasi-equilibrium to occur — fewer nuclei form and the adatoms that subsequently impinge upon the substrate have sufficient time to migrate around the substrate and attach themselves to these existing nuclei. Large crystals tend to form. As an illustration, Pawlewicz [79] found that when rutile films were sputtered above a certain rate, the crystallite size steadily decreased as the rate increased.

3.3.2.7.4 Residual gases

These gases are usually at the temperature of the evaporation chamber and may or may not be introduced deliberately. An example is that at a pressure of 10^{-5} Torr (1.3×10^{-3} Pa), the gas bombardment of a substrate by residual water vapor is sufficient to form a monomolecular layer in 1 second.

Hoffman [76] deliberately introduced oxygen (pressure of 8×10^{-5} Torr, 1.1×10^{-2} Pa) during the deposition of a rhodium layer; the grain size was relatively small. At

lower pressures, the grain size was larger. She attributed the small grain size to an increase in the number of nucleation sites due to the presence of the oxygen.

3.3.2.7.5 Ion and electron bombardment

The effect of electron bombardment is to heat the substrate. This occurs when the coated substrate has sufficiently low surface electrical resistance so that the charge can leak off. Ion bombardment of the substrate is quite common in bias sputtering and ion-aided deposition (see §9.4.5.2). The ions add energy to the surface so that the adatoms are more mobile, thus engendering larger grain size. The disadvantage is that gases may be embedded in the film.

3.4 Process parameters influencing optical properties

This section expands §3.3.2.7 and discusses how the refractive index and other parameters are influenced by the deposition process.

3.4.1 Method of deposition

The data in Table 3-2 illustrate that the index of titania films can vary quite widely, depending upon the process. These processes include chemical deposition, sputtering and vacuum deposition — with and without ion assist. The substrate temperature, deposition rate and residual gases all affect the index. Even given the same deposition conditions, the index of the film may be influenced by what underlies it, whether it be a bare substrate or another layer of different composition.

Table 3-2 Refractive indices of titania films

Refractive index	Wavelength	Deposition method	Reference
2.23	500 nm	Chem 1	Suikovskaia [71]
2.08	500 nm	Chem 2	Nakano [85]
2.35 - 2.45	500 nm	Evap 1	Borgogno [81]
2.33 - 2.62	530 nm	Evap 2	Pulker [76]
2.10 - 2.42	500 nm	Evap 3	Cherepanova [79]
2.51	500 nm	PICVD	Heming [92]
2.52	633 nm	CVD	Desu [92]
2.53	500 nm	IAD	Allen [83]
2.10 - 2.45	1060 nm	Sputter	Pawlewicz [79]

Refractive index (at the specified wavelength) of thin films of titania deposited by different methods. Key to methods: Chem 1: Solvent evaporate tetraethylorthotitanate and bake at 300 °C. Chem 2: Apparently the same as Chem 1: Evap 1. Evaporate either TiO or Ti_2O_3 in presence of oxygen. Evap 2: Substrate temperature 200 °C, 450 °C. Evap 3: Substrate temperature 100 °C, 300 °C. PICVD: Pulsed microwave CVD. IAD: Substrate temperature 200 °C — oxygen ion bombardment during deposition. CVD: Chemical vapor deposition at 460 °C. Sputter: rf sputter titanium in an atmosphere of argon and oxygen.

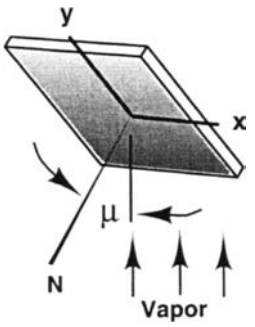


Fig. 3-26 Stream of vapor (arrows) is incident upon the plane substrate at an angle μ measured to the normal *N* of the substrate. The *y* axis lies in the plane of *N* and the vapor stream.

The energy with which the atoms arrive at the substrate is important. Sputtered atoms arrive with approximately 5 eV energy, for example. This allows such adatoms to migrate more freely over the substrate. Larger grains grow — the resulting layer manifests a larger packing fraction.

3.4.2 Vapor impingement angle

In the vacuum evaporation process, it is possible that the molecules of the evaporant hit the substrate tilted at an angle μ (to the normal) as shown in Fig. 3-26. When that impingement angle μ is relatively large — say 60° — then the resultant layer may manifest several unusual optical properties: (1) Its refractive index depends upon the impingement angle; (2) it is optically active [Koppelmann 64]; and (3) it is birefringent.

3.4.2.1 Change of refractive index

There are numerous examples in which the refractive index changes with impingement angle of the vapor. Wecht [91] deposited SiO layers at several impingement angles and measured their indices at 633 nm. Figure 3-27 illustrates that the index varied from 2.04 (at μ = 0) to 1.24 when the impingement angle of the vapor approached 80°. It is possible that this lowering of the refractive index is related to physical density of the films. Figure 3-28 shows that the density of a germanium film decreases about 30% as the impingement angle of the vapor increases from zero to 80° [Nakhodkin 75]. Figure 9-30 shows the lowering of the refractive index with increasing angle of incidence.

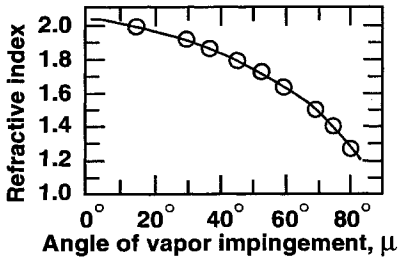


Fig. 3-27 Versus impingement of the evaporant vapor, refractive indices of SiO films at 633 nm wavelength. After Wecht [91].

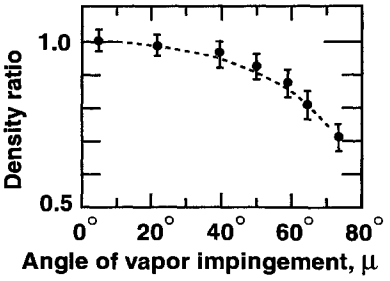


Fig. 3-28 Versus impingement angle of the evaporant vapor, density of a germanium film normalized to its density when the vapor is incident at zero angle. After Nakhodkin [75].

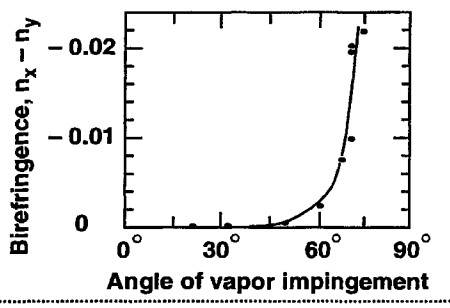


Fig. 3-29 Measured birefringence of an evaporated ZnS film vs impingement angle μ of the vapor. μ and γ appear in Fig. 3-26. After Koppelmann [67].

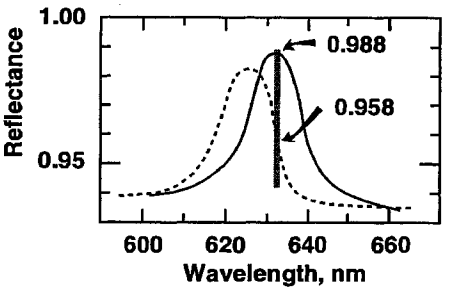


Fig. 3-30 Measured reflectance for two orthogonal linear polarizations of a titania-silica laser mirror in which the outer layers were deposited with the evaporant vapor at a relatively large impingement angle. After van der Wal [80].

3.4.2.2 Birefringence Koppelmann [67] sandwiched a zinc sulfide layer between two silver layers by evaporation in a vacuum. The impingement angle of the vapor for the silver layers was close to zero, but the impingement angles ranged up to 80° for the zinc sulfide. Similar to the aluminum in Fig. 3-21, the grains of zinc sulfide were preferentially oriented (i.e., tilted) along the y axis.

The slab was removed from the vacuum and its spectral transmittance was measured at normal incidence with the linearly polarized light with the direction of the E vector in either the x or y direction, as shown in Fig. 3-25. Its birefringence vs impingement angle is graphed in Fig. 3-29 and shows that the refractive index is higher when the electric field of the linearly polarized light is parallel to the direction of the grains.

Another illustration is that birefringence was introduced deliberately into a multilayer mirror used internally in a He-Ne laser [van der Wal 80]. The outer films of a titania-silica stack were deposited at a high vapor impingement angle, which made them birefringent. Figure 3-30 shows that there is a 3% difference in reflectance between the orthogonal linear polarization directions. The result is that the higher reflectance is sufficient for "lasing" to occur. For the specified optical gain inside the laser tube, however, the lower reflectance does not permit "lasing." This laser outputs linearly polarized light, even though it contains no Brewster windows tilted at angle.

3.4.3 Post deposition environment of a film

The index and other properties of a film may be altered — either deliberately or otherwise — after the film is removed from a deposition chamber.

3.4.3.1 Heat treatment

After deposition, the film can be baked in vacuum, air, nitrogen or other atmospheres. The mechanical stress is usually thereby reduced. The refractive index is often altered, as is evident from the spectral reflectance in Fig. 3-31 of a hafnia film deposited on a fused silica substrate [Smith 78]. Analysis of the shift in the spectral reflectance curve shows that the postbake has decreased the index and slightly increased the metric thickness.

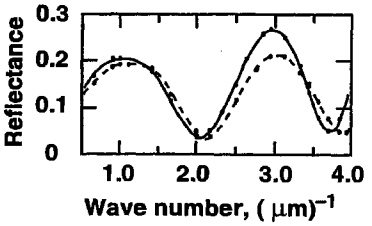


Fig. 3-31 Measured reflectance vs wave number of an electron-beam evaporated hafnia film after deposition (*solid*) and after baking in air 14 h at 400 °C (*dashed*). The substrate is fused silica. After Smith [78].

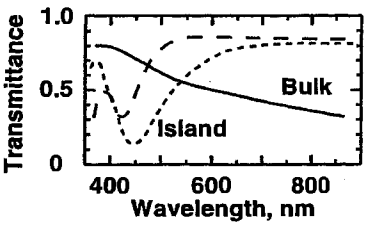


Fig. 3-32 Measured $T(\lambda)$ of an island film of silver 4.5 nm in thickness (*long dash*) and 11.5 nm thickness (*short dash*) on a glass substrate [Baba 93] . Computed $T(\lambda)$ (*solid*) of a silver film of 12 nm thickness using the optical constants in Table 8-6.

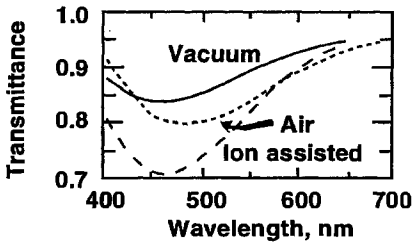


Fig. 3-33 (*Solid*) $T(\lambda)$ of a ZrO_2 evaporated film in the vacuum and (*short dashed*) after exposure to air. (*Long dashed*) ZrO_2 deposited with ion bombardment. $T(\lambda)$ in vacuum and in air are identical in the latter case. After Martin [84].

An island film coating for optical data recording was produced by post-baking (in air) silver at 350 °C for 600 s [Baba 93]. Figure 3-32 illustrates that island films manifest sharp minima in their transmittance curves (*dashed*) at 450 nm, as contrasted to a non-island (i.e., "conventional") film, whose transmittance (*solid*) is four times higher.

3.4.3.2 Adsorption of vapors

Whether due to grain growth (see §3.3.2.4) or other causes, films are often quite porous. This is illustrated by the voids between the grains (*shaded*) in Fig. 3-22. The voids readily adsorb water or other vapors as the coating emerges from the vacuum to a humid atmosphere. Figure 1-57 shows that this is manifested as a spectral shift. Hiraga [74] measured this shift in zirconia and other films by means of a rapid scanning spectrophotometer that monitored the reflectance *in situ* during deposition. Borgogno [84] measured both the R and T with a similar instrument and detected optical inhomogeneities.

Table 3-3 summarizes the refractive index shifts of single and multilayer coatings, as produced by adsorption of vapor. The shift of a multilayer depends upon the design of the stack and, at best, the data in Table 3-3 establish a trend.

3.4.3.2.1 Adsorption of water vapor

In most instances, water vapor adsorption causes the refractive index to increase, which may be explained by the fact that water vapor has a higher refractive index than the vacuum that it displaces. The change of refractive index with humidity is often manifested as a wavelength shift in the spectral transmittance — the slang is "wet-dry" (see Fig. 1-57).

Table 3-3 Shift of density or index by vapor adsorption

Composition	Chemical	*n* shift abs.	*n* shift %	Density	Reference
MgF_2	water	0.01	0.7		Kinosita [69]
MgF_2	ethanol	0.015	1.1		Kinosita [69]
MgF_2	water		4.0		Koch [65a]
MgF_2	water		3.0		Perveev [72]
MgF_2	water	0.01-0.05§		2%-12%	Ogura [75]
MgF_2	water		3.0	12%	Pulker [71]
SiO	water		0.2		Koch [65]
Cryolite	water		1.7		Gibson [83]
Cryolite	water			4.2%	Pulker [71]
Cryolite	water	0.02	1.5		Netterfield [80]
ZrO_2-TiO_2 mixture	water	0.07	3.6		Rabinovitch [82]
SiO_2	water		3.3		Perveev [72]
ZnS	water			1.5%	Pulker [71]
ZnS	water		0.25		Perveev [72]
SiO_2-ZnS	water		2.3		Perveev [72]
ZnS- cryolite	water		4%¶		Holm [80]
SiO_2-TiO_2	water		1%		See Fig. 1-57
SiO_2-MgF_2	water		4		Dobrowolski [72]
SiO_2-MgF_2	water		–1.4*		Dobrowolski [72]
ZrO_2	water		4		Hiraga [74]

For single layers or multilayers (*Composition* column), shift of the refractive index (*n* shift) or shift of physical density (*Density* column) as vapors of the *Chemical* are adsorbed/desorbed by it. All shifts are to higher values, with the exception of Dobrowolski*, where the coating was cemented at 100 °C and water was desorbed. In the case of a multilayer, the refractive index is the mean index.

§Ogura — depending upon the substrate temperature during the deposition of the layers.

¶Estimated by the author, after subtracting the change of metric thickness.

In the first example, a zirconia layer was deposited upon a glass substrate by conventional electron beam evaporation. Its spectral transmittance was measured in the vacuum, which is the *solid* curve in Fig. 3-33. The shift of the $T(\lambda)$ curve

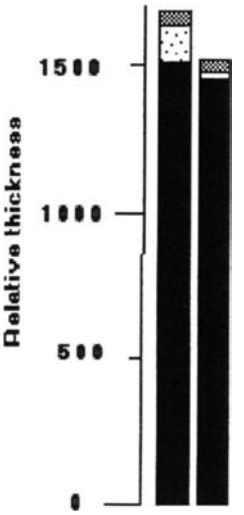


Fig. 3-34 Relative mass thickness (*black*) of layers of magnesium fluoride (M) or cryolite (C) as they absorb water vapor. The reversible water vapor is *dotted* and the irreversible is *shaded.*

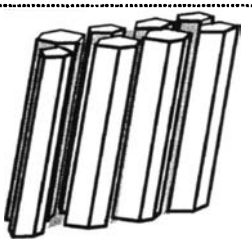


Fig. 3-35 A highly magnified view of the columnar growth (i.e., grains) of a thin film. The shaded regions represent water vapor adsorbed at the grain boundaries.

to lower values with the admission of air indicates that the refractive index of the layer increased as it adsorbed water vapor. The next step was to produce a different film by bombarding it with ions as it was deposited. This densified the layer sufficiently that its wet-dry shift is negligible, as evident from the *long dash* curve in Fig. 3-33.

In the second example, Pulker [71] measured the mass thickness of a magnesium fluoride film with a quartz crystal (see §9.6.1.2). As depicted in Fig. 3-34, the admission of air (with its water vapor) increased the mass thickness by 11%. After again immersing the film to a vacuum of 10^{-6} Torr, about 70% of the water vapor desorbed from the film. There was, however, an obdurate 30% of the vapor that remained locked into the voids between the grains.

The aforementioned results may be interpreted in terms of the grain growth of the film, as depicted in Fig. 3-35. There is a columnar structure consisting of the grains of magnesium fluoride. The water vapor is adsorbed at the boundaries of these grains.

3.4.4 Temperature shift of coatings

There are numerous spectral features of a coating that shift with temperature, such as the cuton wavelength of an edge filter (see §1.5.2.1) or the wavelength of peak transmittance of a bandpass. §1.6.4.2 mentions that such shifts may be attributed to two possible causes: (1) heating causes a loss of water vapor from the coating — the spectral shift of the $R(\lambda)$ curve is to shorter wavelengths; and (2) with the exception of a few semiconductor materials, the refractive index of the film material itself increases with

temperature. The index increase is manifested as a shift of the spectral features to longer wavelengths.

Figure 3-36 illustrates both of the aforementioned effects. A bandpass was constructed from alternate layers of magnesium fluoride and zinc sulfide. The wavelength of the transmittance peak of its passband was tracked as it was heated or cooled. The heating of the bandpass from 20 °C to 250 °C shifts its passband to shorter wavelengths. Thus, the effect of heating it is to "boil" the water vapor out of the voids in the coating materials. Because the addition of water vapor increases the refractive index, a loss decreases the optical thickness of each layer. A smaller optical thickness thus causes the coating to "resonate" at a shorter wavelength. Figure 3-36 shows that cooling of the coating causes it to shift again, but with considerable hysteresis.

The heating of the coating in a vacuum increases the refractive index of a thin film material, similar to a bulk material. There is a shift to longer wavelengths with increasing temperature and a reversible shift to shorter wavelengths as the coating is cooled.

Table 3-4 Temperature shift of optical coatings

Composition	Range	Type	Lambda	Method	Shift	Comment, reference
MgF_2	20° - 200°	S	633 nm	Evap	9×10^{-6}	Lissberger [86]
ZnS	20° - 200°	S	633 nm	Evap	1.5×10^{-5}	Lissberger [86]
$MgF_2 + ZnS$	20° - 380°	BP	800 nm	Evap	1.8×10^{-5}	Lazareva [70]
$MgF_2 + ZnS$	20° - 250°	BP	800 nm	Evap	-5×10^{-5}	Lazareva [70]
$Ta_2O_5 + SiO_2$	20° - 105°	E	520 nm	Reactive	$-40\ 10^{-5}$	Gürtler [88]
$Ta_2O_5 + SiO_2$	20° - 105°	E	520 nm	RIP	7×10^{-5}	Gürtler [88]
SiO_2+ ?	-20° - 70°	BP	1557 nm	PIAD	2×10^{-7}	Scobey [96]
$Ta_2O_5 + SiO_2$	25° - 100°	BP	1560 nm	PIAD	5×10^{-6}	Zöller [96]
$TiO_2 + SiO_2$	25° - 100°	BP	1560 nm	PIAD	1.5×10^{-6}	Zöller [96]
$ZnS + Na_3AlF_6$	20° - 90°	BP	545 nm	Evap	4×10^{-5}	Roche [76]

For layers or multilayers containing the indicated materials(*Composition*), fractional *shift* (in reciprocal K) of the wavelength of a feature of the coating with temperature in the specified *Range*. The coating was fabricated according to the *Method* column: *Evap*: thermal evaporation; *Reactive*: thermal evaporation in an oxygen atmosphere; RIP: low voltage reactive ion plating; PIAD: plasma ion assisted deposition. *Type* refers to the coating design: S: single layer; BP: bandpass filter; and E: edge filter.

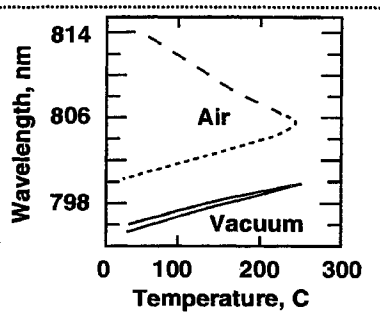


Fig. 3-36 (*Long dash*) Shift with temperature of the peak transmittance of a bandpass as it is heated in air. (*Short dash*) The same bandpass is cooled in air. (*Solid*) A bandpass of the same construction is heated and cooled in a vacuum. The bandpasses are composed of thermally evaporated layers of magnesium fluoride and zinc sulfide. After Lazareva [70].

The temperature shifts of coatings are summarized in Table 3-4. The coatings that manifest negative temperature shifts are, in general, porous — the shifts occur due to the "boiling" of the water vapor from the thin films. There are coatings, such as those produced by RIP [Gürtler 88], PIAD [Zöller 96] and a proprietary PIAD [Scobey 96], that are quite dense and thus relatively free of a wavelength shift with temperature due to water vapor.

3.5 Criteria for thin film material selection

3.5.1 Introduction

What materials should be incorporated in a multilayer? What criteria are useful in selecting them? A response: "Thin film materials with the proper refractive index" may be less than adequate because this ignores other important attributes such as low mechanical stress, physical robustness, low optical absorption and manufacturability.

3.5.2 Producibility

In many instances, the equipment for thin film production already exists and, consequently, the designer is constrained to using those materials that can be deposited with it. The limitations imposed by the equipment are often legion. An example is that refractory oxides such as silica and alumina can be evaporated only from electron beam sources. Most fluoride compounds are difficult to sputter. A paucity of organometallic compounds containing a certain metal, such as scandium, may preclude the production of an oxide layer (such as scandia) by CVD or PCVD.

3.5.3 Refractive index, absorption and inhomogeneity

3.5.3.1 Refractive index

§3.6.2 surveys the optical properties of thin films. In the visible part of the spectrum, refractive indices in the range of 1.32 to 1.34 are the lowest attainable. Although lower indices have been deposited, as, for example, by the sol-gel process, such films are porous and mechanically fragile. The corresponding highest indices available are in the range 2.35 to 2.45, although higher indices may be produced at the price of higher absorption. In the IR, indices as large as 4.2 (for germanium) and 5.2 (for lead telluride) are available. As discussed in §3.6.2.2, dispersion may or may not be important in multilayer design.

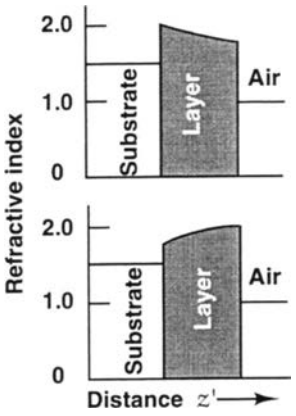


Fig. 3-37 Refractive index vs thickness of a film with a negative refractive index inhomogeneity (*upper*) and positive inhomogeneity (*lower*).

3.5.3.2 Absorption

§3.6.4.2 mentions systems for which highly absorbing coatings have been developed. In many instances, however, absorption is eschewed. Highly reflecting stacks and bandpass filters require films with the lowest absorption — the extinction coefficient k must not exceed 0.0001 in many applications. A small amount of absorption might be acceptable in an ophthalmic antireflection coating but would be intolerable in a laser system.

3.5.3.3 Optical inhomogeneity

Optical inhomogeneity implies that the refractive index of a film changes with thickness. Inhomogeneous films have been produced deliberately by evaporation [Jacobsson 75] or sputtering [Misiano 78] from two sources simultaneously. Such films are components of rugate filters (see §5.3.3.2). In §3.4.3.3.1, the focus is on layers in which the intent was to deposit an optically homogeneous layer, but the process went awry and an inhomogeneity occurred.

Table 3-5 Inhomogeneous nonabsorbing films

Film	n, at substrate	n, at air	Trans. tkns.	**Comment** or ref.
Cryolite	1.32	1.31	30 nm	Netterfield [80]
Cryolite	1.32	1.305		Bourg [61] **A**
CaF_2	1.34	1.28		Bourg [63] **B**
LiF	1.32	1.30	20 nm	Baumeister **C**
Scandia	1.91	1.79	450 nm	Arndt [84] **D**
Titania	2.7	2.3		Harris [79]
Zirconia	2.24	1.95	75 nm	Klinger [85]

For transparent optically inhomogeneous films, index at the substrate (*n, at substrate*), index adjacent to air (*n, at air*) and thickness at which transition takes place — if known. The index is measured in the middle of the visible spectral region.
Comment: (**A**) Substrate temperature 85 °C during deposition. (**B**) Substrate temperature 200 °C during deposition. (**C**) Estimated, from the author's experience in depositing this material. (**D**) Inhomogeneity apparently varies with thickness.

3.5.3.3.1 Optical inhomogeneity of a single layer

Fig. 3-37 depicts a cross section of films in which the index varies with the distance z' from the substrate. There are many instances in which $n(z')$ is a more complex function than quasi-linear. But regardless of the shape of the $n(z')$ curve, an increasing $n(z')$ is termed a *positive inhomogeneity* and decreasing $n(z')$, *negative inhomogeneity.*

During the film deposition, there are many parameters that might change to induce an inhomogeneity, such as substrate temperature, deposition rate or fractionization of an evaporant that is a ternary compound. In the PVD deposition process, a change in the pressure and/or composition of the residual gas also produces a gradient index.

Even when the aforementioned parameters remain relatively constant, Klinger [85] observed a change in the crystalline structure of zirconia films with increasing thickness. This is partially responsible for the index change cited in Table 3-5. Cho [96] observed that ion bombardment of the film during deposition reduces the inhomogeneity.

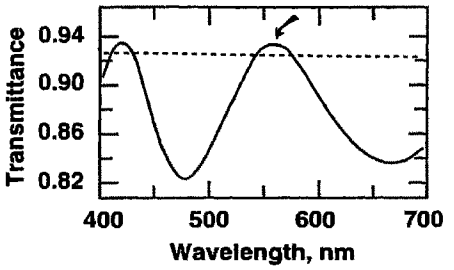


Fig. 3-38 Measured $T(\lambda)$ of an evaporated scandia film (*solid*) and uncoated substrate (*dashed*). The arrow designates the wavelength of halfwave optical thickness, at which the T exceeds that of the uncoated substrate. After Arndt [84].

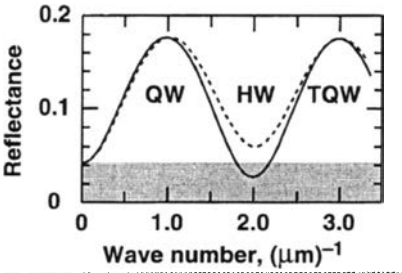


Fig. 3-39 $R(\sigma)$ of nonabsorbing layers deposited on a substrate of refractive index 1.52. (*Solid*) Film has a negative inhomogeneity — a linear change of index from 2.00 (substrate interface) to 1.85 (air interface). (*Dashed*) Positive inhomogeneity — a linear change of index from 1.85 (substrate interface) to 2.00 (air).

Table 3-5 contains conflicting results — undoubtedly related to the process. Borgogno [84] reported a negative inhomogeneity in titania, whereas the inhomogeneity of Englisch [83] is positive.

3.5.3.3.2 Detection of an inhomogeneity

One method of detecting an inhomogeneity in a nonabsorbing layer is to deposit it on a transparent substrate and examine its R or T when its optical thickness is a halfwave (or a multiple thereof). If the film is homogeneous, then §2.4.2.6 states that its R or T is precisely that of the uncoated substrate at these points. The T of a negatively inhomogeneous film exceeds that of the uncoated substrate, as shown in Fig. 3-38 for a scandia layer. As another example, Fig. 3-39 plots the computed reflectance vs wave number of a layer with a negative inhomogeneity (*solid*) and positive inhomogeneity (*dashed*). The halfwave points occur at a wave number of 2 $(\mu m)^{-1}$ The reflectance either undershoots or exceeds that of the uncoated substrate, depending upon the algebraic sign of the inhomogeneity. As illustrated in Fig. 2-50, however, a reflectance that exceeds the uncoated substrate may also be produced by absorption in a homogeneous film.

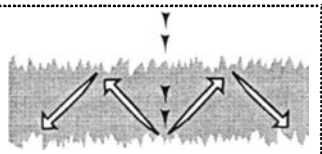


Fig. 3-40 Showing the path of an incident ray (*vertical* arrows) and the scattered light (*white* arrows) internally trapped in a high index layer. The roughness at the layer surfaces is exaggerated for purposes of illustration.

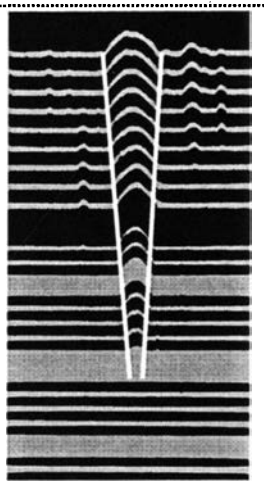


Fig. 3-41 Nodule growth is evident in a scanning electron microscope picture of a cross section of an IR multilayer. Superimposed are the white guidelines that define the nodule. After Aerospace [81].

3.5.4 Scatter and other losses

When light from a plane wave is incident at θ_0 upon a multilayer, scattering redirects this light at other angles. This effect may, or may not, be objectionable, depending upon the application. Consider, for example, the light scattered at small angles from the multilayer cold mirror in the illuminator of the projector in Fig. 1-60. Most of this light still reaches the film gate and therefore the scatter is usually tolerable. In the case of a mirror in a laser cavity, any light scattered out of the small solid angle of the cavity mode is a loss; scatter must be minimized.

Radiation trapping is another way in which scatter increases the loss of a multilayer. This is illustrated in Fig. 3-40, which shows a high index layer (n = 2.30) deposited on a lower index substrate (n_s = 1.52). Roughness at the upper interface scatters light, a portion of which is directed at angles as large as 45°. Because this exceeds the critical angle θ_c

$$\theta_c = \arcsin[1.52\ (2.30)^{-1}] = 41.4°, \qquad (3\text{-}3)$$

the light is trapped in this layer. In other words, the layer functions as an optical waveguide. Regardless whether a waveguide mode is formed or not, the trapped light must propagate laterally through the layer — parallel to the substrate — and is absorbed quite rapidly. Scatter may be attributed to several factors:

- Roughness at the interfaces of the layers;
- scatter in the bulk of the layers; and
- scatter from nodules.

Each is discussed below.

3.5.4.1 Surface roughness

Ebert [79] shows that the scatter loss of single zinc sulfide and thorium fluoride layers did not increase with their thicknesses. This suggests that the surfaces — and/or interfaces between layers — are the principal sources of scatter, rather than the scatter in the bulk of the layers.

3.5.4.2 Bulk scatter

This occurs in the interior of a layer. Few studies have been made to quantify it, although Elson [80] attributes a deviation between theory and experimental results to it. The trapping and waveguiding (see §3.5.4) of light — and its possible rescatter — have not as yet been included in theoretical treatments.

3.5.4.3 Nodules Figure 3-42 shows a fractured layer in which several nodules had grown. The protrusions at the top surface of the layer are also nodules.

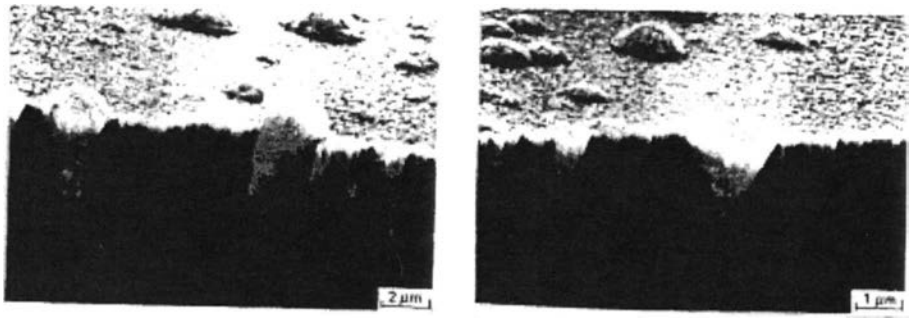


Fig. 3-42 Electron micrograph of a replica of a fractured layer of lanthanum trifluoride, showing several nodules. After Günther [81b].

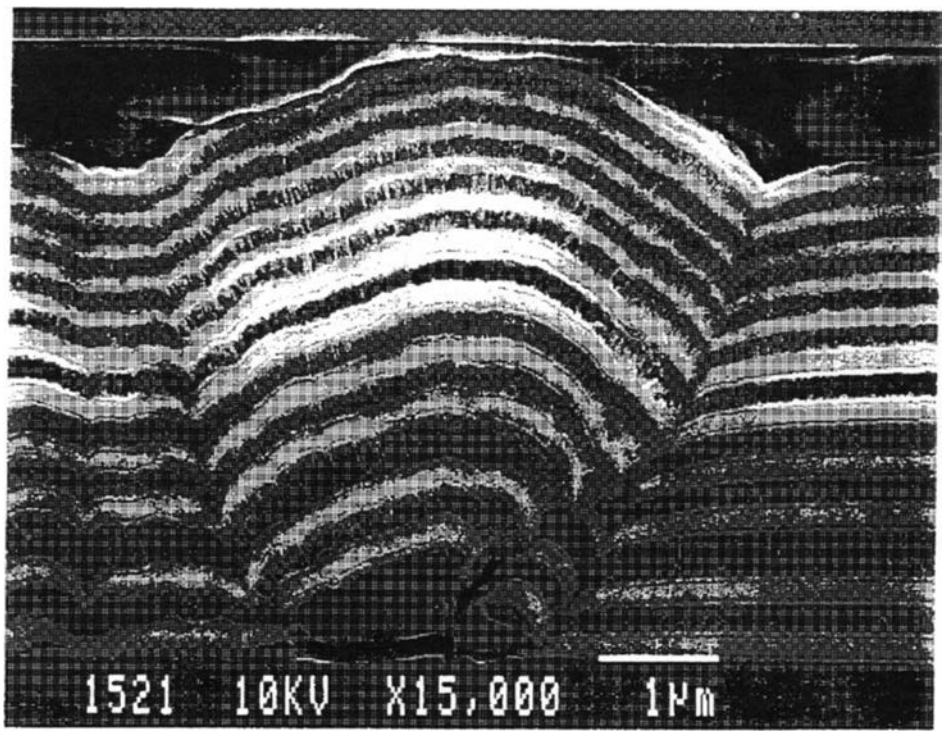


Fig. 3-43 Electron micrograph showing in cross section the seed of a nodule growing in a multilayer of silica (*dark bands*) and niobia (*light bands*). After Deposition Sciences [99].

3.5.4.3.1 Formation of nodules

Günther [81b, 82] concludes that the nodule is produced by preferential nucleation at a defect on the substrate or particulate matter from the evaporation sources. The latter is

termed the "seed" of the nodule. Figure 3-43 shows an electron micrograph of a cross section of a nodule — the seed is quite evident.

Kozlowski [94] investigated a silica-hafnia multilayer and found that the genesis of a nodule was a minute particle of hafnia. It was observed that if the distance the nodule protruded above the top surface exceeded 700 nm, laser damage at 1064 nm occurred.

3.5.5 Mechanical stress

Manifestations of stress in a coating are: (1) The coating may delaminate from the substrate, as shown in the photograph in Fig. 3-44; (2) the coating may adhere to the substrate, but crazes to relieve the stress; and (3) stress may warp or deform a thin substrate.

Mechanical stress often imposes limitations on the selection of coating materials: (1) A highly stressed material might be avoided, although it might have excellent optical properties; and (2) coating materials are often selected because there is compensation of the stress (see §9.7.3.2). As an example of the former criterion, magnesium fluoride is usually avoided as an IR coating material because it crazes when its optical thickness exceeds a quarterwave at 2 μm.

Figure 3-45 shows both the stress and refractive index of layers deposited by vacuum evaporation. With the exception of zinc sulfide, zinc selenide, silica and yttria, all manifest a tensive stress. CVD deposited films are tensively stressed, whereas sputtered oxide layers are compressively stressed.

3.5.6 Thickness limits

Several factors mentioned in the foregoing two sections often limit the thickness of a layer. Quite thick layers may be used in a coating, provided they have low mechanical stress and do not scatter. The classic example of a rough material is calcium fluoride. The scatter of light in a thick layer is intolerable.

The lower limit on the thickness of a layer is usually imposed by its transition from an island film (see §3.3.2.3) to a continuous layer. This occurs at a metric thickness of 10 nm for many materials. Another limitation occurs when optical monitoring is used to deposit a thin layer. Control of the thickness is difficult, to say the least.

3.5.7 Adhesor layers

The adhesor layer, also termed a subbing layer or bonding layer, improves the mechanical adhesion between a layer and substrate, or between a layer and adjacent layers. In many cases, adhesor layers are quite thin and, for that reason, may have little effect upon the optical properties of the coating. Table 3-6 lists adhesor and barrier layers. Another effect of an adhesor layer is to enhance the nucleation (see §3.3.2.2) and thus alter the grain size or possibly the crystalline structure.

Table 3-6 Adhesor and barrier layers

Substrate or adjacent layer	Film to be deposited	Adhesor layer composition	Reference
Oxide glass	Noble metal or Al	Ti or Cr	Holland [56], Ledger [79]
Oxide glass	Ag	SiO_2	Kuznetsov [75] ¶
Oxide glass	Ag	Al_2O_3	Volgunova [83]
Oxide glass	Ag	Al_2O_3 ZnO Ag_2O	Colbert [53, 54]
Oxide glass	Al	Ag + Ag_2O	Miller [63]
Oxide glass	Au	Ti	Rozhdestvenskiy [70]
Oxide glass	Au	Ti-Mo§	Nicolet [78]
Oxide glass	Au	Ti Zr Hf V Nb	Nicolet [78]
Oxide glass	Au	Al_2O_3 PdO NiO Cu_2O	Colbert [53, 54]
Oxide glass	Au	PbF_2 SnF_2 CdF_2	Singh [79]
Oxide glass	Cu	Au_2O_3 Cu_2O Ag_2O	Colbert [53, 54]
Oxide glass	SiO_2	Au_2O_3	Maya [96]
Plastic	MgF_2	Thin Al	Scharf [67]
Plastic	MgF_2	SiO or Al	Libbert [67]
Plastic	MgF_2	Al_2O_3	Allen [82]
Plastic	TiO_2	TiO	Lee [92]
As_2S_3 glass	MgF_2 or SiO	Ge	Wozniak [62]
As_2Se_3 glass	MgF_2 or SiO	Ge	Wozniak [62]
As_2S_3 glass	ThF_4	PbF_2	Jupnik [58]
As_2Se_3 glass	ThF_4	PbF_2	Jupnik [58]
Germanium	ZnS or ZnSe	MgO	Hacker [90]
Germanium	ZnS or ThF_4	Y_2O_3	Lubezsky [83]
Gold*	ThF_4	ZnSe	Ledger [79]
MgF_2*	ZnS	$ThOF_2$	Dimmick [46a, 46b]

The adhesor layers are intended for a substrate, unless otherwise noted by an asterisk *. The multiple entries on one line indicate alternate materials.

§To inhibit diffusion, the sequence of the layers are: substrate, Ti layer, Mo layer, gold.

¶Prior to the deposition of silver by the "wet chemical method," the surface was treated with a $PdCl_2$ solution.

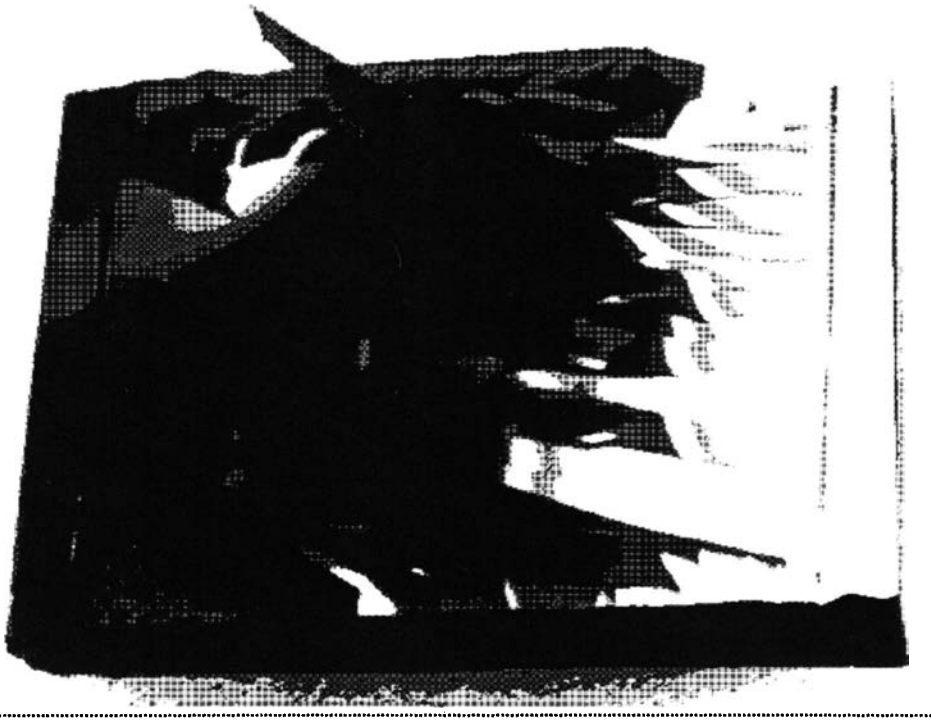


Fig. 3-44 PbF_2-ZnSe multilayer that peeled from the substrate. A few of the strips were 10 mm in length. After Edwards [79].

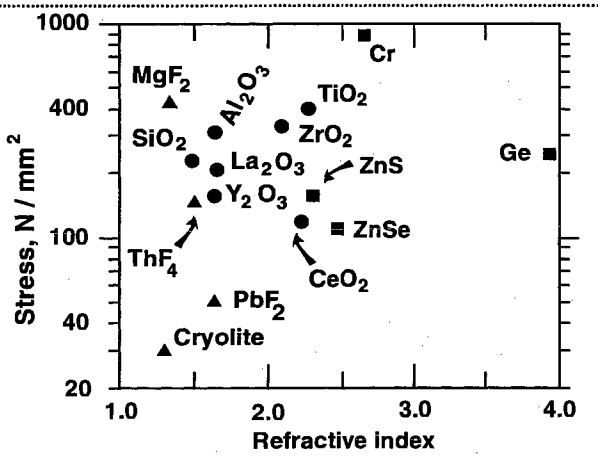


Fig. 3-45 Mechanical stress and typical refractive index of layers deposited by evaporation. With the exception of SiO_2, ZnS, ZnSe and Y_2O_3, all stresses are tensive. Stress data from Ennos [66].

3.5.8 Sacrificial layers

The role of a sacrificial layer is illustrated by the following story. Two men were camping and had been warned that a man-eating bear inhabited the region. One of them looked out of the tent and saw a bear in the distance — walking towards them. As he frantically donned his running shoes, his buddy said, "Isn't it silly to put on your running shoes? It is known that a person cannot run faster than a bear." "It is not necessary for me to run faster than the bear" he responded, "just faster than you."

A typical manifestation of a sacrificial layer is one more chemically reactive than other materials in the coating. Consider a silver layer used as a

reflector, as shown in Fig. 1-73. A copper layer contiguous to the silver may be used as the sacrificial layer [Dar-Yuan Song 85]. Because the copper is chemically more reactive than the silver, moisture that penetrates the system attacks it, rather than the silver. Palladium serves the same purpose.

In the next example, a zinc oxide layer is to be sputter deposited on top of a silver layer. The ZnO is formed by sputtering zinc metal in the presence of oxygen. The problem is that until the zinc oxide layer is sufficiently thick, the oxygen converts some of the silver into silver oxide. The solution is to deposit a thin layer of titanium metal (often called a "subbing layer" or "foundation layer") on the silver. It protects the silver.

3.5.9 Other criteria

Previous sections list criteria for the selection of a coating material, which include: refractive index, amount of optical absorption, optical homogeneity, level of mechanical stress and amount of scatter. Other criteria for the selection of a coating material are:

3.5.9.1 Compatibility with the substrate

Certain substrates — notably plastic and esoteric glasses — cannot be heated. Because the time-honored method of producing mechanically hard oxide layers is by heating the substrate during deposition, this may restrict the use of certain coating materials.

3.5.9.2 Availability

Let us suppose that a layer of zirconium tetrafluoride is to be deposited by evaporation in a vacuum. When a bottle of that material arrives from a supplier, it is most likely in the form of a white powder. Any attempt to evaporate that powder from a "boat" (see Fig. 9-13) or from an electron gun (see Fig. 9-14) will meet with disaster. The reason is that heating releases the gas trapped in the powder, which in turn ejects the dust of the evaporant into the vacuum. A miniature volcano occurs.

The aforementioned problem is avoided by, prior to evaporation, outgasing the material and compacting it into tablets. Unless a coating house has its own facilities to do so or can purchase this processed material, that material can then be evaporated only with great difficulty — or perhaps not at all. The preparation of a sputtering target requires similar processing.

3.5.9.3 Cost

Metals such as rhodium, iridium or gold, as well as various rare earth compounds, are often too costly to coat over a large area.

3.5.9.4 Robustness

A material might be used because of its mechanical robustness or its ability to survive in a hot, humid environment. The rhodium coating on a dental mirror, for example, survives immersion in the autoclave several times daily.

Table 3-7 Thin films composed of mixtures of dielectric materials

Components	Lambda in μm	Refractive index range	Usage *n*	*k*	s	o	Reference and comment
Cryolite ZnS	0.4 -	1.45 - 2.10	*				Jacobsson [75]
MgF_2 ZnS	0.4 -	1.85	*				Chopra [79] **QH**
MgF_2 ZnS	0.4 -	1.70 - 1.95	*				Dobrowolski [78b] **A**
MgF_2 SiO_2	0.4 -	1.39 - 1.43			*		Kurdock [78], Austin [72b]
MgF_2 SiO	0.4 -	1.6 approx.				*	Libbert [67] **B**
ThF_4 ZnS	0.4 -	1.59 - 2.33	*		*		Austin [72b] **C**
CeF_3 ZnS	0.4 -	1.65 - 2.35	*				Fugiwara [63a] **QHQ**
MgF_2 Ge	2.0 -	1.80 - 3.20	*				Jacobsson [75]
ZnS Ge	2.0 -	2.60 - 3.50	*				Jacobsson [75]
SrF_3 CeF_3	8 - 12	1.30 - 1.50		*	*		Lubezsky [83] **D**
YF_3 CaF_2	0.3 - 12	1.23 - 1.55	*				Jacobson [94]

Thin films composed of dielectric mixtures, showing the wavelength range (*Lambda* column) in which it is useful and range of indices attained. The usages are: *n, k, s* and *o* — to modify the refractive index, extinction coefficient, mechanical stress and other application, respectively. **QH** and **QHQ** mean that material was used in quarter-half and quarter-half-quarter antireflection coatings, respectively. Comments: **A** Component of an optical minus filter. **B** Mixture adhered to a plastic substrate. **C** Scatter of mixtures was measured. **D** Component of an IR antireflection coating.

3.5.9.5 Adsorption of water

An example is a bandpass filter with ultranarrow bandwidth. Coating materials are needed whose optical properties are minimally unaffected by temperature and humidity.

3.5.9.6 Toxicity

An example is that notwithstanding desirable properties of berylia as an UV coating material [Ebert 82], its toxicity would preclude its use.

3.5.9.7 Laser damage

Coating materials are selected for their ability to resist damage from the intense irradiation of a laser beam. The requirements change radically with the wavelength and whether the laser is CW or pulsed. The tests of Günther [84] show that the damage threshold can vary by at least a decade, depending upon the pulse width and beam diameter.

Table 3-8 Thin films composed of mixtures — oxides

Components	Lambda in μm	Refractive index range	Usage *n*	*k*	s	o	Reference and comment
CeF_3 CeO_2	0.4 -	1.65 - 2.10	*				Fugiwara [63b] **QHQ**
CeF_3 CeO_2	0.4 -	1.63 - 2.30	*				The film was sputtered. Zheng [92]
CeF_3 CeO_2	0.4 -	1.68 - 1.95	*				Jacobsson [75], Katsube [76] **QHQ**
SiO_2 Ta_2O_5	0.63 -	1.78 - 2.05	*				Kobayashi [83] **A**
SiO_2 TiO_2	0.4 -	1.62 - 2.00	*				Misiano [78] **QHQ**
SiO_2 CeO_2	0.4 -	1.65 - 2.00	*				Misiano [78] **QHQ**
SiO_2 CeO_2	0.3 - 0.4	?		*			Rheinberger [70] **B**
Al_2O_3 Ta_2O_5	0.4 -	1.63 - 1.80	*				Ganner [83] **QHQ**
Al_2O_3 TiO_2	0.4 -	2.00	*				Louderback [7172] **QHQ**
Al_2O_3 MgO	0.4 -	1.66 - 1.76	*				Heaney [81] **C**
ZrO_2 TiO_2	1.06	2.00	*			*	Stetter [76], Rabinovitch [82] **D**
ZrO_2 Ta_2O_5	0.4 -	2.05	*			*	Wille [74] **QHQ**
ZrO_2 Ta_2O_5	0.4 -	2.05	*			*	Kozawa [81] **QHQ**
$Hf_xSi_yO_2$	0.22-		*				The film was sputtered. Edlou [92]
HfO_2 Ta_2O_5	0.4 -	2.05	*			*	Englisch [83]
CeO_2 Pr_6O_{11}	0.4 -	?	*				Kraus [62]
CeO_2 TiO_2	0.4 -	?	*				Kraus [62]
PrF_3 MgO	8 - 12	1.60		*	*		Kurdock [74] **E**
PrF_3 La_2O_3	8 - 12	1.60		*	*		Kurdock [74] **E**
WO_3 ZnS	?	?			*		Kurdock [78] **F**
TeO_2 Bi_2O_3	0.4 -	2.15		*		*	Lieberman [72] **G**

Thin films composed of dielectric mixtures containing oxygen, showing the wavelength range (*Lambda* column) in which it is useful and range of indices attained. The usages are: *n, k, s* and *o* — to modify the refractive index, the extinction coefficient, the mechanical stress and other applications, respectively. **QH** and **QHQ** mean that material was used in quarter-half and quarter-half-quarter antireflection coatings, respectively. Comments: **A** Sputtered film used in optical waveguides. **B** Level of UV absorption was varied. **C** Range of values represents irreproducibility of index. **D** Mixtures reduce the index inhomogeneity in zirconia. **E** Index estimated from transmittance curves. **F** One of many mixtures developed to aspherize optical surfaces. **G** Sputtered film — the tellurium dioxide reduced the absorption of the bismuth oxide.

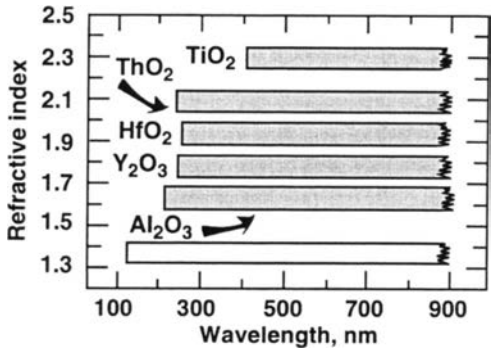


Fig. 3-46 Representative refractive indices and range of transparency of optical coating materials for the UV, visible and near IR. The lowest nonshaded bar represents magnesium fluoride.

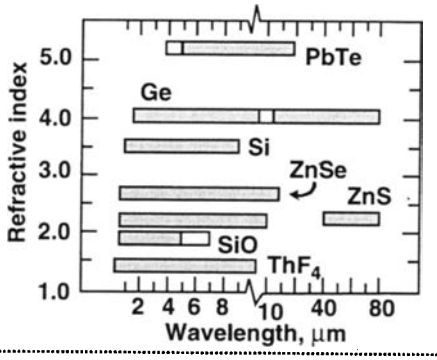


Fig. 3-47 Representative refractive indices and range of transparency of optical coating materials for the IR. A layer of the material is absorbing in the nonshaded spectral regions. The abscissa changes scale at 10 μm.

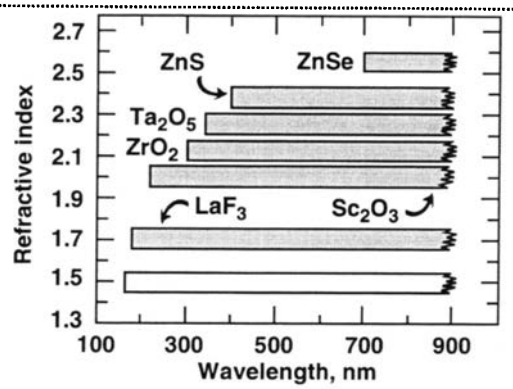


Fig. 3-48 Representative refractive indices and range of transparency of optical coating materials for the UV, visible and near IR. The lowest nonshaded bar represents thorium fluoride.

3.6 Survey of coating materials

3.6.1 Introduction

Given the criteria for coating material selection in §3.4, the designer and producer of the coating might select a material from this book or from other compendia [Holland 56, Ritter 75, Borisevich 71, Pulker 84]. Figs. 3-46, 3-47 and 3-48 show the approximate index (*approximate* due to dispersion) and range of transparency of coating materials. The survey article of Black [68] reviews IR thin film materials and relates their optical properties to parameters such as the band gap, etc.

Strehlow [73] reviews bulk materials. The real and imaginary parts of $\hat{\varepsilon}$ in Eq. (2-69j) are related by a dispersion relationship [Hodgson 70]. A consequence is that all materials are absorbing in at least one spectral region. In general, the higher index materials such as germanium, silicon and lead telluride are transparent only in the IR, whereas a lower index substance such as magnesium fluoride is transparent to wavelengths as short as 120 nm.

3.6.2 Dielectric coating materials

3.6.2.1 Survey

In excess of 100 materials used in thin films appear in Table 3-10, which is relegated to an appendix — §3.7. The layers were deposited by either vacuum evaporation or sputtering. Most are binary compounds and do not contain any water of hydration if they are in crystalline form. Table 3-9 omits compounds that are highly soluble in water —

this excludes most of the alkali halides. Each material in Table 3-10 is relatively transparent in either the UV, visible or IR. The range of transparency is arbitrary, since it depends upon the film thickness and the application.

There also may exist absorption bands within the range of transparency, as, for example, the 2.5 μm – 2.8 μm absorption of water prevalent in many oxide coating materials. Most of the coating materials in Table 3-10 are inorganic, although there are exceptions [Barr 48, Hollahan 74, Shen 76].

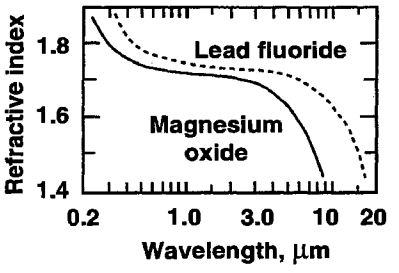


Fig. 3-49 Refractive index vs wavelength of bulk magnesium oxide (*solid*) and lead fluoride (*dashed*). Data from references cited in Table 3-10.

3.6.2.2 Dispersion

This is discussed in general terms in §2.1.3. Dispersion data may or may not be essential to design a multilayer, depending upon the level of the dispersion and range of wavelengths.

An illustration is the dispersion of magnesia and lead fluoride over several decades of wavelength (Fig. 3-49). In the spectral region 1 μm to 2 μm, the index varies by less than 2%. Consequently, many designers would neglect the dispersion of such materials in the design of a filter that functions in this band. The dispersion is, however, more important in the spectral regions near 0.25 μm and 10 μm.

3.6.3 Mixtures — mostly all-dielectric

Mixtures of coating materials have been concocted to: (1) modify the refractive index; (2) either increase or decrease the absorption; and (3) reduce the mechanical stress. The latter is accomplished by mixing tensively and compressively stressed materials. Mixtures of dielectric materials are listed in Tables 3-7 and 3-8. The addition of a metal to a dielectric film does not guarantee that the mixture will be highly absorbing. An example is that the inclusion of a small amount of gold in a dielectric film increased the refractive index in the IR without any measurable increase in its absorption [Stensland 69].

3.6.3.1 Mixtures — proprietary

Table 3-9 lists coating materials consisting of mixtures sold by commercial vendors. The refractive indices and wavelengths at which the materials are transparent are indicated, if known. Is it possible that the composition of these materials was anticipated by Shakespeare (MacBeth, act IV)?

"Eye of newt and toe of frog,	Adder's fork and blind worm's sting,
Wool of bat and tongue of dog,	Double, double toil and trouble;
Lizard's leg and howlet's wing,	Fire burn and cauldron bubble."

Table 3-9 Proprietary mixtures of coating materials

Name of material	Possible component(s)	Refractive index @ spectral region	Transmittance range	Comment	V
Substance H1	TiO_2	2.10 @ 500 nm	0.40 μm—7.0 μm		E
Substance H2	TiO_2 Pr_6O_{11}	2.10 @ 500 nm	0.42 μm—5.0 μm		E
Substance H4	TiO_2	2.10 @ 500 nm	0.40 μm—7.0 μm		E
Substance M1	Al_2O_3 Pr_6O_{11}	1.70 @ 500 nm	0.40 μm—9.0 μm		E
Substance A	SnO_2	2.00 @ 500 nm	0.40 μm—9.0 μm	(a)	E
YF_3 1A	YF_3 CaF_2	1.45 in the IR	1.00 μm—12 μm		E
Paso I		1.71 @ 550 nm	0.30 μm—9.0 μm		B
Paso II		1.81 @ 550 nm	0.30 μm—9.0 μm		B
Paso II		1.92 @ 550 nm	0.40 μm—9.0 μm		B
Lida	TiO_2¶	2.13 @ 550 nm	0.35 μm—10.0 μm		B
Pisa		2.11 @ 550 nm	0.30 μm—10.0 μm		B
Ida	TiO_2¶	2.18 @ 550 nm	0.40 μm—11.0 μm		B
Dimo		2.05 @ 550 nm	0.35 μm—7.0 μm		B
Mira		1.47 @ 550 nm	0.30 μm—12.5 μm	(b)	B
P-3000	Pr_6O_{11} TiO_2	2.0 @ 550 nm			Ce
Spinel	MgO Al_2O_3 ZrO_2	1.7 @ 550 nm			Ce
Y-1056	Y_2O_3 ThO_2				Ce
Z-1077	ZrO_2 MgO	2.0 @500 nm			Ce
Z-1092	TiO_2 ZrO_2	2.1 @500 nm			Ce
IRB	BaF_2¶	1.35 @ 10 μm	0.30 μm—13 μm		Ce
IRX	CeF_3	1.43 @ 10 μm	0.30 μm—12 μm		Ce
OH-5	TiO_2 ZrO_2	2.09 @ 500 nm	0.38 μm—6 μm		Op
OM-4	Al_2O_3 ZrO_2	1.67 @ 500 nm	0.35 μm—5 μm		Op
OH-10	Pr_6O_{11} TiO_2	2.05 @ 500 nm			Op
Glass	Ca B Na K O	Presumably near 1.50		(c)	S

V (Vendor) column:
B — Balzers Aktiengesellschaft (Fürstentum Liechtenstein). Sales Literature, circa 1993.
Ce — Cerac Corporation, Milwaukee, Wisconsin. Sales literature (1995).
E — E. M. Chemicals, Hawthorne, New York (in the United States), E. M. Merck GmbH (Darmstadt, Germany). Sales literature (1996).
Op — Optron Corporation (division of Canon Corp.), Ibaraki 302, Japan. Distributed in the US by Wacom, Spring Valley, New York. Sales literature (1996).
S — Schott Corporation. Mainz, Germany.
(a) An indium-tin oxide mixture with a typical k value of 0.05 in the visible part of the spectrum.
(b) This material is intended as a component of an antireflection coating in the IR.
(c) See Aussenegg [73].
¶ Speculation on the part of the author.

3.6.4 Absorbing materials

3.6.4.1 Introduction

This subject is vast because, if for no other reason, it is usually easier to deposit an absorbing film than a nonabsorbing one. Absorbing films are used in coatings such as:

(1) Metallic reflectors, such as aluminum or silver. These are discussed in §1.7.4.
(2) Semitransparent films for beamdividers, bandpasses and neutral densities. Examples are cited in §1.6.2.3, §1.8.2.1.3, §1.9.3 and Chap. 8.
(3) Absorbing coatings — the subject of §1.9.4 and §8.7. The discussion centers on this last category.

3.6.4.2 Absorbers: sources of information

A large area (many square km!) of architectural glass is coated yearly with absorbing films such as chromium, titanium nitride and silver. Both the deposition equipment and the process may be purchased. The purveyors of coating materials sell absorbers for the production of sunglasses. Chrome-silica mixtures [Anders 67a] have also been used.

Absorbers for photothermal energy converters have received considerable attention. Bogaerts [83] cites 80 such coatings, for example. Books such as Agnihortri [81] and review articles [Hahn 78] describe the production and characterization of such coatings. The dark mirror (shown in Fig. 1-216) has been widely used. Bastien [78] used both homogeneous and inhomogeneous concentrations of nickel in a matrix of alumina to fabricate such a mirror.

The dispersion model (see §2.11.1) predicts that the absorption should be dependent upon wavelength. IR transparent semiconductors such as silicon and germanium absorb at wavelengths in the visible. The converse is true of a highly doped semiconductor such as a mixture of indium oxide and tin oxide, which is quasi-dielectric (with $n \approx 2$ and $k \approx 0.05$) in the visible, yet behaves like a metal at wavelengths longer than 3 μm. Turner [65] observed that a dielectric acts like a metal in the wavelength region near its reststrahlen band and used such "metals" to fabricate IR multilayer filters.

3.6.4.3 Dielectrics with low absorption

The foregoing discussion centered on materials with substantial absorption. Where are data on the extinction coefficients — the k values — of the dielectric coatings in Table 3-10? The reference to the literature often provides k as well as n data. In the IR, the k of many coating materials has been measured at 2.8 μm, 3.8 μm and 10.6 μm for use in high power lasers.

3.6.5 Commonly used coating materials

At least many hundreds of optical coating materials have been deposited, yet only a few score are in common use because they are mechanically and chemically stable. The information in Table 3-10 is obtained from the patent literature, catalogs of the purveyors of coating materials, publications and the author's personal experience.

3.6.5.1 Vacuum UV

This spectral region extends from 110 nm (transmittance limit of bulk lithium fluoride) to 180 nm, where oxygen gas is nonabsorbing. Magnesium fluoride and aluminum have been deposited in ultra-high vacuum systems for bandpasses at 122 nm [Schroeder 62]. Thorium fluoride is transparent to 160 nm and has been used in dielectric reflectors. Malherbe [74] produced all-dielectric reflectors at 180 nm with magnesium fluoride and lanthanum fluoride.

3.6.5.2 UV

The oxides alumina, yttria, scandia and silica are useful at wavelengths longer than 220 nm. Although the mechanical stress often causes a multilayer reflector to delaminate from the substrate, there is little alternative but to use the fluorides of thorium, magnesium or the rare earths at shorter wavelengths. Zirconia and hafnia have been incorporated in reflectors at wavelengths longer than 300 nm. The problems of producing multilayer stacks in this region are not only the absorption but also the mechanical stress and scatter. Notwithstanding its mechanical softness, lead fluoride is usable to wavelengths as short as 260 nm.

3.6.5.3 Visible and near IR

Titania, tantala, zirconia, hafnia, silica, alumina and magnesium fluoride are widely used ingredients of these coatings. These materials are mechanically robust. Cryolite and zinc sulfide are used in interference filters protected by cementing.

3.6.5.4 IR

Silicon and germanium are used as high index materials. Zinc sulfide, zinc selenide, silicon monoxide and rare earth fluorides are commonly used as lower index films. Fig. 5-26 shows that a multilayer composed of lead fluoride and zinc selenide is transparent to 17 μm. Table 3-10 cites materials transparent at even longer wavelengths. Hard carbon is a mechanically robust single-layer antireflection coating for germanium. Smith [68] reviews far IR coatings.

3.7 (Appendix) List of useful coating materials

The optical properties of these materials appear in Table 3-10 on the following four pages.

Table 3-10 Optical properties of coating materials

Film	Index in the wavelength (μm) region					Transp'cy		Reference, note and/or comment
	0.35	0.55	1.0	1 - 8	10.0	Low	Up	
AlF_3	1.385	1.385	1.385			0.2	10	Heitmann [70]
AlN				2.10		0.33		Borisevich [71]
Al_2O_3	1.66	1.63	1.60			0.15	9	Cox [64a], Newnam [78] §
Sb_2O_3	2.18	2.00		1.93		0.3	12	Honcia [61], Shirokshina [58], Jenkins [58], Butterfield [73]
Sb_2Se_3						0.6	16	Schneeberger [66], Ghosh [79]
Sb_2S_3			3.0	2.70	2.38	0.8	24	Valeev [65], Ghosh [79]
As_2Se_3					2.80	0.77	10	Borisevich [71], Pedinoff [79]
As_2S_3			2.66	2.41	2.38	0.6	12	Borisevich [71], Pedinoff [79]
As_2Te_3				3.90		1.1		Borisevich [71]
BaF_2					1.38	0.25	14	Baer [76a]
BeO		1.72		1.72		0.18	6	Ebert [82], Borisevich [71] *
BiF_3			1.75		1.65			Moravec [79]
Bi_2O_3		2.55	2.35		2.2	0.38	14	Traylor [92], Lieberman [72] [1]
BN		2.0	1.75			0.25	9	Pawlewicz [82], Sainty [88]
CdI_2					2.9	5.0	60	Smith [68a]
CdSe				2.45	2.42	0.7	25	Borisevich [71]
CdS			2.42	2.30	2.27	0.5	18	Moss [59], Hall [55]
CdTe			2.80		2.64	0.86	28	Borisevich [71], Braunstein [78], Aranda [84]
C				1.9		2.0	12	Bubenzer [83] [2]
CaF_2		1.40	1.39		1.315	0.15	9	Traylor [92]
CeF_3	1.67	1.63	1.62	1.59	1.57	0.28	12	Hass [59], Pastor [78], Golubovic [76]
CeO_2		2.22				0.4	5	Seeser [75]
Chiolite		1.33				0.13	9	Canon [91] [3]
Cryolite		1.35				0.13	9	Heitmann [68a] §
DyF_3	1.55	1.53				0.20		Lingg [90]
Dy_2O_3		1.88				0.28		Dement'ev [77]
ErF_3	1.55	1.54				0.16		Lingg [90]
Er_2O_3	1.98	1.88	1.87			0.27		Marcinow [81], Dement'ev [77]
EuF_3						0.35		Lingg [90]
Eu_2O_3	1.97	1.86	1.84			0.27		Marcinow [81], Dement'ev [77]
GdF_3	1.61	1.59				0.28		Kolbe [91], Lingg [90]
Gd_2O_3	2.10	1.92	1.91			0.3		Marcinow [81], Truszkowska [76], Dement'ev [77]
GaSb			3.90	3.8		1.77	3.5	Borisevich [71]
Ge				4.20	4.20	1.7	25	Tomlin [76], Denton [72] [4]

The footnotes and explanations of the column headings appear on page 3-38.

[1] Lieberman [72] sputtered bismuth oxide.

[2] Hard carbon is a mixture of carbon and hydrocarbons. It is quite resistant to abrasion.

[3] The compositions of cryolite and chiolite are Na_3AlF_6 and $Na_5Al_3F_{14}$, respectively. Both materials are slightly water soluble.

[4] The film refractive index exceeds that of the bulk material.

Table 3-10 Optical properties of coating materials — continued

Film	Index in the wavelength (μm) region 0.35	0.55	1.0	1 - 8	10.0	Transp'cy Low	Up	Reference, note and/or comment
Ge_3N_4		2.11				0.8	9	Pawlewicz [82]
GeO_2		1.60				0.28	8	Pawlewicz [82]
HfF_4		1.57	1.55		1.46	0.25	12	Traylor [92]
HfO_2	2.15	2.02	1.98			0.25		Smith [79] [1] §
HoF_3	1.54	1.52				0.16		Lingg [90]
Ho_2O_3		1.96				0.25		Dement'ev [77]
InI_2				2.65		4.0	70	Smith [68a]
In_2O_3		1.95				0.35	1.5	Vainshtain [71], Vossen [77] [2]
In_2S_3				1.90		2.0	>20	Smith [68a]
LaF_3	1.68	1.58	1.57	1.52	1.52	0.14	12	Hass [59], Malherbe [74], Traylor [92] §
La_2O_3	1.95	1.90	1.82			0.26	11	Smith [79], Hass [59], Dement'ev [77], Kurdock [74] [3]
LiF	1.40					0.12	8	Haltom [79] [4]
$PbCl_2$			1.90			0.3	14	Koch [62b] *
PbF_2	1.80	1.75			1.73	0.25	18	Canon [91], Hass [59], Kraatz [78] [5] *
PbI_2		2.95	2.62			0.5		Milosklavskii [80], Buckman [75] [6]
$PbMoO_4$			2.08			0.4		Edwards [69] [7]
PbTe				6.0		4.0	40	Piccioli [74], Smith [68a]
LiF			1.40			0.12	9	Borisevich [71], Koch [62b]
LuF_3	1.52	1.48				0.16		Lingg [90]
Lu_2O_3		1.93				0.24		Dement'ev [77]
$MgAl_2O_4$			1.61			0.25		Pawlewicz [80]
MgF_2	1.39	1.38		1.36		0.13	10	Koch [62b], Borisevich [71], Kraatz [78] §
MgO	1.74	1.70	1.68	1.63		0.23	9	Bradford [72], Kraatz [78], Borisevich [71]
MoO_3		1.90						Canon [91]
NdF_3	1.68	1.60	1.58		1.60	0.25	12	Hass [59], Traylor [92]
Nd_2O_3	2.10	1.95	1.90			0.25	8	Hass [59], Dement'ev [77] [8]
Nb_2O_5		2.1-2.4				0.32	8	Dobrowolski [92], Aagard [75], Pawlewicz [82], Ondera [84] [9]

The footnotes and explanations of the column headings appear on page 3-38.

[1] Unless considerable (and expensive) effort is made to purify it, hafnia contains about 2% zirconia.
[2] A semiconductor; its reflectance in the IR depends upon the level of doping [Vossen 77].
[3] Dement'ev [77] notes that the film is "unstable in air."
[4] Haltom [85] deposited a layer and 48 hours later noted a change in refractive index from 1.32 to 1.40 .
[5] See Fig. 3#19. Kühn [91] measured the scattering of the films and observed their columnar structure.
[6] At a temperature of 150 C, Buckman [75] subjected the film to photolysis with light at 488 nm.
[7] Calculated from the measured reflectance of a stack of alternate layers of $PbMoO_4$ and cryolite.
[8] Dement'ev [77] cites an index of 2.5 in the visible.
[9] All authors prepared the film by sputtering. The higher refractive index is from Dobrowolski [92].

Table 3-10 Optical properties of coating materials — continued

	Index in the wavelength (μm) region					Transp'cy		
Film	0.35	0.55	1.0	1 - 8	10.0	Low	Up	Reference, note and/or comment
PrF_3	1.57	1.55				0.22		Lingg [90]
Pr_6O_{11}	2.0	1.9	1.83			0.35		Hass [59], Dement'ev [77] [1]
SmF_3	1.58	1.55				0.20		Canon [91], Lingg [90]
Sm_2O_3		1.90				0.34		Canon [91], Dement'ev [77]
Sc_2O_3	1.95	1.90	1.88			0.23	12	Rainer [82], Heitmann [73] §
Si			3.91	3.42		1.0	9	Pulker [64], Borisevich [71] [2]
Si:H				2.0		1.0	5	Perrin [79], Pawlewicz [82] [3]
SiC				2.95		1.5	11	Pawlewicz [82]
Si_3N_4		1.92				0.25	9	Netterfield [86], Pawlewicz [82]
SiO		2.00	1.90	1.85		0.5	8	Ritter [62, 71], Hass [54] [4]
Si_2O_3		1.57	1.55			0.3	8	Ritter [62, 71], Bradford [72]
SiO_2	1.48	1.46	1.45			0.16	8	Cox [64a], Eriksson [85] [5]
NaF	1.35		1.33		1.32	0.13	15	Borisevich [71], Baer [76b] §
SrF_2				1.34	1.36			Kurdock [74], Baer [76a]
Ta_2O_5	2.31	2.16	2.09			0.3	8	Smith [79], Pawlewicz [82], Aoki [92], Edlou [92]
TFE		1.42				0.8	6	Hollahan [74a, 74b] [6]
Te			5.00		4.30	1.0		Valeev [65]
TeO_2	2.20	1.90				0.3		Honcia [61], Lieberman [72]
TbF_3	1.55	1.53				0.23		Lingg [90]
Tb_2O_3		2.05				0.33		Dement'ev [77]
TlBr				2.20			40	Smith [68a] *
TlI				2.60			50	Smith [68a] *
ThF_4	1.53	1.52	1.51	1.48	1.35	0.16	13	Eastman [71], Heitmann [68b], Baer [76a] [7] * §
$ThOF_2$	1.55					0.18		Eastman [71] [1] *
ThO_2	1.82	1.78	1.77			0.25		Cox [64a][13] * §
TmF_3	1.53	1.50				0.16		Lingg [90]
Tm_2O_3		1.95				0.29		Dement'ev [77]
	Index in the wavelength (μm) region					Transp'cy		

[1] Dement'ev [77] notes that the film "interacts with water."

[2] Index in the near IR depends strongly upon the residual oxygen pressure during deposition.

[3] Pawlewicz [82] sputtered silicon in a hydrogen atmosphere. The refractive index ranges from 2.0 to 3.6, depending upon the hydrogen concentration in the film.

[4] An increase in the partial pressure of oxygen during deposition can decrease its refractive index to 1.70.

[5] Silica is highly absorbing at wavelengths near 10 μm. *n* and *k* may be extracted from Eriksson [85].

[6] Organic coating produced by plasma decomposition of a tetrafluorethlene monomer. Teflon AF is a trademark of Dupont and is deposited by solvent evaporation. Quaranta [93] sputtered it.

[7] Thorium compounds are mildly radioactive and even more so if the supplier has failed to remove the "daughter" decay products. Care should be taken in handling them.

Film	0.35	0.55	1.0	1 - 8	10.0	Low	Up	Reference, note and/or comment
SnO_2			2.00			0.4	1.5	Borisevich [71], Canon [91]
TiO_2		2.32	2.20			0.4	6	Pulker [76] [2]
WO_3		2.2				0.4		Nishide [95]
$Y_4Al_2O_9$		1.80	1.76			0.25		Pawlewicz [80]
YF_3	1.52	1.51	1.49	1.45		0.20		Jacobson [94] §
Y_2O_3	1.95	1.88	1.83			0.25		Heitmann [73], Marcinow [81], Borgogno [80] [3] §
YbF_3	1.56	1.53	1.48		1.48	0.20	11	Jacobson [94], Traylor [92], Lingg [90]
Yb_2O_3		1.95				0.28	11	Dement'ev [77]
ZnO		2.10	2.08	1.89		0.4		Borisevich [71], Canon [91]
ZnSe		2.65		2.42	2.42	0.55	18	Heitmann [65], Kraatz [78], Braunstein [78]
ZnS		2.36	2.27	2.24		0.4	14	Pelletier [82], Braunstein [78]
ZrO_2	2.15	2.05	2.00			0.3	8	Klinger [85], Pawlewicz [82], Edlou [92], Smith [79] [4] §

Table 3-10 Optical properties of coating materials — explanation of column headings and footnotes

Optical properties of thin films produced by evaporation in a vacuum or by sputtering. In a few instances other deposition methods, such as chemical, are used. The deposition method is unknown if Borisevich [71] is referenced.

"Transp'cy" refers to the lower (Low) and upper (Up) range of transparency. This range is relative, since it depends upon the coating process and film thickness. Within the "transparent" range, for example, there might be absorption band(s), as, for example, at 2.7 μm, due to water vapor.

Footnotes

* The extreme toxicity of a few of the aforementioned compounds is to be noted and appropriate safety precautions should be taken.

§ At $\lambda = 248$ nm, Rainer [85] and Lowdermilk [84] provide n, k values and laser damage data.

---

[1] Eastman [71] cites $n = 1.65$ at $\lambda = 160$ nm, $n = 1.55$ at $\lambda = 250$ nm.

[2] Index at 1 μm extrapolated using Eq. (2-2c).

[3] The mechanical stress of yttria has been reported as both tensive and compressive.

[4] Klinger [85] showed that a film in excess of 70 nm metric thickness manifests an inhomogeneous refractive index. Unless considerable (and expensive) effort is made to purify it, zirconia contains about 2% hafnia.

Chapter 4
Reflection reducing coatings

4.0 Introduction

Reduction of the reflectance R of a surface produces either of two effects: (1) If the material below the surface is opaque, the radiant absorption increases (see §8.5 and §4.4.3); or (2) if the material is transparent, such as a slab of glass, its overall transmittance is enhanced. For the latter effect, English lacks the proper verb equivalent to the Russian просветление *(prosvetleniya)*, which translates as "clarification," "brightening" or "to make more transparent." Although *antireflection* ("to be against reflection") is adequate as an adjective or noun, it is quite awkward as a verb. This void in the English language could be filled by the transitive verb *dereflect* — it is not distant from the German *Entspiegelung*. It could appear in a sentence: "SF-10 glass is dereflected at 633 nm by a two-layer antireflection coating." The verb stems from Latin: *reflectare* ("to reflect") and *de* ("removal of"). This and the acronym AR (for antireflection) are used in this chapter which surveys antireflection coating design. Procedures are emphasized and, where possible, mathematical details are relegated to appendices.

4.0.1 Development of the antireflection coating

The development of the antireflection coating began in the 1930s. Landmarks in its production are:

- 1939 Single-layer coatings of magnesium fluoride or calcium fluoride are deposited in production quantities at Carl Zeiss (Jena, Germany) and at Bausch and Lomb (Rochester, New York).
- 1941 - 1960 Vast numbers of optical surfaces are coated with single-layer antireflection coatings of magnesium fluoride.
- 1960 - 1980 Three-layer "quarter-half-quarter" antireflection coating (and its variants) slowly replace the single-layer antireflection coating. This coating was invented by Geffcken [40].
- 1980 - Multiple-layer antireflection coatings containing hard oxide layers are developed and mass produced.

4.1 Antireflection coating design by computer optimization

4.1.1 Can optimization produce an adequate antireflection coating?

Readers may find it unnecessary to delve into this chapter further than §4.1. Why? It is possible that they will discover that computer optimization (see §10.2) functions adequately as a design tool. In such an instance, the theory beyond §4.1 could be of only academic interest. Optimization of an antireflection coating often proceeds quite rapidly because usually there are relatively few layers.

4.1.2 Design procedure

Emancipation from antireflection design theory in the remainder of this chapter requires: (1) An optimization computer code and digital computer; (2) a willingness to be adventuresome and to explore new ideas; (3) manifest "street sense" and "intuition," which is often more valuable than analytical abilities; (4) an apotheosis of empiricism; (5)

an ability and willingness to deposit relatively thin layers, possibly with a quartz crystal (see §9.6.1.2) for thickness control.

4.1.2.1 Starting design for optimization

Freedom from antireflection design theory is achieved when the starting design is of the Coursen variety (see §10.2.4.5)

air **F** (**L H**)^q *substrate* .

A general rule is that at the top of the stack there should be a single layer **F** with as low refractive index as possible. Its thickness should be a quarterwave centered (in frequency — see §4.2.4) in the spectral region where the antireflection coating is to function. The remainder of the stack consists of many thin layers. The exponent q above should be increased until, after optimization, it is apparent that the defect [see Eq. (10-2a)] is no longer decreasing appreciably.

The procedure in the foregoing paragraph should be repeated with several indices of refraction — the lowest defect is selected. The examples below clarify this point.

4.1.2.2 Choices in optimization

It is preferable that targets of the defect function be defined at equal increments of *wave number* rather than *wavelength*. This is especially important when the span of wavelengths is large, such as 2:1. The optimization program should discard any layers of negative or zero thickness. After optimization, any ultrathin layers should: (1) be discarded; or (2) increased in thickness until they are producible.

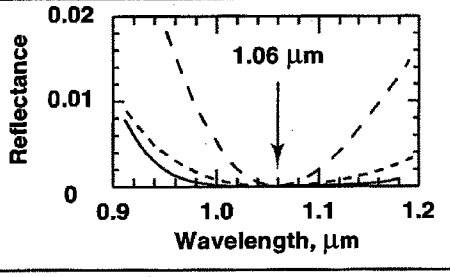


Fig. 4-1 $R(\lambda)$ of antireflection coating (*solid*)
air 1.149**L** 0.933**T** 0.354**L** 0.4345**T** *glass*, (*short dash*)
air 1.077**L** 1.191**H** 0.363**L** 0.319**H** *glass* and (*long dash*)
air 1.067**L** 1.079**S** 1.2**L** 0.443**S** *glass* .
Refractive indices of **T**, **H**, **S**, *glass* and **L** are, respectively, 2.25, 1.95, 1.62, 1.508 and 1.44. Optical thicknesses are λ_0 /4 at λ_0 of 1 μm.

quite low at the design wavelength.

4.1.3 Examples

4.1.3.1 Antireflection coating with relatively narrow spectral bandwidth

This antireflection coating is to dereflect BK-7 at 1.06 μm. Optimization targets are zero R at four equally spaced points between 1.03 μm and 1.09 μm.

4.1.3.1.1 The coating contains layers of titania and silica

The starting design consists of q periods of alternating silica and titania layers

air 0.1**L** (0.1**T** 0.1**L**)^q *glass* ,

each of which has an optical thickness of 25 nm. The refractive indices are 2.25 and 1.44. A single **L** layer remains after optimization unless q exceeds 5. When q is 6, the four layers remain after optimization. The solid curve in Fig. 4-1 shows that R is

Table 4-1 Comparison of narrowband antireflection coatings

ℓ	ℓ '	$\mathcal{D}$	Final design
17	2	0.008	*air* 1.059L 1.059S *glass*
25	3	0.002	*air* 1.059L 1.059S 1.059L *glass*
33	4	0.00057	*air* 1.067L 1.079S 1.12L 0.443S *glass*
41	6	0.00046	*air* 1.063L 1.067S 1.0775L 1.217S 0.234L 0.4714S *glass*

Defect $\mathcal{D}$ [see Eq. (10-2a)] of antireflection coatings with ℓ ' layers whose designs are listed above. Thicknesses and refractive indices are captioned in Fig. 4-1. ℓ is the layer count in starting design: *air* 0.1L $(0.1S\ 0.1L)^q$ *glass* .

4.1.3.1.2 The coating contains layers of hafnia and silica

Hafnia layers (index 1.95) are substituted for titania layers in §4.1.3.1.1 and the optimization program is again unleashed, with q equal to 6. Figure 4-1 (*short-dash* curve) shows that bandwidth of low reflectance is narrower than that of the titania-silica stack.

4.1.3.1.3 The coating contains layers of alumina and silica

Layers of alumina (index 1.62) and silica are used. Table 4-1 illustrates that the defect function steadily decreases as the layer count in the optimized design increases from 2 to 6. Figure 4-1 (*long dash* curve) shows that its bandwidth is narrower than those of the titania-silica or hafnia-silica stacks mentioned above.

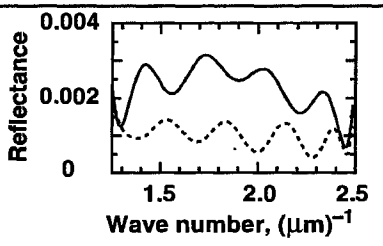


Fig. 4-2 $R(\sigma)$ of the antireflection coating (*solid*)
air 0.578F 0.364T 0.082L 0.584T 0.285L 0.1134T 0.683L 0.214T 0.171L 0.683T 0.112L 0.32T 0.284L 0.1T *glass,*
and (*short dash*)
air 0.540F 1.18H 0.29L 0.238H 0.256L 1.26H 0.191L 0.141H *glass* .
Refractive indices of T, H, *glass*, L and F are 2.25, 1.95, 1.52, 1.45 and 1.38, respectively. Optical thicknesses are λ_0 /4 at λ_0 of 1 μm.

4.1.3.1.4 Choice of design

Which design should be selected? This depends upon such factors as: (1) ability to control refractive index and thickness in a reproducible manner; (2) low absorption and/or high laser damage threshold of the coating materials; (3) cost of producing coating. In the last instance, reduction of the number of layers may or may not be important, depending upon the deposition process.

4.1.3.2 Antireflection coatings with relatively broad spectral bandwidth

The goal in the first example is low R over a 2:1 bandwidth on BK-7, using materials compatible in the visible portion of the spectrum. The band of low R extends from 1.25 $(\mu m)^{-1}$ to 2.5 $(\mu m)^{-1}$ — which corresponds to 400 nm to 800 nm. Starting designs are

air 0.5F $(0.1H\ 0.1L)^{16}$ 0.1H *glass*

and

air 0.5**F** (0.1**T** 0.1**L**)^{14} 0.1**T** *glass* ,

depending on whether hafnia **H** or titania **T** is used for the high index. Layer **F** is a quarterwave (at 500 nm) of magnesium fluoride and **L** is silica. Figure 4-2 depicts $R(\lambda)$ and Fig. 4-3 shows, for the hafnia-silica antireflection coating, thicknesses of layers — before and after optimization.

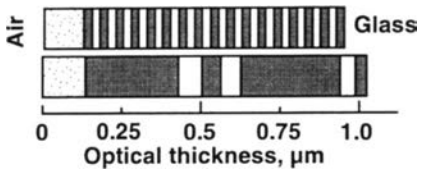


Fig. 4-3

Optical thicknesses of starting design (*upper*) and final design (*lower*) of the *short dash* antireflection coating captioned in Fig. 4-2.

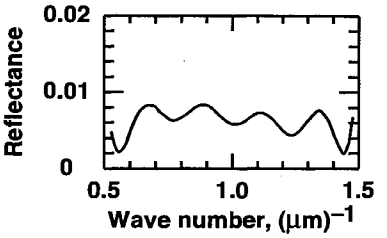


Fig. 4-4 $R(\sigma)$ of the antireflection coating

air 1.12**F** 0.417**H** 0.262**L** 1.3956**H** 0.135**L** 0.607**H** 0.479**L** 0.216**H** 2.2312**L** *glass* .

Refractive indices of **H**, *glass*, **L** and **F** are 2.10, 1.52, 1.45 and 1.38, respectively. Optical thicknesses are λ_0 /4 at λ_0 of 1 μm.

A further example is an antireflection coating covering a 3:1 bandwidth in the near IR and visible. The starting design contains 62 layers, which optimization reduced to nine. Figure 4-4 illustrates that its $R(\lambda)$ is below 1% over most of the band. Another example of optimization appears in §4.8.5.1.

4.1.3.3 Antireflection coating for a germanium substrate

The goal is to reduce R of a germanium slab in a wavelength span from 7.7 μm to 12.3 μm — the FLIR band. A restriction is that thorium fluoride — with its relatively low refractive index — may not be used for the outer layer. Although a few coating houses have substituted rare-earth fluorides such as lanthanum fluoride or yttrium fluoride for thorium fluoride in such coatings, a more difficult task is chosen here of using films of only germanium (**H**; n=4.2) and zinc sulfide (**L**; n = 2.2). The starting design has a low index outer layer

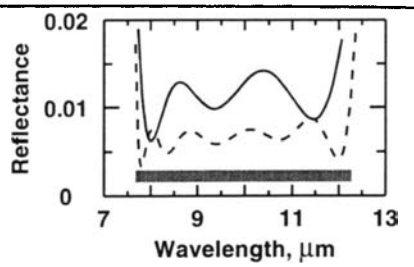


Fig. 4-5 $R(\lambda)$ of (*solid*)
air 0.959**L** 1.053**H** 0.305**L** 0.327**H** 1.156**L** 0.298**H** 0.301**L** *Ge*
and (*dashed*)
air 0.954**L** 1.014**H** 0.4613**L** 0.187**H** 1.0971**L** 1.097**H** 0.153**L** 0.522**H** 1.089**L** 0.431**H** 0.231**L** 1.09**H** 0.626**L** 0.109**H** 1.149**L** 0.399**H** 0.197**L** *Ge*
where refractive indices of **H**, *Ge* and **L** are 4.2, 4.0 and 2.2, respectively. Optical thicknesses are λ_0 /4 at λ_0 of 10 μm. The shaded band covers 7.7 μm to 12.3 μm.

air **L** (0.33**H** 0.33**L**)^q *germanium*
where quarterwave optical thickness is defined at 10 μm. Figure 4-5 shows the results of optimizing at 197 equally spaced points in wave number in the spectral range 810 cm^{-1} to 1300 cm^{-1}. A nine-layer starting design (q = 4 above) results in a seven-layer antireflection coating whose $R(\lambda)$ averages 1.2% over that spectral region. For the 17-layer design with lower $R(\lambda)$ in Fig. 4-5, its starting design has 31 layers. Aguilera [88] lists starting designs other than Coursen's and the results of optimizing them.

4.1.4 Antireflection coating that functions at nonnormal incidence

§4.9.5.3 mentions an antireflection coating that was produced by computer optimization. Its performance is quite superior to those produced by analytical methods.

4.1.5 Optimization or analytical design — which should be used?

The question remains: Should analytical or optimization design methods be used to produce antireflection coatings? As noted earlier, the user of optimization regards the computer code as a tool — albeit a quite complex one.

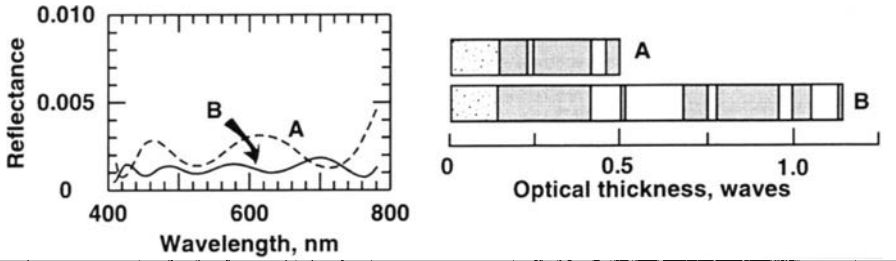


Fig. 4-6 (*Left panel*) $R(\lambda)$ of (*dashed*)
air 0.563**F** 0.3358**H** 0.0617**L** 0.682**H** 0.182**L** 0.1538**H** *glass*
and (*solid*)
air 0.536**F** 1.105**H** 0.3572**L** 0.0513**H** 0.6646**L** 0.2803**H** 0.101**L** 0.7083**H** 0.1717**L** 0.216**H** 0.3177**L** 0.0492**H** *glass*
where refractive indices of **H**, *glass*, **L** and **F** are 2.10, 1.52, 1.45 and 1.38, respectively. Optical thicknesses are λ_0 /4 at λ_0 of 1.0 μm.
(*Right panel*) Optical thicknesses (in waves at 1.00 μm) of the antireflection coatings.

The user of this tool can empirically investigate results produced by a given input. The antireflection coating with the 2:1 bandwidth in §4.1.3.2 serves as an example. The starting design is

air 0.5**F** (0.1**Z** 0.1**L**)^q 0.1**Z** *glass* .

where the refractive indices of the F, L and Z layers are 1.38, 1.45 and 2.10, respectively. $R(\lambda)$ becomes smaller as the number of layers in the starting design grows from six (q = 3) to 34 (q = 16).

Figure 4-6 depicts $R(\lambda)$ of a few antireflection coatings produced by optimizing the aforementioned starting designs. A doubling of the number of layers from six to twelve reduces rms reflectance fourfold. Figure 4-7 also shows that the total optical thickness of the "better" design is double that of the other antireflection coating.

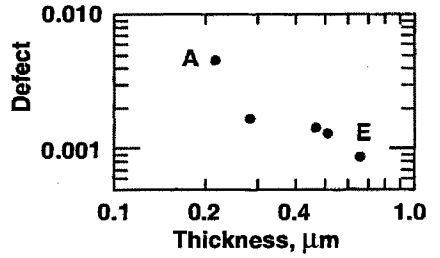


Fig. 4-7 Defect function [see Eq. (10-2a)] vs total optical thickness of antireflection coatings with low R over a 2:1 bandwidth. The designs are the same as those in Fig. 4-6, but with different numbers of layers.

4.1.5.1 Empiricism

The user of this complex "tool" — the computer — plots the rms defect function vs total optical thickness of coatings, as shown in Fig. 4-7. This hints at the number of layers needed to produce a given level of reflectance. This and other types of analyses lead the investigator to select the proper starting design. It is emphasized that data in Fig. 4-7 pertain to a *particular* computer code and set of optimization targets. Results completely unlike those in that Fig. might be produced from different targets and/or code. Willey [93] studied designs of antireflection coatings that minimized $R(\lambda)$ on BK-7 over a 2.25:1 bandwidth [Thelen 93]. His empirical data show, for example, the minimum number of layers needed to produce a certain level of rms R over that bandwidth.

Suppose that optimization produces useful antireflection designs. A question remains: Is a detailed knowledge of antireflection design needed? A quote from the English poet Thomas Gray (1716 – 1771) is appropriate: "...where ignorance is bliss, 'Tis folly to be wise."

4.2 Design methods and evaluation of antireflection coatings

4.2.1 Design methods Table 4-2 lists methods of antireflection coating design.

4.2.1.1 Choice of method

Preference of a designer for any design method may be purely personal or dictated by the ability of coating plants to produce: (1) refractive indices of arbitrary values — i.e., intermediate to those available (see §3.6.2); or (2) thin layers of precisely controlled thicknesses.

Example: The antireflection coating in §4.2.2.1 has an outer layer of index 1.34. This can be achieved only with a mechanically soft and water-soluble material such as cryolite. Hence, this design is unacceptable in many applications.

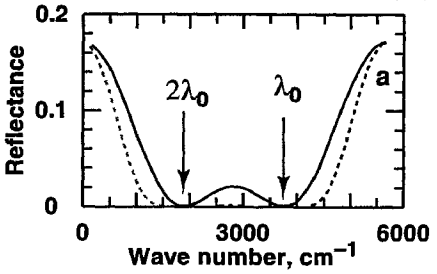


Fig. 4-8 $R(\sigma)$ of an antireflection coating (*dashed*)
air L_1 H_1 L_2 H_2 *ZnSe*
and (*solid*)
air L_1' H_1' *ZnSe* .
Refractive indices of L_1, H_1, L_2, H_2, L_1', H_1' and *ZnSe* are 1.11, 1.36, 1.77, 2.17, 1.34, 1.80 and 2.41, respectively. Optical thicknesses are λ_0 /4 at λ_0 of 3.533 μm.

4.2.2 Evaluation of an antireflection coating

Methods of evaluating antireflection coatings are:

- Zero R (or nearly zero, at least) is attained at one or more wavelengths.
- Average R over a prescribed wavelength region is minimized.
- Radiant absorption, as well as reflectance, is minimized.

4.2.2.1 Example: zero reflectance

An antireflection coating for the IR is deposited on a zinc selenide substrate ($n_s = 2.41$). It should transmit at both the fundamental frequency (5.30 μm = 1886 cm^{-1}) of a laser and at its doubled frequency (λ = 2.65 μm = 3774 cm^{-1}). There are many methods of producing zero R at these wavelengths. One is to deposit two layers of equal optical thickness

air L_1' H_1' *ZnSe* .

Using Eq. (4-12a) in §4.6.3.1, it is found that $n_{outer} = 1.34$ (sodium fluoride) for the outer layer and $n_{inner} = 1.80$ (silicon monoxide) for the inner. Fig. 4-8 depicts its $R(\lambda)$.

Table 4-2 Methods of antireflection coating design

Method	Reference	This book
Admittance (or impedance) matching	Muchmore [48], Blaisse [50]	§2.7.3.4, §4.2.3
Roots of admittance polynomial	Young [61], Knittl [76], Thelen [89]	
Partitioning of stack	Thetford [69], Rabinovitch [79]	§2.6.4, §4.4.4
R / T polynomial — its solution	Pohlack [52], Kard [56]	§2.7.4.3, §4.5.2.3
Graphical — vector addition of amplitudes	Geffcken [40], Anders [67b], Turner [50]	§2.8.3
Graphical — circle diagram	Berning [63], Rock [69], Apfel [72]	Fig. 4-21
Graphical — admittance chart	Macleod [86]	Fig. 4-21
Optimization	Aguilera [88], Willey [93]	§4.1

References to antireflection coating design methods in this book and elsewhere.

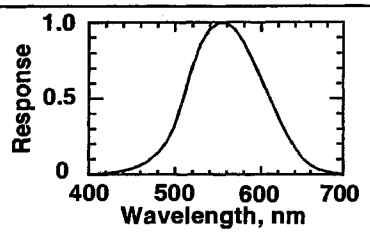


Fig. 4-9 Photopic response of the human eye — i.e., C.I.E. standard observer.

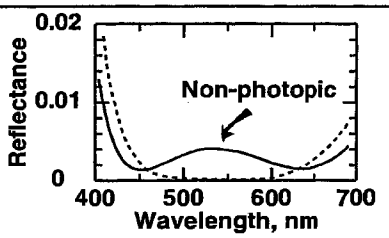


Fig. 4-10 $R(\lambda)$ of the antireflection coating (*dashed)*
air 99.3**F** 60.7**H** 12.1**L** 42**H** 91.8**S** 12**L** *glass*
and the non-photopic antireflection coating (*solid*)
air 92.2**L** 66.75**H** 10**L** 45.7**H** 38.4**L** 10**H** *glass* .
The refractive indices of **H**, **S**, *glass*, **L** and **F** are 2.10, 1.62, 1.52, 1.45 and 1.38, respectively. Layer thicknesses are metric in nm.

4.2.2.2 Example: antireflection coating with low photopic reflectance

An antireflection coating is deposited on BK-7 that must manifest the lowest photopic reflectance, i.e., R as perceived by the human eye when the antireflection coating is illuminated by light manifesting equal radiance at all wavelengths. This photopic response is shown in Fig. 4-9 and is the weighting factor [i.e., the Wi in Eq. (10-2a)] for computer optimization.

A modification of the $\mathbf{Q}_1\mathbf{D}\mathbf{Q}_3$ "quarter-half-quarter" coating (see §4.8.3) is used as one of the starting designs, with zirconium as the halfwave **D** layer and aluminum oxide as the $\mathbf{Q}_3$ layer contiguous to substrate. The modifications are: (1) A layer of silica is interposed between alumina layer and substrate so that the alumina does not react chemically with the substrate [Günther 81a]; (2) a thin silica layer is inserted in the center of the thick (i.e., halfwave) zirconia layer so that it does not become optically inhomogeneous as it grows thicker [Klinger 85]. The outer layer is magnesium fluoride.

The dashed curve in Fig. 4-10 shows that $R(\lambda)$ of the coatings rise considerably at 420 nm and at 670 nm. This does not add appreciably to the photopic reflectance, since the human eye is relatively insensitive at those wavelengths.

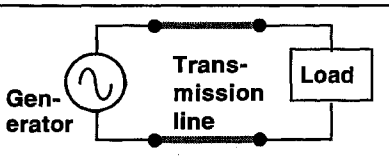


Fig. 4-11 A section of transmission line is interposed between a high frequency electrical generator and a load impedance.

4.2.2.2.1 Example: antireflection coating with green reflectance

The antireflection coating described in §4.2.2.2 contains magnesium fluoride and thus is better suited to be deposited by vacuum evaporation, rather than by sputtering. The reflected color of the coating is a deep magenta. The problem is that, in producing it, slight errors in layer thicknesses alter the reflected color from a reddish magenta to a bluish magenta, which makes the color difficult to control.

The solution is to deposit a coating that has is deliberately designed with a green reflected color, as shown by the solid curve in Fig. 4-10. Experience has shown that when layer thickness errors alter the color of such a coatings, it still retains it green

reflected color, as perceived by the human eye. This six-layer coating contains only two oxide coating materials and, therefore, is a candidate to be deposited by planar magnetron sputtering.

4.2.2.3 Both absorption and reflectance minimized

Baer [76b] developed a three-layer antireflection coating at 10.6 μm for potassium chloride substrates using the coating materials arsenic trisulfide **H** and sodium fluoride **L**. Although two layers

air $\mathbf{L_1\ H_2}$ *substrate*

(i.e., the "two-layer unequal thickness" system — see §4.3) would have produced low enough R, addition of a third layer

air $\mathbf{H_1\ L_2\ H_3}$ *substrate*

allowed thicknesses of the three layers to be reduced substantially, thereby lowering fourfold the absorption of this coating.

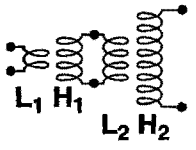


Fig. 4-12
Nonabsorbing layers of equal optical thickness
$\mathbf{L_1\ H_1\ L_2\ H_2}$
are modeled as an electrical circuit containing two cascaded lossless transformers.

4.2.3 Electrical impedance mismatch and optical coating

Impedance matching is used to design electrical matching networks [Collin 72]. Optical coatings are designed by analogy — see §2.2.5.3 and Young [63] or Knittl 76]. An illustration is in Fig. 4-11, which depicts a high frequency electrical signal generator with a characteristic impedance Z_0 (300 Ω in this example) driving a load of impedance Z_s (125 Ω). The section of electrical transmission line has an impedance Z_1. By conversion of each impedance into an admittance, this is equivalent to a generator of admittance 3.33×10^{-3} (i.e., reciprocal of 300) siemens and load admittance of 8.0×10^{-3} siemens. Application of the principle of scaling (see §2.4.3.4.1) allows us to divide both admittances by 3.33×10^{-3}. The result is that this is equivalent to a generator of normalized admittance of 1.00 driving a load of normalized admittance 2.41.

Following §2.2.5.3, the aforementioned generator and load are equivalent to a light source in air ($n_0 = 1.00$) emitting waves that impinge upon a substrate of index 2.41. Equation (2-12e) provides the R caused by the air-substrate (i.e., generator-load) admittance mismatch

$$R = \frac{(1-2.41)^2}{(1+2.41)^2} = 0.171 = 17.1\% \ . \qquad (4\text{-}1a)$$

This uncoated substrate reflects 17.1% of incident radiant flux. By analogy, the admittance mismatch in the electrical circuit reflects 17.1% of incident power back to the generator. Reflected power may be reduced by insertion of a transformer or alternatively by interposing a section of transmission line of proper impedance and length.

4.2.3.1 Antireflection viewed as impedance transformer

Reflected power is reduced by insertion of a transformer between load and generator. §2.7.3.4 shows that a sequence of equal thickness layers **H L** functions as an admittance transformer that produces a higher admittance. This is a *step-up* transformer. Reversal of the order to produce the sequence

$\mathbf{L}_1' \; \mathbf{H}_1'$

transforms admittance downwards — this is precisely the two-layer antireflection coating whose $R(\lambda)$ is graphed in Fig. 4-8. This is a *step-down* transformer. At a frequency where each optical thickness is 90°, it transforms substrate admittance of 2.41 downwards to 1.33, where it matches more closely that of incident medium, air. The foregoing result is obtained by substituting the aforementioned indices into Eq. (2-49d). The normalized admittance Y at λ_0 is

$$Y = (1.34)^2 \, (1.80)^{-2} \, 2.41 = 1.33 \ , \tag{4-1b}$$

which produces reflectance of

$$R = \frac{(1 - 1.33)^2}{(1 + 1.33)^2} = 0.02 = 2\% \ , \tag{4-1c}$$

which is well below the 17.1% in Eq. (4-1a).

The foregoing design is termed a *step-down* design because refractive index (i.e., optical admittance) *steps down* from a relatively high value of 2.41 to the inner layer ($n_2 = 1.80$) and thence to the outer layer ($n_1 = 1.34$), which is closer to incident medium index of 1.00.

4.2.3.1.1 Cascaded transformers and nonproducible refractive indices

Figure 4-8 (*dashed*) depicts $R(\lambda)$ of four layers of equal optical thickness. Its design is

air $\mathbf{L}_1 \; \mathbf{H}_1 \; \mathbf{L}_2 \; \mathbf{H}_2$ *substrate* .

Its electrical analog is a cascade of two step-down transformers, as depicted in Fig. 4-12. Compared to the two-layer antireflection coating, the four-layer has lower R and broader passband.

The problem is that refractive index 1.11 of its outer layer $\mathbf{L}_1$ is unobtainable — a problem first encountered by Muchmore [48]. Nevertheless, Epstein [52] simulated such low indices with equivalent layers. The step-up-step-down design (see §4.8) is an alternative that avoids inclusion of such low indices.

4.2.4 Symmetry of the reflectance vs wave number curve

Figure 4-5 plots $R(\lambda)$ (vs wavelength) and Fig. 4-8 graphs $R(\sigma)$ (vs wave number). Frequency (or wave number) is often preferred to wavelength for the reasons mentioned in §1.2.3. Another virtue is that if there is symmetry of a reflectance curve, it is evident when plotted in frequency.

Example: The coating in Fig. 4-8, which is centered at $\sigma_0 = 2830$ cm^{-1} contains only quarterwave layers and thus its transmittance can be found from Eq. (2-50b). A consequence is that R vs phase thickness is symmetrical about $\beta = 90°$. It follows that $R(\sigma)$ is symmetrical about σ_0, provided refractive indices of layers and substrate are nondispersive. Its "$2\lambda_0$" reflectance minimum is at 1887 cm^{-1} = (5.3 μm)$^{-1}$. The wave number difference is 2830 – 1887 = 943 cm^{-1}. Symmetry requires that the other zero of the R / T polynomial be located at 2830 + 943 = 3773 cm^{-1}.

4.3 Multiple quarterwave and other narrowband designs

4.3.1 Introduction

Coatings in this section attain nearly zero R over a relatively narrow bandwidth. They are often used in laser systems and are termed "vee-coats" in the coating industry. Coating design is often dictated by materials that resist laser damage.

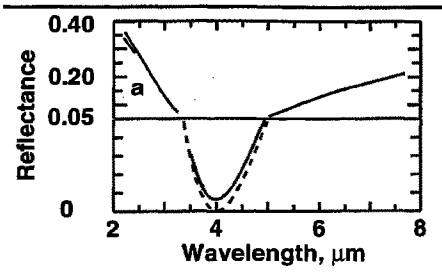


Fig. 4-13 $R(\lambda)$ of an antireflection coating (*solid*)
air **L** *Ge*
and (*short dash*)
air 1.64**Z** 0.68**T** 1.64**Z** *Ge* .
The refractive indices of *Ge*, **Z**, **L** and **T** are 4.0, 2.23, 1.85 and 1.48, respectively. Optical thickness are multiples of λ_0 /4 at λ_0 of 1 μm. The ordinate changes scale at 0.05.

4.3.2 Single layers

Equation (2-23j) specifies index n_1 of a layer of $\lambda_0/4$ optical thickness that produces zero R. Its thickness can also be an odd multiple of a quarterwave, such as $3\lambda_0/4$ or $5\lambda_0/4$. The thicker the layer, however, the narrower the spectral region over which R is relatively low. It was recognized many years ago that for BK-7 glass, $n_1 = \sqrt{1.52} = 1.23$ is an unattainable index. Use of MgF_2 ($n_1 = 1.38$) or cryolite ($n_1 = 1.35$) is a compromise. Fig. 1-23 shows that minimum R of close to 1.2% is attained.

Both silicon monoxide ($n_1 \approx 1.85$) and zinc sulfide ($n_1 \approx 2.23$) perform reasonably well in dereflecting germanium ($n_s = 4.0$) in the near IR, as illustrated in Fig. 4-13. At point *a* — near 2 μm — the layer is a halfwave in optical thickness and R approaches 36% — that of an uncoated substrate. An equivalent layer (see §2.6.7) with $n_1 = 2.00$ and 90° phase thickness (at 4 μm) is created using a rare earth fluoride and zinc sulfide. Figure 4-13 shows that its R is quite close to zero at 4 μm.

Philipp [70] decomposed silane in a plasma to deposit a single-layer antireflection coating of silicon nitride ($n_1 \approx 1.95$) on both gallium arsenide ($n_s \approx 3.8$) and silicon ($n_s \approx 3.8$). Ravesz [73] deposited single layers of Ta_2O_5 or Nb_2O_5 to match the surface of a silicon solar cell to the cover cement ($n_0 \approx 1.39$). Jupnik [58] deposited single-layer antireflection coatings of $ThOF_2$ on arsenic trisulfide glass using a thin PbF_2 layer to improve adhesion. Finally, designers should remember that an equivalent layer may be used at either normal or nonnormal incidence, as illustrated by Figs. 4-13 or 4-54.

One feature of the single-layer is important, especially when the antireflection coating is required to perform over a broad range of wavelengths and/or incidence angles [Pohlack 52a]:

> Reflectance (for randomly polarized flux) of any single-layer antireflection coating never exceeds that of the uncoated substrate, regardless of wavelength and incidence angle. The layer must be nonabsorbing.

The foregoing is not true for multiple-layer antireflection coatings in which the index of the films exceeds that of the substrate. This is illustrated in Fig. 1-39.

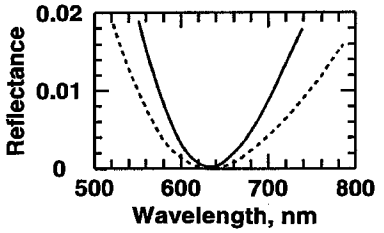


Fig. 4-14 $R(\lambda)$ of an antireflection coating (*solid*)
air **L T L** *glass*
and (*dashed*)
air **F S** *glass* .
The refractive indices of **S**, *glass*, **T**, **F** and **L** are 1.70, 1.52, 1.50, 1.38 and 1.35, respectively. The optical thicknesses are λ_0 /4 at λ_0 of 633 nm.

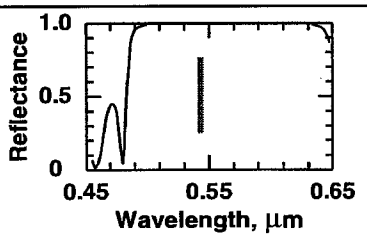


Fig. 4-15 $R(\lambda)$ of reflector (*solid*) captioned in Fig. 4-16.
The vertical line is at λ = 543 nm.

4.3.3 Multiple layers of quarterwave optical thickness

These designs are not widely used because the spectral region over which they function is relatively narrow. Examples of antireflection coatings that dereflect at λ_0 are

air $Q_1\ Q_2\ Q_1\ Q_2$ *substrate* (even layer count)

and

air $Q_1\ Q_2\ Q_1\ Q_2\ Q_1$ *substrate*(odd layer count),

where each layer has an optical thickness of $\lambda_0/4$.

As an example, suppose a three-layer coating with films of index 1.35 (adjacent to air and substrate) and 1.50 (center layer) are applied to a BK-7 substrate. Its reduced admittance y at λ_0 [by the use of Eq. (2-48c)] is

$$y = \frac{(1.35)^2\ (1.35)^2}{(1.50)^2\ 1.517} = 0.97\ , \qquad (4\text{-}1d)$$

which produces a reflectance well below 0.1%, as plotted (*solid*) in Fig. 4-14.

4.3.3.1 Double quarterwave

In the 1960s, the double quarter antireflection coating was deposited on optical components in laser systems. The procedure is to choose the outer layer index to be as low as possible and find the inner layer index from

$$n_2 = n_1 \sqrt{n_s / n_0}\ . \qquad (4\text{-}2a)$$

As an example, if n_1 = 1.38 (magnesium fluoride), a layer of index n_2 = 1.70 (alumina) is required for BK-7. Figure 4-14 (*dashed*) shows its $R(\lambda)$. There are several problems in producing this coating. The first is that quite high temperatures must be used in a coating chamber to achieve the higher index of 1.70 for the alumina layer. A significant drawback is that MgF_2 has a relatively low laser damage threshold.

R at λ_0 of double quarterwave — or any multiple quarterwave coating with an even number of layers — depends upon only the *ratio* of layer indices. Indices of 1.45 and 1.78 could be used, for example, in lieu of the 1.38 and 1.70 mentioned in the previous paragraph. The higher indices, however, the narrower the bandwidth of low reflectance.

4.3.3.2 Reflector with low reflectance at single IR wavelength

A total reflector (see §1.7.1.2) is designed for a green helium-neon laser with R at λ_0 = 543 nm of at least 99.95%. R should be less than 0.2% at 3.39 μm. This low R is required because of competition of transitions between energy levels in the neon gas at 543 nm and 3.39 μm.

The high R is achieved by depositing alternating high and low refractive indices on BK-7 (see §5.2.4.3.1). It is tempting to produce low R at 3.39 μm by laying down one

or more thick layers at the top of stack. This should be avoided because thick layers increase absorption at 543 nm. An alternative design is

air $(\mathbf{E}_L\ \mathbf{E}_H)^2$ *glass*

where $\mathbf{E}_L$ and $\mathbf{E}_H$ are equivalent layers (see §2.6.7) with phase thickness of 90° at 3.39 μm. At that wavelength, their indices are 1.74 and 1.90, respectively. In other words, this is the $\mathbf{Q}_1\ \mathbf{Q}_2\ \mathbf{Q}_1\ \mathbf{Q}_2$ design of §4.3.3.

What is its R at 3.39 μm? Using Eq. (2-48b), reduced admittance y is

$$y = \frac{(1.74)^2 \times (1.74)^2 \times 1.52}{(1.90)^2 \times (1.90)^2} = 1.07\ , \qquad (4\text{-}2b)$$

which results in R of below 0.2%. Figure 4-15 shows that this stack also has high reflectance at 543 nm.

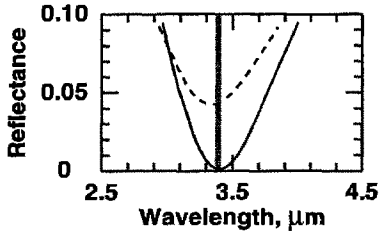


Fig. 4-16 $R(\lambda)$ of (*solid*)
air $[(0.612\mathbf{L}\ 0.816\mathbf{H}\ 0.612\mathbf{L})^3$
$(0.408\mathbf{L}\ 1.224\mathbf{H}\ 0.408\mathbf{L})^3]^2$ *glass*
and (*dashed*)
air $(0.5\mathbf{H}\ \mathbf{L}\ 0.5\mathbf{H})^{12}$ *glass*
where refractive indices of **H**, *glass* and **L** are 2.25, 1.52 and 1.45, respectively. Optical thicknesses are $\lambda_0/4$ at λ_0 of 543 nm. The vertical line is at λ = 3.39 μm.

The procedure of §2.11.3 synthesizes a design for $\mathbf{E}_L$ as

$(0.612\mathbf{L}\ 0.816\mathbf{H}\ .612\mathbf{L})^3$

where optical thickness of **L** and **H** is a QWOT at 543 nm. The design of $\mathbf{E}_H$, as well as the entire stack, appears in the caption to Fig. 4-16. This Fig. shows $R(\lambda)$ of the aforementioned multilayer, as well as $R(\lambda)$ of an "edge filter" type design (see §5.4.3) with the same number of layers. The latter has considerably higher R at 3.39 μm. In practice, the foregoing design is only an approximation because no attention is given to dispersion of either substrate or coating materials.

4.3.4 Two layers of unequal optical thickness

Previous sections discuss coatings with prescribed optical thicknesses; the task is to find refractive indices needed to produce an antireflection coating at λ_0. The opposite is now considered. Indices of two layers are given and their thicknesses are adjusted to attain R = 0 at λ_0, using both vector and analytical design methods. Such antireflection coatings are widely used in laser systems.

4.3.4.1 Design using vector addition of amplitudes

Design using vector addition of amplitudes is presented first because it is intuitive. The procedure is:

- Choose two indices — subject to restrictions mentioned in §4.3.4.2.1. If the lower index of those materials is less than the square root of substrate index, then that layer may, or may not, be adjacent to the substrate. Otherwise, the film with the higher of the two indices must be contiguous to substrate.
- Compute Fresnel amplitude reflection coefficients at each interface using Eq. (2-8f).
- There are three interfaces and hence three vectors. They form a triangle, provided R = 0. Determine the internal angles of each triangle.

- Convert each angle in the foregoing step into a phase thickness β of a layer, remembering that 2β is represented in a vector diagram (see §2.8.3.2).
- Convert phase thickness into optical thickness using Eq. (2-4d).

Example: Figure 4-17 (*right panel*) shows that films with indices of 1.44 and 1.92 are deposited on substrate with $n_s = 1.507$. Vector triangles appear in left panels of Fig. 4-17. In solution *a*, the outer layer is thinner than a quarterwave and the inner is thicker. The converse is true for solution *b*.

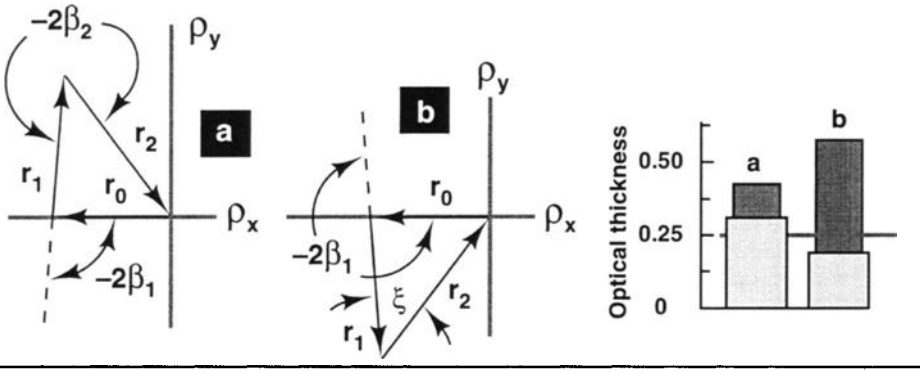


Fig. 4-17 (*Panels a* and *b*) Vector addition of amplitudes calculation of ρ of two-layer antireflection coatings, with $\rho = 0$. Outer layer index is 1.44 and inner index is 1.92. Substrate index is 1.507. Angle ξ is "$-2\beta_2$." (*Right panel*) Optical thicknesses of the layers in the antireflection coatings.

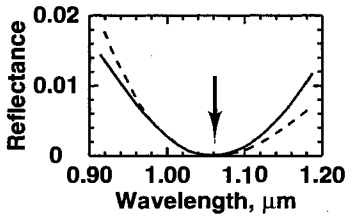


Fig. 4-18 $R(\lambda)$ of (*solid*)
air 0.76**L** 1.54**H** *glass*
and (*dashed*)
air 1.24**L** 0.46**H** *glass*
where refractive indices of **H**, *glass* and **L** are 1.92, 1.507 and 1.44, respectively. Optical thicknesses are λ_0 /4 at λ_0 of 1.06 μm.

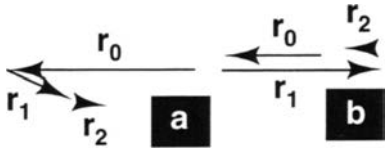


Fig. 4-19 (Panel *a*) Amplitude reflection coefficients at each interface r_0, r_1 and r_2 for the sequence of refractive indices 1.00/1.80/2.24/2.41. (Panel *b*) The sequence of indices is 1.00/1.35/2.24/2.41.

Angles in Fig. 4-17 are translated into layer thicknesses. Figures 2-84 and 2-85 show that the "$-2\beta_2$." may actually be a complementary angle, depending upon the sign of the Fresnel coefficient. In Fig. 4-17, only one of the internal angles is an actual double phase thickness. This is illustrated by the fact that in both designs, the complementary angle is the double phase thickness of the outer layer.

Vector design provides insight into the sensitivity of reflectance to errors in layer thickness. Thicknesses of both layers are smaller in the left panel of Fig. 4-17. This design is, therefore, less sensitive to percentage errors in layer thicknesses. Figure 4-18 depicts $R(\lambda)$ of those two antireflection coatings at a wavelength center of 1060 nm.

4.3.4.2 Numerical design procedure: thicknesses of the layers

Symbols n_0, n_1, n_2 and n_s refer to the indices of the incident medium, outer layer, inner layer and substrate, respectively. n_L and n_H refer to the lower and higher of the two indices, respectively, regardless of the order in which they are deposited. §4.3.4.2.1 and §4.3.4.2.2 contain a procedure for finding layer thicknesses.

4.3.4.2.1 Existence of a solution

No solution exists if certain constraints on refractive indices are violated. This is illustrated by panel *a* in Fig. 4-19, which shows a vector diagram for a coating with indices *air*/1.80/ 2.24/2.41, where 2.41 is the substrate index. Zero R cannot occur because that vector triangle fails to close — r_0 is simply too long compared to other vectors in panel *a*. Panel *b* shows vectors for the sequence *air*/1.35/2.24/2.41. No solution exists because of the excessive length of r_1. Grebenshchikov [46] and Schuster [48] derived constraints for the indices. These constraints are depicted graphically in Fig. 4-20 for n_s of 2.41.

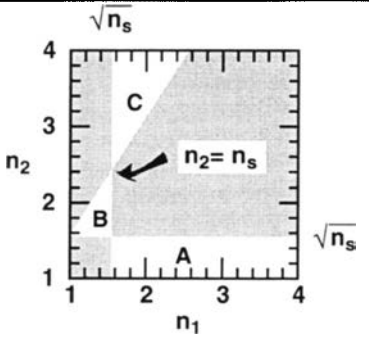


Fig. 4-20 For an incident medium of air and two nonabsorbing layers deposited upon substrate of index 2.41, regions *A*, *B* and *C* (all nonshaded) are allowed refractive indices for zero R. The index n_1 is next to air. After Schuster [48].

Forbidden regions for indices are shaded. Given those indices, constraints are:

1. Either n_1 or n_2 must exceed $\sqrt{n_s\, n_0}$.
2. If $n_2 < \sqrt{n_s\, n_0}$, then n_1 must exceed $\sqrt{n_s\, n_0}$
3. If $n_1 < \sqrt{n_s\, n_0}$, then an upper bound for n_2 is n_b

$$n_b = n_1 \sqrt{n_s / n_0} \ . \qquad (4\text{-}3)$$

4. If $n_1 > \sqrt{n_s\, n_0}$, then a lower bound for n_2 is given in Eq. (4-3).

Example: For the indices listed in the caption to panel *b* in Fig. 4-19, $\sqrt{n_s}$ is 1.55 and hence $n_1 < \sqrt{n_s}$ when $n_1 = 1.35$. The upper bound for n_b in Eq. (4-3) is 2.10. A solution does not exist because n_2 of 2.24 exceeds 2.10.

4.3.4.2.2 Computation of optical thicknesses of the layers

There are several methods of finding layer thicknesses. For example, Cox [64b] expresses them as arccosines of phase angles. Equations below follow from Kard theory in §2.6.5.2, as shown in an appendix, §4.10.1. Only those solutions are considered in which the thickness of at least one of the layers is less than a quarterwave at λ_0.The procedure below produces zero R at λ_0 :

(1) Compute the argument

$$\kappa \equiv \frac{(n_0 - n_s)\,(n_s\, n_0 - n_2^2)\, n_1^2}{(n_1^2\, n_s - n_2^2\, n_0)\,(n_s\, n_0 - n_1^2)} \quad . \tag{4-4}$$

No solutions for thicknesses exist if κ is negative.

(2) Assuming $\kappa > 0$, angles ξ_1 and ξ_2 (in radians) are found from

$$\xi_1 = \arctan(\sqrt{\kappa}) \qquad \text{and} \tag{4-5a}$$

$$\xi_2 = \arctan\left[\frac{(n_s - n_0)\, n_1\, n_2}{(n_1^2\, n_s - n_2^2\, n_0)\,\sqrt{\kappa}}\right] . \tag{4-5b}$$

(3) Compute

$$y' = n_L^2\; n_s / (n_H^2\; n_0) . \tag{4-5c}$$

(4) If y' is less than 1.00, then go to step (7). If y' exceeds 1.00, execute steps (5) and (6), but do not execute step (7).

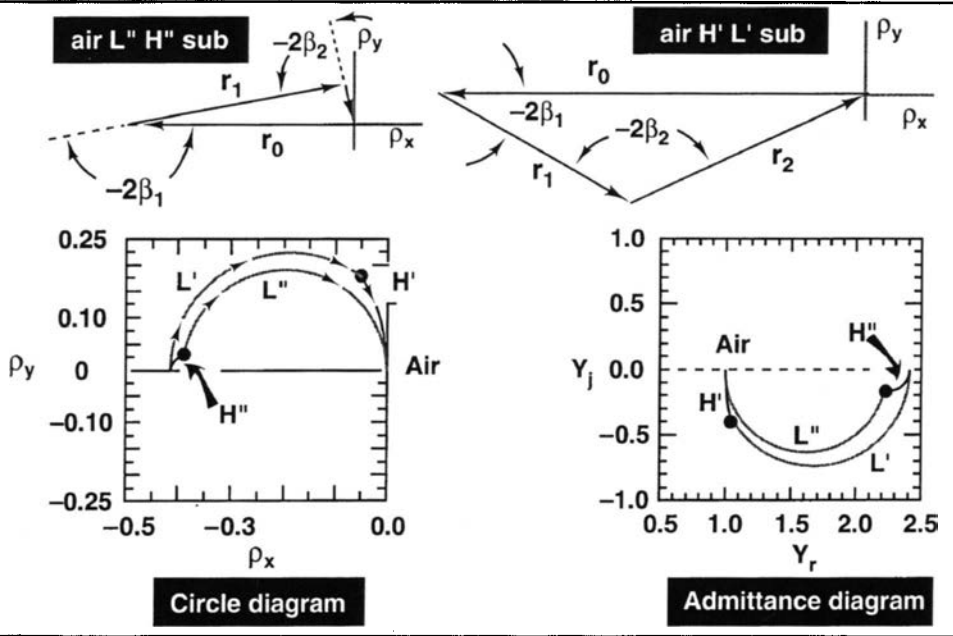


Fig. 4-21 At wavelength λ_0, vector diagrams (*upper panels*), circle diagram (*lower left* — see §10.1.2.3) and admittance diagram (*lower right* — see §10.1.2.2) of the antireflection coatings *air* **L" H"** *ZnSe* and *air* **H' L'** *ZnSe* where the indices of **L**," **H**," **L**, **H'** and *ZnSe* are 1.50, 2.24, 1.50, 2.24 and 2.41, respectively. Optical thickness of **L**," **H**," **H'** and **L'** is 0.128 λ_0, 0.236 λ_0, 0.033 λ_0 and 0.185 λ_0, respectively. The solid black dot is at the interface between layers.

(5) If $n_0 < n_s$, set $n_1 = n_H$ and $n_2 = n_L$, so that the lower index layer is adjacent to the substrate. Compute angles in step (2). The layer thicknesses are

$$\text{QWOT}]_1 = 2\,\lambda_0\,\xi_1\,/\pi \quad \text{or} \quad n_1\, h_1 = \lambda_0\,\xi_1\,/(2\,\pi) \tag{4-5d}$$

$$\text{QWOT}]_2 = 2\,\lambda_0\,\xi_2\,/\pi \quad \text{or} \quad n_2\, h_2 = \lambda_0\,\xi_2\,/(2\,\pi) \; . \tag{4-5e}$$

(6) Swap layer indices, so that $n_2 = n_H$ and $n_1 = n_L$, so that the higher index is adjacent to the substrate. Re-execute steps (2) and (5), which provides the second solution.

(7) There are two solutions. The higher index layer is adjacent to the substrate. One solution is

$$\text{QWOT}]_1 = 2\,\lambda_0\,(\pi - |\xi_1|)/\pi \quad \text{or} \quad n_1 h_1 = \lambda_0\,(\pi - |\xi_1|)/(2\,\pi) \qquad (4\text{-}5f)$$

$$\text{QWOT}]_2 = 2\,\lambda_0\,|\xi_2|/\pi \quad \text{or} \quad n_2 h_2 = \lambda_0\,|\xi_2|/(2\,\pi) \qquad (4\text{-}5g)$$

and another solution is

$$\text{QWOT}]_1 = 2\,\lambda_0\,|\xi_1|/\pi \quad \text{or} \quad n_1 h_1 = \lambda_0\,|\xi_1|/(2\,\pi) \qquad (4\text{-}5h)$$

$$\text{QWOT}]_2 = 2\,\lambda_0\,(\pi - |\xi_2|)/\pi \quad \text{or} \quad n_2 h_2 = \lambda_0\,(\pi - |\xi_2|)/(2\,\pi)\,. \qquad (4\text{-}5i)$$

Example one Figure 4-21 depicts vector diagrams, admittance diagrams and circle diagrams for layers with indices 1.50 and 2.24 deposited upon a zinc selenide substrate ($n_s = 2.41$). According to step (7) in above,

$$y' = n_L^2\; n_s/(n_H^2\; n_0) = (1.50)^2\,(2.24)^{-2}\,(2.41) = 1.08\;. \qquad (4\text{-}5j)$$

Since y' exceeds 1.00, there are both "**H**-outer & **L**-inner" and "**L**-outer & **H**-inner" solutions, as given in steps (5) and (6) above. These solutions lie in areas *A* and *B* in Fig. 4-20. Figure 4-22 graphs $R(\lambda)$. Stack center is at 3.33 μm.

Example two If a Schuster plot similar to Fig. 4-20 were drawn for a substrate of index 1.507, then layer indices in Fig. 4-17 would lie in region *C* of that Fig. A consequence is that the higher index must be adjacent to the substrate. Thicknesses are given in step (7) above — one optical thickness is always greater than a quarterwave and the other is less than a quarterwave.

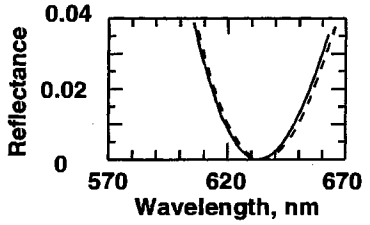


Fig. 4-23 $R(\lambda)$ of an antireflection coating (*dashed)*
air 0.32**L** 0.336**H** 1.32**L** 0.664**H** *metal*
and (*solid*)
air 0.403**H** 1.38**L** 0.664**H** *metal*
L and **H** represent indices of 1.45 and 2.30, respectively. Optical thicknesses are $\lambda_0/4$ at λ_0 of 633 nm. Optical constant of *metal* is $\hat{n} = 2 - j\,3$.

4.4 All-dielectric antireflection coatings deposited upon metallic layers

4.4.1 Introduction

Foregoing sections use dielectric layers to reduce *R* of a dielectric (i.e., nonabsorbing) substrate. The next task is to dereflect a metallic "under layer" at wavelength λ_0. This layer need not be opaque, provided its admittance can be computed or measured. Although it is possible to use thin absorbing layers in a *dark mirror* design (see §8.5), antireflection coatings containing only dielectric layers are considered in Chap. 4.

4.4.2 Design procedure

These steps are executed:

(1) Select two materials for this antireflection coating — designated as high index layer **H** and low index **L**.

(2) Deposit an **H** layer on the metal. The thickness of **H** is obtained from Eq. (2-46a) so that normalized admittance Y_1 of overcoated metal is pure real.

(3) Use Eq. (2-46b) to compute normalized

admittance of this overcoated system. In most cases, the negative root in Eq. (2-46b) is used, which results in a normalized admittance of less than one. This is termed the "pseudo-substrate."

(4) Add as many layers as necessary to dereflect this pseudo-substrate. Since it has an "index" (i.e., normalized admittance), that is pure real, any antireflection coating described in §4.2 or §4.3 may be used.

4.4.3 Antireflection coating deposited upon a transition metal

Optical constant of opaque substrate is $\hat{n}_s = n_s - jk_s = 2 - j3$, which is typical of a transition metal such as chromium or nickel. It is overcoated with a layer of index 2.30. §4.4.4.2 explains that it is preferable to use a high index (as compared to a low index) next to metal.

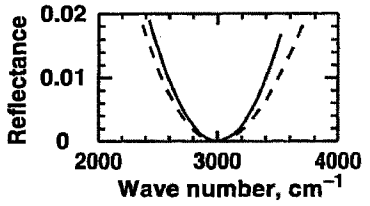


Fig. 4-22 $R(\sigma)$ of an antireflection coating (*dashed*)
air **L**" **H**" *ZnSe*
and (*solid*)
air **H**' **L**' *ZnSe*
where indices and thicknesses appear in the caption to Fig. 4-21.

Since the metal is opaque, its normalized admittance is its optical constant. The following substitutions are made in Eq. (2-46a): $n_s \rightarrow Y_r$ and $-k_s \rightarrow Y_j$. Execution of steps (2) and (3) in §4.4.2, Eq. (2-46a) produces optical thickness of 0.166 waves that converts $\hat{n}_s$ to a pure real admittance of $Y_1 = 0.621 + j\,0$. The next step is to dereflect at λ_0 this pseudo-substrate of "index" 0.621.

Solution one: A double quarter antireflection coating is used (see §4.3.3.1). As discussed in §2.7.3.2 and §2.7.3.4, a layer of higher index (which is adjustable) must be adjacent to air for the purpose of stepping upwards (to air) this relatively low pseudo-substrate "index" of 0.621. An **L** layer is adjacent to the pseudo-substrate. Solving Eq. (4-2a) for the *outer* layer index provides

$$n_1 = n_2\, n_s^{-1/2} = 1.45\,(0.621)^{-1/2} = 1.84\ . \qquad (4\text{-}6a)$$

The foregoing is an equivalent layer (see §2.6.7). $R(\lambda)$ of the final antireflection coating appears (*dashed*) in Fig. 4-23.

Solution two: A two-layer unequal-thickness antireflection coating (see §4.3.4.2.2) is applied to the overcoated substrate, using indices of 1.45 and 2.30. Figure 4-23 (*solid*) graphs its $R(\lambda)$.

4.4.4 Antireflection coating deposited upon a semitransparent metal

A semitransparent silver layer of 30 nm metric thickness is deposited upon a glass substrate. It is dereflected at $\lambda_0 = 650$ nm, where the optical constant of silver is approximately $0.172 - j\,4.196$. Refractive indices of **L** ($n_L = 1.35$) and **H** ($n_H = 2.25$) are identical with those of Table 8-3. This antireflection coating is designed by the "spacerless" partitioning method (see §2.6.4.3).

Solution: The following steps are executed:

(1) Position a high index layer next to the silver layer **M**. Preference of high index over low index is justified in §4.4.4.2.

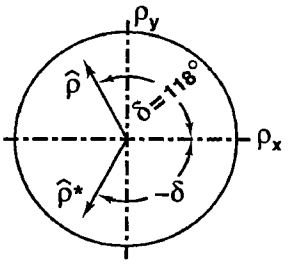


Fig. 4-24 At 650 nm, amplitude reflection coefficient $\hat{\rho}$ (and its complex conjugate) of the single layer
medium **M** *glass*
where refractive indices of incident *medium* and *glass* are 2.25 and 1.52, respectively. Optical constant of **M** is $0.172-j\,4.196$ and its thickness is 30 nm.

(2) R and phase shift upon reflection δ are computed for
pseudo-incident **M** *glass* .
The *pseudo-incident* medium index is 2.25, same as that of layer $\mathbf{H}_1$ that is contiguous to silver when the design is completed. A computer code (based upon theory in §2.4.4) is used to find R and δ, with the result: $R = 0.748$ and $\delta = 118°$. The amplitude reflection coefficient $\hat{\rho}$ (measured *inside* medium of index 2.25) is

$$\hat{\rho} = \sqrt{0.748}\ \exp(j\ 118°) = 0.865\ \exp(j\ 118°) \quad \text{.(4-6b)}$$

Fig. 4-24 represents $\hat{\rho}$ as a vector in the amplitude reflection plane.

(3) Design a dielectric stack that manifests amplitude reflection coefficient $\hat{\rho}^*$ This is the complex conjugate of $\hat{\rho}$ found in step (2) above. Its design is
pseudo-incident stack *emergen*
where *pseudo-incident* has the same meaning as above and *air* is the emergent medium.

(4) The stack mentioned in step (3) should manifest a target reflectance of 74.8%, which translates into a standing wave ratio (see §2.3.2) V of

$$V = \frac{1+\sqrt{0.748}}{1-\sqrt{0.748}} = 13.8\ . \qquad (4\text{-}6c)$$

(5) As a trial design, consider the quarterwave stack
pseudo-incident **L H L H** *emergent* ,
which at λ_0 has a reduced admittance [using Eq. (2-48b)] of

$$y = (1.35)^4\ (2.25)^{-3} = 0.0576 = (17.4)^{-1}\ . \qquad (4\text{-}6d)$$

Its V of 17.7 exceeds the target of 13.8. This leads us to
pseudo-incident **L H L E** *emergent*
where the index of equivalent layer **E** is adjusted so that its V is precisely 13.8.

$$\frac{n_L^4}{n_H^3\, n_E^2} = \frac{(1.35)^4}{(2.25)^3\, n_E^2} = (13.8)^{-1}\ . \qquad (4\text{-}6e)$$

Solution of Eq. (4-6e) for n_E leads to a value to 2.01. This equivalent (of 90° phase thickness) is synthesized using the procedure of §2.11.3. The phase shift upon reflection of the above stack is 0° — a fact used in step (6) below.

(6) Add layer $\mathbf{H}_1$ to the top of stack. Since $\mathbf{H}_1$ has the same refractive index as *pseudo-incident*, this does not alter its R. The design of the partition is
pseudo-incident $\mathbf{H}_1$ **L H L E** *emergent* .

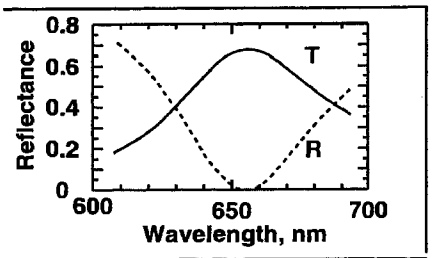


Fig. 4-25 $R(\lambda)$ (*solid*) and $T(\lambda)$ (*dashed*) of the metal-dielectric stack

air 0.208**L** 0.546**H** 1.208**L** **H** **L** 0.65**H** **M** *glass*

where indices of **L**, *glass* and **H** are 1.35, 1.52 and 2.25, respectively. Optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 650 nm. Optical constant of **M** is $\hat{n} = 0.172 - j\,4.196$ and its metric thickness is 30 nm.

The phase thickness β_1 of $\mathbf{H}_1$ is chosen so that phase shift upon reflection of the foregoing is 360° – 118° = –118° = 242°, as mandated by Eq. (2-45i). Remembering that phase thickness and δ are rotations in *opposite* directions, we have

$$-2\beta_1 = -118° \quad . \tag{4-6f}$$

Thus the fractional *optical thickness* of $\mathbf{H}_1$ is 59°/360° = 0.164 at λ_0 and its QWOT is 0.65. Amplitude reflection coefficient $\hat{\rho}$ of the foregoing stack appears in Fig. 4-24 and is the complex conjugate of the silver immersed in a medium of high index material.

(7) The order of the layers of the stack in step (6) is reversed and then laid on top of the silver. Figure 4-25 shows $R(\lambda)$ of the final design

air **E** **L** **H** **L** $\mathbf{H}_1$ **M** *glass* .

This coating is resurrected in §8.2.4.4.

4.4.4.1 Reflectance and transmittance of a dereflected metal

Rules mentioned in §2.9.2 are illustrated by dereflection of this semitransparent absorbing film. If reflectance R_1 and transmittance T_1 of the uncoated silver are known, transmittance T_2 of silver covered by an antireflection coating may be calculated quite easily. The following proof is based upon net flux ratio Ψ.

A computer program based upon the theory in §2.4.4.3 is used to find R_1 (= 67.4%) and T_1 (= 26%) of the uncoated silver. Starting from Eq. (2-55f), Ψ_1 of silver is

$$\Psi_1 = \frac{T_1}{1-R_1} = \frac{0.260}{1-0.674} = \frac{0.260}{0.326} = 0.80 \quad . \tag{4-6g}$$

Ψ_2 is the net flux ratio of silver overcoated with the six-layer dielectric stack (see caption to Fig. 4-25). From the multiplicative properties of Ψ (see §2.9.2.1),

$$\Psi_2 = \Psi_s \Psi_1 \tag{4-6h}$$

where Ψ_s is the net flux ratio of the dielectric stack. According to §2.9.2.2, Ψ_s is 1.00. In other words, addition of nonabsorbing layers "0.208**L** 0.546**H** 1.208**L** **H** **L** 0.65**H**" (see caption to Fig. 4-25) does not alter the net flux ratio Ψ_1. From Eq. (4-6h) it follows that

$$\Psi_2 = \Psi_1 = \frac{T_2}{1-R_2} = 80\% \tag{4-6i}$$

where R_2 and T_2 are reflectance and transmittance of silver *with* the antireflection coating. Since R_2 of the dereflected stack is close to zero, T_2 is 80%, since Ψ_2 equals T_2 when R_2 is 0.

Fig. 4-26 summarizes the partitioning of R, T and A for bare silver and overcoated silver. When reflectance R_1 of 67% of the silver is reduced to (nearly) zero by the antireflection coating, the 67% is partitioned between increased *transmittance* and *absorption* of the dereflected stack. This dereflection causes *both* T and A to increase threefold.

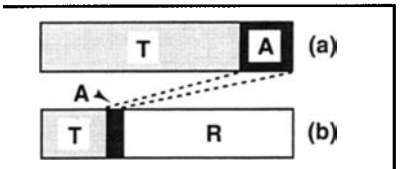


Fig. 4-26 At λ_0, reflectance R, transmittance T and absorption A of: (*a*) the antireflection coating captioned in Fig. 4-25; and (*b*) the single **M** layer captioned in Fig. 4-25 is deposited on glass.

4.4.4.2 Advantage of using a high index layer next to the metal

Both antireflection coatings in §4.4 are designed with an **H** layer next to the metal. Why? In fact, an antireflection coating could be designed with an **L** layer contiguous to the metal, but it usually contains more layers. The reason is evident when the antireflection coating is analyzed with a vector diagram.

Consider, for example, the three-layer antireflection coating captioned in Fig. 4-23. Three reflections that occur at the interfaces between dielectric layers cancel the strong reflection occurring at interface between the metal and $\mathbf{H}_1$ layer. The final question is, why is it preferable to have an **H** layer, rather than an **L** layer, is contiguous to metal? Eq. (2-16e) shows that the magnitude of the Fresnel amplitude reflection coefficient is less when a higher index **H** layer is used, rather than **L**.

4.5 Coatings with broader spectral bandwidth — maximally flat designs

4.5.1 Introduction

Coatings in §4.3 dereflect a substrate at a single wavelength with no effort to broaden the spectral region of low R. A maximally flat design provides both low R and broader bandwidth. The prototype maximally flat coating is composed of layers of equal optical thickness. Expressions are simplified by definition of R / T ratio F in Eq. (2-24b) and identity

$$F_s \equiv R_s / T_s \tag{4-7a}$$

for an uncoated substrate. The mathematical form of R / T vs phase thickness β of a maximally flat coating with ℓ layers is

$$F = F_s (\cos\beta)^{2\ell} . \tag{4-7b}$$

It has the feature that not only is R zero when $\beta = 90°$, but, in addition, first and higher derivatives with respect to $\cos^2\beta$ are also zero

$$\left[\frac{\partial F}{\partial(\cos^2\beta)}\right]_{\beta=90°} = \left[\frac{\partial^2 F}{\partial(\cos^2\beta)^2}\right]_{\beta=90°} = \left[\frac{\partial^3 F}{\partial(\cos^2\beta)^3}\right]_{\beta=90°} = 0 \tag{4-7c}$$

up to the ℓth derivative.

Table 4-3 lists approximate refractive indices of such coatings. For an antireflection coating that matches a germanium substrate (n_s = 4.0) to air, indices are attainable, with n_1 = 1.414 and n_2 = 2.828. For three layers, the index of the outer layer is quite low (1.19) and, therefore, is unattainable. This same lack of attainability is true when the layer count is larger than three.

So, if indices are unattainable in so many cases, why is there interest in maximally flat systems? One reason is that these antireflection coatings are useful in the design of bandpasses. It is not uncommon in such designs, for example, that a pseudo-substrate with an index as large as 33 must be matched to air over a relatively narrow bandwidth. Figure 4-27 depicts $R(\lambda_0 / \lambda)$ of antireflection coatings with one, two and three sections. Broader bandwidth of the three section system is evident. §7.4.3 shows that although a

few indices of such designs are attainable, others must be simulated with equivalent layers (see §2.6.7).

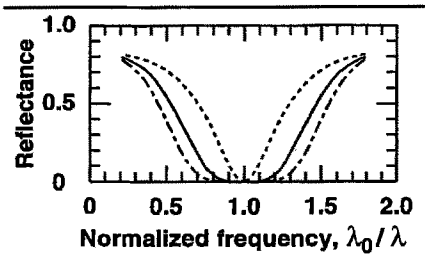


Fig. 4-27 $R(\lambda_0/\lambda)$ of maximally flat antireflection coatings (*short dashed*)
incident Q $_s$ *substrate*,
double layer antireflection coating(*solid*)
incident Q'Q" *substrate*,
and triple layer antireflection coating (*long-short dashed*)
incident $Q_1 Q_s Q_2$ *substrate*.
Refractive indices of *incident*, Q_1, Q', Q_s, Q", Q_2 and *substrate* are 1.52, 2.23, 3.28, 7.1, 15.3, 22.5 and 33, respectively. Optical thicknesses are $\lambda_0/4$.

4.5.2 Quasimaximally flat designs

4.5.2.1 Introduction

These are antireflection coatings in which several quarterwave or halfwave layers broaden the spectral bandwidth. Such antireflection coatings are not truly maximally flat but rather approximate maximally flat performance over a relatively narrow bandwidth.

A way of introducing this subject is to analyze the performance of

air Q D *substrate*

where Q and D are layers of optical thickness $\lambda_0/4$ and $\lambda_0/2$, respectively. Index n_1 of the Q layer is provided by Eq. (2-23j). This produces zero R at λ_0 because layer D is absentee (see §2.6.3.1) at that wavelength. A substrate of index 1.90 is chosen in this example. Such a high index is rarely encountered in the visible portion of the spectrum, except in exotic optical glasses or a copper chloride crystal used for second harmonic generation. Index of Q layer is 1.38.

Table 4-3 Refractive indices of maximally flat antireflection coatings

ℓ	n_1	n_2	n_3	n_4	n_5
2	$n_0^{3/4}\ n_s^{1/4}$	$n_0^{1/4}\ n_s^{3/4}$			
3	$n_0^{7/8}\ n_s^{1/8}$	$n_0^{1/2}\ n_s^{1/2}$	$n_0^{1/8}\ n_s^{7/8}$		
4	$n_0^{15/16}\ n_s^{1/16}$	$n_0^{11/16}\ n_s^{5/16}$	$n_0^{5/16}\ n_s^{11/16}$	$n_0^{1/16}\ n_s^{15/16}$	
5	$n_0^{31/32}\ n_s^{1/32}$	$n_0^{13/16}\ n_s^{3/16}$	$n_0^{1/2}\ n_s^{1/2}$	$n_0^{3/16}\ n_s^{13/16}$	$n_0^{1/32}\ n_s^{31/32}$
6	$n_0^{63/64}\ n_s^{1/64}$	$n_0^{57/64}\ n_s^{7/64}$	$n_0^{21/32}\ n_s^{11/32}$	$n_0^{11/32}\ n_s^{21/32}$	$n_0^{7/64}\ n_s^{57/64}$

Approximate refractive indices of maximally flat antireflection coatings. Each of the ℓ layers has identical optical thickness. Layer one is adjacent to the incident medium of index n_0 and the last layer is contiguous to substrate of index n_s. $n_6 = n_0^{1/64}\ n_s^{63/64}$.

An appendix (see §4.10.2) derives one of the equations in Table 4-3 and also demonstrates its extension beyond six layers.

Table 4-4 lists an equation for the index of D, which is 2.23 in this example. $R(\lambda)$ of substrate coated with only a Q layer appears in Fig. 4-28, thus illustrating that addition of halfwave D "layer" broadens the bandwidth. The quotation marks are used in foregoing

sentence because this layer is synthesized by two contiguous equivalent layers (see §2.6.7) of 90° phase thickness, as shown in caption to Fig. 4-28.

Table 4-4 Quasimaximally flat antireflection coatings

Design	Condition for $R = 0$ at λ_0	Condition for quasimaximally flat	Design example
Q D	$n_1 = n_0^{1/2}\ n_s^{1/2}$	$n_2 \approx n_0^{-1/4}\ n_s^{5/4}$	1.38 2.23 1.90
Q_1 Q D	$n_2 = n_0^{-1/2}\ n_1\ n_s^{1/2}$	$n_3 \approx n_0^{3/4}\ n_1^{-1}\ n_s^{5/4}$	1.35 1.86 1.65 1.90
Q_1 Q Q*	$n_2 = n_0^{-5/4}\ n_1^2\ n_s^{1/4}$	$n_3 \approx n_0^{-3/4}\ n_1^{-1}\ n_s^{3/4}$	*1.38* 2.11 1.89 1.52
Q_1 D Q	$n_3 = n_0^{-1/2}\ n_1\ n_s^{1/2}$	$n_2 \approx n_0^{-7/4}\ n_3^3\ n_s^{-1/4}$	*1.38* 2.37 1.70 1.52
Q_1 Q_2 Q Q¶	$n_3 = n_0^{1/2}\ n_1^{-1}\ n_2\ n_4\ n_s^{-1/2}$	$n_4 \approx n_0^{5/4}\ n_1^{-2}\ n_2\ n_s^{3/4}$	*1.38 2.30* 2.24 1.65 1.52
Q_1 D_1 Q Q‡	$n_4 = n_0^{1/2}\ n_1^{-1}\ n_3\ n_s^{1/2}$	$n_3 \approx n\backslash O\backslash AL(^{-9/4}{}_{,0})$ $n_1^4\ n_2^{-1}\ n_s^{1/4}$	*1.38 2.30* 1.75 1.56 1.52

Quasimaximally flat antireflection coatings. **Q** and **D** represent layers of optical thicknesses $\lambda_0/4$ and $\lambda_0/2$, respectively. Indices are ordered from incident medium to substrate and indices of layers with a subscript (i.e., Q_1 or D_1) are constrained, as shown by italics in design examples.
*For the Q_1 Q Q design, both indices n_2 and n_3 must satisfy equations for $R = 0$ at λ_0.
¶For the Q_1 Q_2 Q Q design, both indices n_3 and n_4 must satisfy their respective equations for R to be zero at λ_0. Calculate n_4 first.
‡For the Q_1 D_1 Q Q design, both indices n_2 and n_3 must satisfy their respective equations for R to be zero at λ_0. Calculate n_3 first.

4.5.2.2 Analysis using the vector addition of amplitudes

Why does this **Q D** design manifest a broader passband than a single layer? A vector diagram (see §2.8.3) furnishes the insight to answer this question. Assume that the index of the **D** layer is chosen properly. Figure 4-29 depicts vector diagrams for a single layer. At the tuned wavelength λ_0 (when phase thickness β is 90°), the vectors are collinear. At a wavelength 17% shorter than λ_0 (when β is 105°) vectors r_0 and r_1 (which are of equal length) are opening up like a pair of scissors — a sizable residual amplitude reflection coefficient $\hat{\rho}$ is evident.

Figure 4-30 shows vectors at the aforementioned wavelengths for the **Q D** design. At $\lambda_0 / \lambda = 1.17$ vector r_2 reverses direction, resulting in a $\hat{\rho}$ that is much closer to the origin than its counterpart in Fig. 4-29. This reversal is due to two factors: (a) Phase thickness of the **D** layer is double that of **Q** layer, since the former is a halfwave; and (b) sign of r_2 is positive, because there is a dense-to-rare reflection at the substrate interface.

4.5.2.3 Analysis using *R* / *T* polynomial

Indices in Table 4-4 are approximate values obtained from Kard theory (see §2.7.4.3). Derivation of indices for the **QD** design illustrates how they are obtained for a more general case. This **QD** coating is actually a degenerate three-layer design, in which n_2 =

n_3. Equation (2-51c) is the R/T polynomial for this coating. A maximally flat coating at $\beta = 90°$ is achieved when $a_2 = 0$. Use of approximation for a_2 in Table 2-15 results in

$$a_2 = 0 \approx 3u_0 - 4u_1 + 4u_2 \ . \quad (4\text{-}8a)$$

Equation (2-23j) translates from index into log index [using Eq. (2-29c)] as

$$u_1 = u_0/2 \ . \quad (4\text{-}8b)$$

Combining Eqs. (4-8a) and (4-8b) provides

$$u_2 \approx -u_0/4 \ . \quad (4\text{-}8c)$$

n_2 is found by translating the natural log index into index

$$n_2^4 \approx n_s^5 \, n_0^{-1} \ . \quad (4\text{-}8d)$$

One algebraic step transforms the foregoing into entry one in Table 4-4.

Justification for setting a_2 to zero to produce a quasimaximally flat design is evident by inspection of the R / T polynomial. Equation (2-51c) simplifies to

$$F = a_0^2 \cos^6\beta + a_1^2 \cos^4\beta \sin^2\beta \ . \quad (4\text{-}8e)$$

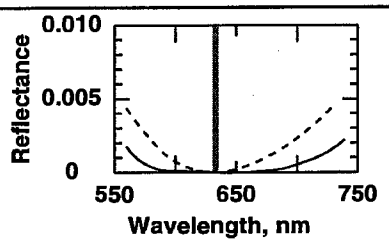


Fig. 4-28 $R(\lambda)$ of antireflection coatings at 633 nm (*dashed*)
air **L** *sub*
and (*solid*)
air 1.1**L** 0.769**H** 0.2**L** 0.769**H** 0.1**L** *sub*
where indices of **L**, *sub* and **H** are 1.38, 1.9 and 2.30, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$ at λ_0 of 633 nm.

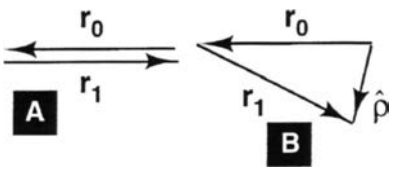


Fig. 4-29 Amplitude reflection coefficient $\hat{\rho}$ approximated graphically at (*A*) $\lambda_0/\lambda = 1.00$ and (*B*) $\lambda_0/\lambda = 1.17$ of the design *air* **L** *sub*, where **L** and *sub* have the same meaning as in Fig. 4-28.

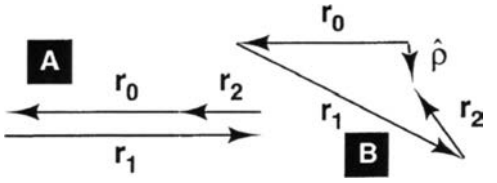


Fig. 4-30 Caption to Fig. 4-29 obtains, with the exception that coating is
air **L D** *sub*
where **D** represents layer of index 2.23 and optical thickness of $\lambda_0/2$.

This antireflection coating is not maximally flat owing to the presence of term containing a_1. It approximates a maximally flat antireflection coating, however, because that term is quite small when β is close to 90°. This is illustrated by the fact that when $\beta = 80°$, $(\cos 80°)^4$ is 0.001, which means that its contribution is still quite small. This would not be true if the a_2 term were not zero, in which case a $(\cos 80°)^2$ term would dominate and increase R substantially.

4.5.2.4 Constraint of the refractive indices

It is also possible to constrain any index. In the following design, the index of outer layer Q_1 is constrained

air Q_1 **Q Q** *substrate* .

Many pairs of refractive indices may be selected for the two remaining layers that produce zero R at λ_0. A unique solution is obtained by making the $R(\lambda_0/\lambda)$ curve

quasimaximally flat by setting both a_2 and a_3 to zero in Eq. (2-51c), similar to the procedure in §4.5.2.3. Solution of linear equations leads to the entry in Table 4-4.

Figure 4-31 depicts $R(\lambda)$ of three antireflection coatings from Table 4-4. Figure 4-32 shows $R(\lambda)$ when intermediate index layers are replaced with equivalent layers (see §2.6.7).

4.6 Coatings with zero reflectance at two or more wavelengths

4.6.1 Introduction

Antireflection coatings are designed in which $R = 0$ at two or more wavelengths. Goal is to find analytical solutions, rather than depending upon an optimization computer code. An appendix (see §4.10.3) discusses mathematical details and distinction between antireflection coatings that achieve zero R with *simultaneous solutions* or *manifold solutions* of R / T polynomial. Ensuing sections present at least one example of each type of solution.

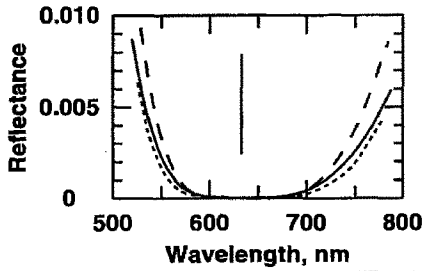


Fig. 4-31 $R(\lambda)$ of quasi-maximally flat antireflection coatings from Table 4-4. Each refractive index below is a QWOT at 633 nm (*long dash*)
air/1.38/2.11/1.89/*sub*,
(*solid*)
air/1.38/2.37/2.37/1.70/*sub*
and (*short dash*)
air/1.38/2.30/2.30/1.75/1.56/*sub* .
The *sub* refractive index is 1.52.

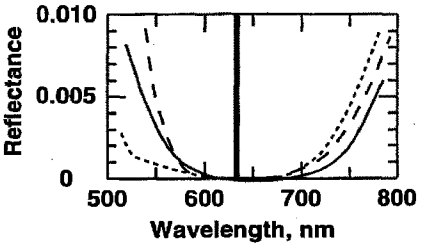


Fig. 4-32 $R(\lambda)$ of the antireflection coating (*long dash*)
air 1.176**L** 0.612**H** 0.452**L** 0.41**H** 0.276**L** *sub*,
(*solid*)
air **L** 2**H'** 0.356**L** 0.258**H** 0.356**L** *sub*
and (*short dash*)
air **L** 2.308**H** 0.347**L** 0.52**H**
0.536**L** 0.212**H** *sub*
where indices of **H'**, **H**, *sub* and **L** are 2.37, 2.30, 1.52 and 1.38, respectively. Optical thicknesses are $\lambda_0/4$ at λ_0 of 633 nm.

4.6.2 Quarter-half coating

Design of this coating is identical to that in §4.5.2.1.

Zeroes of the R / T polynomial are *simultaneous solutions*, as shown in an appendix — §4.10.4. There are two zeros in the $R(\beta)$ curve symmetrical about $\beta = 90°$. Whether a solution may be found depends upon finding attainable refractive indices. For a given β, for example, the **Q** layer index may be too low and/or the **D** layer index could be too high.

The following design procedure is used:

(1) Given wavelengths λ_1 and λ_2 (assuming $\lambda_2 > \lambda_1$) at which $R = 0$, compute ratio κ

$$\kappa \equiv \lambda_2 / \lambda_1 \ . \qquad (4\text{-}9a)$$

(2) Choose the index of **Q** layer as the lowest available — an index that satisfies Eq.-(2-23j), if possible. Then find u_1 — the natural log of the index — from Eq.-(2-29c). Next, solve for index n_2 of the **D** layer

$$u_2 \approx [u_1 \pm \sqrt{u_1^2 + 8(u_0 - u_1)^2}\ ]/4 \qquad (4\text{-}9b)$$

and thence find n_2

$$n_2 = n_s \exp(2u_2)\ . \qquad (4\text{-}9c)$$

(3) If the foregoing index is attainable, find ratio κ_1 of wavelengths of zeros from

$$\kappa_1 = (\pi - \mu)/\mu \qquad (4\text{-}9d)$$

where

$$\mu \approx \arctan[\sqrt{u_0/(3u_0 - 4u_1 + 4u_2)}\ ]\ . \qquad (4\text{-}9e)$$

Equation (2-29c) is used to translate refractive indices into u values.

(4) If κ obtained in step (1) is less than κ_1, then no solution exists with available indices.

(5) Substitute increasingly larger values for index n_1 into equations in step (2) and obtain values of n_2 that should also steadily increase. Eventually, n_2 should exceed the highest attainable. Compute wavelength ratio κ_2 for this condition.

(6) If the index n_1 is attainable, then compute index ratio κ_2 as is done in step (3). If $\kappa_1 \leq \kappa \leq \kappa_2$, then the solution is bracketed and a few iterations with trial values of n_1 in steps (2) and (3) will uncover a solution.

(7) Choose optical thickness by finding a center wavelength λ_0 equidistant in frequency between λ_1 and λ_2

$$\lambda_0 = [\ (\lambda_1^{-1} + \lambda_2^{-1}\ )/2]^{-1}\ . \qquad (4\text{-}10)$$

Optical thicknesses of **Q** and **D** are $\lambda_0/4$ and $\lambda_0/2$, respectively.

Example: A glass substrate (n_s = 1.70) is to be dereflected at two laser lines: λ_1 = 458 nm and λ_2 = 699 nm. Ratio of wavelengths is κ = 1.526:1. An index of 1.38 is the lowest for the outer layer. A film pair with indices 1.38 and 2.04 produces κ_1 = 1.36. The highest index available for the halfwave is 2.30. A film pair with indices 1.57 and 2.30 produces κ_2 = 1.59. There is a pair of attainable refractive indices that will dereflect the substrate adequately because $\kappa_1 < \kappa < \kappa_2$, i.e., 1.36 < 1.526 < 1.59.

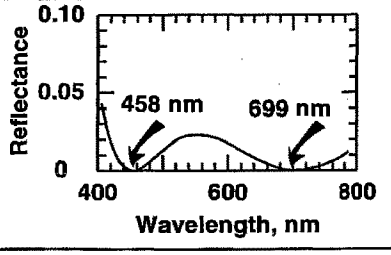


Fig. 4-33 $R(\lambda)$ of the antireflection coating

air **Q D** *glass* .

Indices of **D**, *glass* and **Q** are 2.18, 1.70 and 1.52, respectively. The optical thicknesses of **D** and **Q** are $\lambda_0/2$ and $\lambda_0/4$, respectively, at λ_0 of 553 nm.

After a few iterations, indices of 1.52 and 2.18 are found for outer and inner layers, respectively. $R(\lambda)$ of this antireflection coating appears in Fig. 4-33. At the center wavelength of 553 nm [obtained from Eq. (4-10)], the **D** layer is absentee (see §2.6.3.1) and R is the same as the substrate overcoated with only the **Q** layer.

4.6.3 Zeroes of reflectance — manifold solutions

Restrictions placed on refractive indices of layers (see §2.7.4.4 and § 4.10.5) make it relatively easy to derive equations for indices that produce coatings with zero R at specific wavelengths. Optical thickness of each **Q** layer in §4.6.3 is $\lambda_0/4$.

4.6.3.1 Q Q design: zero *R* at two wavelengths

Layer indices are constrained by

$$n_1 n_2 = n_0 n_s \ . \tag{4-11a}$$

Selection of n_1 determines the ratio of wavelengths at which $R = 0$ occurs.

The following procedure is used:

(1) Given wavelengths λ_1 and λ_2 (assuming $\lambda_2 > \lambda_1$) at which zero R is to be produced, compute ratio κ

$$\kappa = \lambda_2 / \lambda_1 \ . \tag{4-11b}$$

(2) Compute difference ξ between 90° and phase thickness β at which zero R occurs

$$\xi = 90°(\kappa - 1)/(1+ \kappa) \ . \tag{4-11c}$$

(3) Find u_1 (natural log of the index for outer layer)

$$u_1 \approx u_0 \ (3 - \tan^2\xi \ )/4 \tag{4-12a}$$

where

$$u_0 \equiv 0.5 \ \ln(n_0 / n_s) \ . \tag{4-12b}$$

(4) Calculate n_1 from

$$n_1 = n_s \exp(2 u_1) \ . \tag{4-12c}$$

(5) Obtain λ_0 from Eq. (4-10). Optical thicknesses are QWOTs at λ_0.

Example: Following the specifications of §4.2.2.1, a substrate of index 2.41 is dereflected at λ_1 = 2.65 μm and at λ_2 = 5.30 μm. Use of the procedure above shows that indices of 1.34 and 1.80 are needed. Both layers are QWOTs at λ_0 = 3.53 μm.

A few comments are appropriate regarding Eq. (4-12a). When ξ equals zero, $R = 0$ is at $\beta = 90°$ — the coating is maximally flat (see §4.5.1). When $\xi = \beta = 45°$, there is a 3:1 ratio between wavelengths where $R = 0$. In this instance, all layers of the antireflection coating have identical index, which is the square root of substrate index. In other words, the phase thickness at which the root occurs is bounded: $45° \le \beta \le 90°$.

Phase thicknesses — β values — at which $R = 0$ are visualized in vector diagrams. Figure 4-34 depicts such diagrams for germanium ($n_s = 4.0$) coated with layers of indices 1.80 and 2.22. $R(\lambda_0/\lambda)$ is also depicted. Vectors are co-linear at $\beta = 90°$ (at *C*). Vectors r_0 and r_2 are of equal length, due to Eq. (4-11a). At $\lambda_0/\lambda = 0.56$ (at *B*), vector polygon closes and $R = 0$. At frequency *A* vectors are unwinding and R is increasing rapidly.

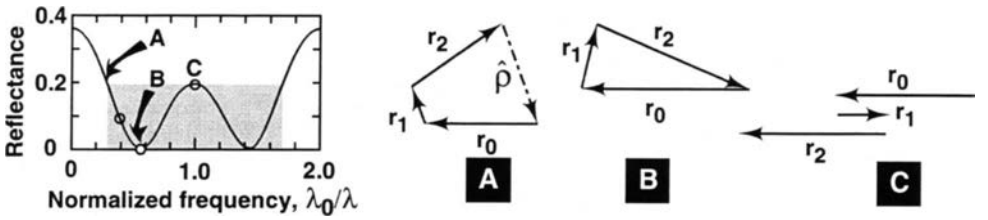


Fig. 4-34 $R(\lambda_0/\lambda)$ and vector addition of amplitudes computation of $\hat{\rho}$ for the antireflection coating

air Q_1 Q_2 *Ge* .

Indices of *Ge*, Q_2 and Q_1 are 4.00, 2.22 and 1.80, respectively.

The layers are of optical thickness $\lambda_0/4$.

4.6.3.2 Q Q Q design: zero *R* at three wavelengths

This antireflection coating contains three layers of equal optical thickness. It is deceptive to imply that three *independent* wavelengths exist at which R can be zero. Due to the limitations imposed by symmetry mentioned in §4.2.4, roots are symmetrical in wave number about $(\lambda_0)^{-1}$ λ_0 is determined by adjusting the layer thicknesses. Only outer layer index n_1 need be determined because the indices are constrained by

$$n_1 n_3 = n_0 n_s \qquad (4\text{-}12d)$$

and

$$n_2 = \sqrt{n_0 n_s} \quad . \qquad (4\text{-}13a)$$

The procedure for finding n_1 is identical to that in §4.6.3.1, with the exception that the following is used in lieu of Eq. (4-12a)

$$u_1 \approx u_0 \, (7 - \tan^2\xi \,)/8 \quad . \qquad (4\text{-}13b)$$

Comments pertaining to the Q Q Q design are apropos. When ξ equals zero, there exist three (degenerate) roots at $\beta = 90°$. This antireflection coating is maximally flat — its indices appear in Table 4-3. When $\xi = 60°$, then $\beta = 30°$ and all layers degenerate to the same index — that of the center layer. Thus the phase thickness at which the root occurs is bounded: $30° \le \beta \le 90°$.

Phase thicknesses — β values — at which $R = 0$ are depicted in vector diagrams. Figure 4-35 depicts such diagrams of coatings deposited on a germanium (n_s = 4.0) substrate. The outer layer index increases from 1.3 (panel *A*) to 1.414 (panel *B*) and finally to 1.6 (panel *C*). β at which $R = 0$ shifts from 55° to 45° and finally to 37.5°. Fig. 4-35 shows corresponding reflectance curves. An appendix (see §4.10.5) derives Eqs. (4-12a) and (4-13b).

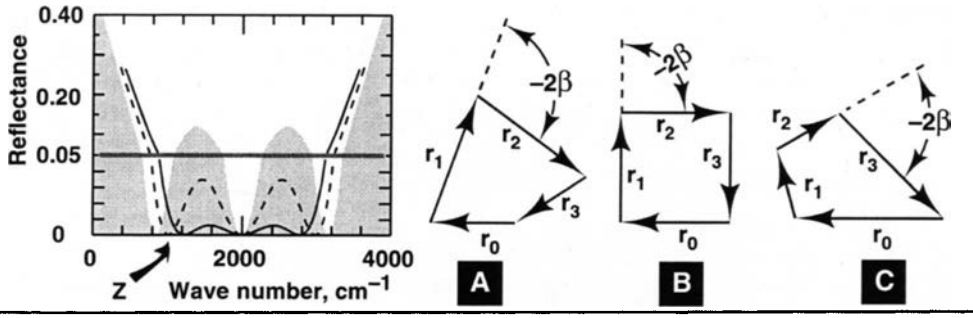


Fig. 4-35 At wave numbers (near the arrow *Z*) where $R = 0$, vector addition of amplitudes of
air Q_1 Q_2 Q_3 *Ge*
where indices of Q_2 and *Ge* are 2.00 and 4.00, respectively. The indices of Q_1, Q_3 are (*panel A*) 1.30, 3.08, (*panel B*) 1.414, 2.828 and (*panel C*) 1.60, 2.50. The optical thickness of each layer is 1.25 μm. (*Left*) $R(\sigma)$ of foregoing designs: *A* — *solid*, *B* — *dash* and *C* — *shaded*. The ordinate changes scale at 0.05.

4.6.3.3 Q Q Q Q design: zero *R* at four wavelengths

This antireflection coating contains four layers of equal optical thickness. For phase thickness β in the range $0 \le \beta \le 180°$, there are three possibilities: (1) There are four degenerate roots at $\beta = 90°$ — maximally flat condition obtains (see §4.5.1); (2) two pairs of degenerate roots are symmetrically situated about $\beta = 90°$; and (3) four roots occur in symmetrical pairs about $\beta = 90°$. Only two of the four refractive indices are variables, due to constraints

$$n_4 = n_0\, n_s / n_1 \qquad (4\text{-}14)$$

and

$$n_3 = n_0\, n_s / n_2 \ . \qquad (4\text{-}15)$$

An example illustrates the aforementioned three possibilities. The first step is to expand the R / T polynomial [from Eq. (2-51d)] in powers of sinβ

$$(u_0 + a_2 + a_4)\sin^4\beta - (a_2 + 2\,u_0)\sin^2\beta + u_0 = 0 \ . \qquad (4\text{-}16a)$$

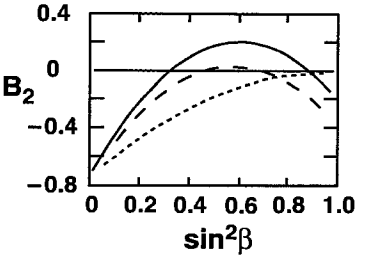


Fig. 4-36 $\sqrt{R/T}$ vs $\sin^2\beta$ of *air* $Q_1\ Q_2\ Q_3\ Q_4$ *sub* . The layers are of equal optical thickness and *sub* has index of 4.00. Indices (starting at outer layer) are: (*solid*) 1.37/1.75/2.28/2.93 , (*long dash*) 1.32/1.55/2.58/3.03 and (*short dash*) 1.09/1.54/2.59/3.66 .

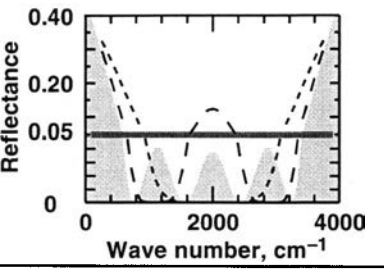


Fig. 4-37 $R(\sigma)$ of the four-layer coatings captioned in Fig. 4-36. Each layer has optical thickness of 1.25 μm. The ordinate changes scale at 0.05. For the shaded curve, the refractive indices of the layers are: 1.37/1.75/2.28/2.93 .

4.6.3.3 Q Q Q Q design: zero R at four wavelengths

This antireflection coating contains four layers of equal optical thickness. For phase thickness β in the range $0 \le \beta \le 180°$, there are three possibilities: (1) There are four degenerate roots at $\beta = 90°$ — maximally flat condition obtains (see §4.5.1); (2) two pairs of degenerate roots are symmetrically situated about $\beta = 90°$; and (3) four roots occur in symmetrical pairs about $\beta = 90°$. Only two of the four refractive indices are variables, due to constraints

$$n_4 = n_0\, n_s / n_1 \qquad (4\text{-}14)$$

and

$$n_3 = n_0\, n_s / n_2 \ . \qquad (4\text{-}15)$$

An example illustrates the aforementioned three possibilities. The first step is to expand the R / T polynomial [from Eq. (2-51d)] in powers of sinβ

$$(u_0 + a_2 + a_4)\sin^4\beta - (a_2 + 2\,u_0)\sin^2\beta + u_0 = 0 \ . \qquad (4\text{-}16a)$$

Table 2-15 provides approximate values for "a" coefficients

$$a_2 \approx 14\,u_0 - 12\,u_1 - 4\,u_2 \qquad (4\text{-}16b)$$

$$a_4 \approx u_0 - 4\,u_1 + 4\,u_2 \ . \qquad (4\text{-}16c)$$

Equation (4-16a) is a second degree polynomial in $\sin^2\beta$. Figure 4-36 plots this polynomial for several combinations of indices. The *short-dash* curve represents the maximally flat case. The polynomial intersects $R = 0$ only once — at $\beta = 90°$. The *solid* curve represents an antireflection coating with widely separated zeros. Zeros of the *long-dash* curve, on the other hand, are quite close together and approximate a degenerate root condition. In fact, roots are degenerate if $n_1 = 1.32$ and $n_2 = 1.50$. Figure 4-37 depicts

$R(\sigma)$ of antireflection coatings captioned in Fig. 4-36. Degenerate root condition occurs when

$$(a_2 + 2\,u_0)^2 - 4\,u_0\,(u_0 + a_2 + a_4) = 0 \ . \qquad (4\text{-}16d)$$

If the goal is to design a coating with degenerate roots, n_1 may be specified, whence u_1 is computed. Equation (4-16d) then results in a quadratic for u_2, which may be solved.

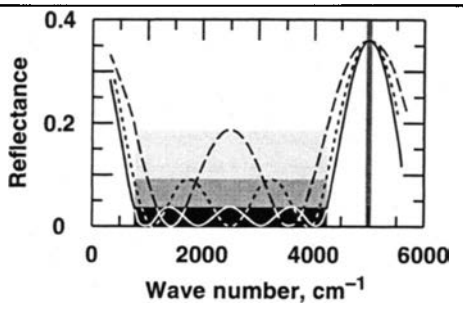


Fig. 4-38 $R(\sigma)$ of two (*long dash*), three (*short dash*) and four layers (*solid*) of optical thickness 1.00 μm deposited on substrate of index 4.00. Refractive indices are
air 1.783/2.244 ,
air 1.544/2.00/2.591 and
air 1.364/1.746/2.291/2.932 . The shaded area is the bandwidth — the region of equal ripple in reflectance.

4.7 Chebyshev antireflection coatings

4.7.1 Introduction: their attributes

These designs are quite similar to those of §4.6.3 — layers have identical optical thickness. A unique feature of a Chebyshev antireflection coating is that its R manifests equal ripple in its band of low reflectance — its passband. This bandwidth B appears in Fig. 4-38. Although certain Chebyshev antireflection coatings are nonproducible because required indices are unattainable, it is still necessary to review them because of their use in bandpass filters. A mathematical definition of a Chebyshev antireflection coating appears in an appendix, §4.10.6.

Given the substrate and incident medium indices, a Chebyshev antireflection coating with ℓ layers is characterized by: (1) its normalized bandwidth; and (2) maximum ripple of R in its passband. The foregoing parameters are not independent, but related, as shown below.

4.7.1.1 Spectral bandwidth

Normalized bandwidth w_q is related to the width of the spectral region of low reflectance. This is the wavelength interval (λ_1 to λ_2) or wave number interval (σ_1 to σ_2) in which R does not exceed a maximum R_{max} — its maximum ripple. w_q is defined as

$$w_q \equiv \frac{(\sigma_2 - \sigma_1)}{\text{average wave number}} = \frac{2\,(\sigma_2 - \sigma_1)}{\sigma_2 + \sigma_1} \qquad (4\text{-}17a)$$

or

$$w_q \equiv \frac{2\,(\lambda_2 - \lambda_1)}{\lambda_2 + \lambda_1} \ . \qquad (4\text{-}17b)$$

Since the antireflection coating is centered at $\beta_0 = 90°$, Eq. (4-17a) translates into phase thickness as

$$w_q \equiv 2\,(\beta_2 - \beta_1)/\pi \ . \qquad (4\text{-}17c)$$

Whether expressed in frequency or wavelength, w_q is confined to the interval

$$0 \le w_q \le 2 \ . \qquad (4\text{-}17d)$$

Example: The spectral region of equal ripple of the Chebyshev antireflection coating in Fig. 4-38 extends from 775 cm^{-1} to 4225 cm^{-1}. From Eq. (4-17a) w_q is

$$w_q = 2\,(4225 - 775)\,/\,(4225 + 775) = 1.38 \ . \qquad (4\text{-}17e)$$

The wavelength ratio $\lambda_2{:}\lambda_1$ is 5.5:1.

4.7.1.2 Ripple in the reflectance in the passband spectral region

The level of passband ripple depends upon the number of layers and passband width. For a given substrate index, ripple decreases as w_q remains constant and more layers are used. In Fig. 4-38, w_q of 1.38 is identical for all coatings. The ripple decreases from 19% R for the two-layer antireflection coating to 4% for the four-layer.

Table 4-5 Chebyshev polynomials

$\mathcal{T}_q(x)$	Polynomial
$\mathcal{T}_0$	1
$\mathcal{T}_1$	x
$\mathcal{T}_2$	$2x^2 - 1$
$\mathcal{T}_3$	$4x^3 - 3x$

$\mathcal{T}_q(x)$	Polynomial
$\mathcal{T}_4$	$8x^4 - 8x^2 + 1$
$\mathcal{T}_5$	$16x^5 - 20x^3 + 5x$
$\mathcal{T}_6$	$32x^6 - 48x^4 + 18x^2 - 1$
$\mathcal{T}_7$	$64x^7 - 112x^5 + 56x^3 - 7x$

Chebyshev polynomials of the *first* kind of order q, not to be confused with Chebyshev polynomials of the *second* kind in Table 5-15.

Tables for design of microwave filters may be used to find ripple [Matthaei 64]. It follows from Eq. (4-37a) in §4.10.6 that passband ripple (in terms of $F \equiv R / T$) for a coating with ℓ layers is

$$F = \frac{F_s}{[\mathcal{T}_\ell (1/C_0)]^2} \tag{4-18a}$$

where $\mathcal{T}_\ell$ is a Chebyshev polynomial of the first kind of order ℓ (see Table 4-5). F_s appears in Eq. (4-7a) and C_0 is related to fractional bandwidth by

$$C_0 \equiv \sin(w_q \, \pi / 4) . \tag{4-18b}$$

A more intuitive definition of C_0 is

$$C_0 \equiv \cos\beta_0 \tag{4-18c}$$

where β_0 is the phase thickness at the edge of the passband.

Example: The Chebyshev antireflection coating captioned in Fig. 4-38 (coated on a germanium substrate) is centered at 2500 cm^{-1} and its passband edge is at 775 cm^{-1}. β_0 is found from the proportion: $90°:\beta_0 :: 2500:775$. β_0 is 27.9° and [from Eq. (4-18a)] $\cos\beta_0$ is 0.884. The reciprocal of 0.884 is 1.132. F_s is 0.5625. Using Table 4-5, $\mathcal{T}_2(1.132) = 2(1.132)^2 - 1 = 1.5628$. It follows from Eq. (4-18a) that $F = 0.5625/(1.5628)^2 = 0.23$. Use of Eq. (2-24c) shows that its maximum R in the passband is 18.7%.

4.7.2 Procedure for design

These steps are followed:

(1) Select wavelengths λ_1 and λ_2 at edges of band of low reflectance.

(2) Compute w_q from Eq. (4-17b), C_0 using Eq. (4-18b) and u_0 from Eq. (4-12b).

(3) Assume that a two-layer antireflection coating is used.

(4) Find the passband ripple using Eqs. (4-18a), (4-18b) and (2-24c).

(5) If the ripple is sufficiently small, then proceed to step (6). If not, increment number of layers by one and repeat step (4).

(6) Find the natural log of the index of the ith layer from

$$u_i \approx \frac{u_0}{\kappa_1}\left[\kappa_2 - \frac{C_0^2\, P(C_0)}{Q(C_0)}\right] . \qquad (4\text{-}19a)$$

Table 4-6 defines coefficients κ_1 and κ_2, as well as polynomials $P(C_0)$ and $Q(C_0)$.

Table 4-6 Coefficients for Chebyshev antireflection coatings

ℓ	i	κ_1	κ_2	$P(C_0)$	$Q(C_0)$
2	1	4	3	1	$2 - C_0^2$
3	1	8	7	3	$4 - 3\,C_0^2$
4	1	16	15	$8 - C_0^2$	$8 - 8\,C_0^2 + C_0^4$
4	2	16	11	$8 - 5\,C_0^2$	Same
5	1	32	31	$20 - 5\,C_0^2$	$16 - 20\,C_0^2 + 5\,C_0^4$
5	2	16	13	$20 - 15\,C_0^2$	Same
6	1	64	63	$48 - 18\,C_0^2 + C_0^4$	$32 - 48\,C_0^2 + 18\,C_0^4 - C_0^6$
6	2	64	57	$144 - 126\,C_0^2 + 7C_0^4$	Same
6	3	32	21	$48 - 54\,C_0^2 + 11\,C_0^4$	Same

Coefficients in Eq. (4-19a) for computing the natural log of the index of the ith layer in a Chebyshev antireflection coating containing ℓ layers. C_0 appears in Eq. (4-18c).

(7) Given u_i, find the index n_i of the ith layer using Eq. (2-53b).
(8) Find the remaining indices using the conditions of symmetry outlined in §2.7.4.4.

Example: For a three-layer Chebyshev antireflection coating deposited on silicon (n_s = 3.60), find the index of the outer layer when w_q is 1.30. Using Eq. (4-18b), C_0 is evaluated as 0.853 and u_0 is – 0.6405. From Eq. (4-19a) we have

$$u_1 \approx \frac{u_0}{8}\left[7 - \frac{3\,C_0^2}{4 - 3\,C_0^2}\right] = \frac{-0.6405}{8}\left[7 - \frac{2.183}{1.817}\right] = -\,0.464 . \qquad (4\text{-}19b)$$

Equation (4-12c) provides an index of 1.42.

4.7.3 Comparison with maximally flat antireflection coating

A maximally flat antireflection coating may be viewed as a Chebyshev antireflection coating in which roots are degenerate at $\beta = 90°$, which is equivalent to letting $w_q \to 0$ or alternatively $C_0 \to 0$. As $C_0 \to 0$, for example, Eq. (4-19a) simplifies considerably and u_1 for a two-layer antireflection coating reduces to

$$u_1 \approx 3u_0/4 . \qquad (4\text{-}20a)$$

Use of Eq. (4-12b) and conversion into index produces the entry in Table 4-3.

4.8 Step-up and step-down of admittance

4.8.1 Introduction

Previous sections describe antireflection coatings in which refractive index steps down from 4.0 (for a germanium substrate) to air, as shown in Fig. 4-39. A distressing fact is

that a few of the indices are unattainable if w_q shrinks much below 1.50 for a germanium substrate. They are unattainable for all w_q values when the substrate is BK-7. This is avoided by an alternative — the step-up, step-down design is depicted in Fig. 4-39.

The concept of a coating as an admittance transformer (see §4.2.3) is used to analyze its performance. Its design is

air $\mathbf{Q}_1$ **D** $\mathbf{Q}_3$ *substrate* .

The relatively low normalized admittance of substrate is stepped upwards by a quarterwave layer $\mathbf{Q}_3$. Center layer **D** is of higher index; its optical thickness is a halfwave. This **D** layer does not alter the admittance produced by $\mathbf{Q}_3$. Outer layer $\mathbf{Q}_1$ of index 1.38 steps admittance down to air.

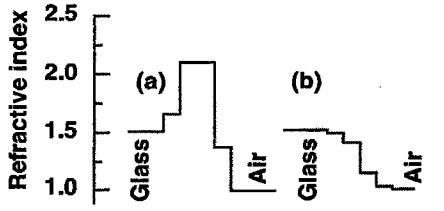


Fig. 4-39 For a BK-7 glass substrate, refractive index vs optical thickness of: (*a*) step-up-step-down antireflection coating (see Fig. 4-41); and (*b*) step-down antireflection coating (indices from Table 4-3).

4.8.2 Quasimaximally flat quarter-half-quarter

This design appears in Table 4-4. At λ_0, the optical thickness of the center layer **D** is a halfwave. It is, therefore, absentee (see §2.6.3.1) at λ_0. The condition for zero R to occur at λ_0 is given in Eq. (4-2a), provided n_3 appears on the left of that equation.

For $n_1 = 1.38$ and a BK-7 substrate, the index of 1.70 for n_3 in Table 4-4 is difficult to achieve. The center layer index of 2.37 is equally difficult; hence it is better to compromise with a lower index, such as 2.30.

4.8.3 Quarter-half-quarter: wavelengths of zero reflectance

It is possible to design an antireflection coating with two zeros in its reflectance — symmetrically located on a wave number scale about (λ_0).[1] An appendix (see §4.10.7) derives the equations used below. The procedure is as follows:

(1) Given the wavelengths λ_1 and λ_2 (assuming $\lambda_2 > \lambda_1$) at which R is to be zero, compute phase thickness β_1 at which the root occurs

$$\beta_1 = 180°(1+ \lambda_2/\lambda_1)^{-1} . \qquad (4\text{-}20b)$$

(2) Choose an index for the outer layer. Compute its log index u_1 from

$$u_1 = 0.5 \ln(n_1 / n_s) . \qquad (4\text{-}20c)$$

(3) Compute u_0 using Eq. (4-12b).

(4) Find u_3 from

$$u_3 \approx \frac{u_1 (2 S^4 - 6 C^2 S^2) - u_0 (S^4 - 6 C^2 S^2 + C^4)}{2 S^4 - 6 C^2 S^2} \qquad (4\text{-}20d)$$

where

$$S \equiv \sin\beta_1 \quad \text{and} \quad C \equiv \cos\beta_1 . \qquad (4\text{-}21a)$$

(5) Find n_3 from

$$n_3 = n_s \exp(2 u_3) . \qquad (4\text{-}21b)$$

(6) Determine u_2 from

$$u_2 \approx [C^2 (4 u_0 - 2 u_1 - 4 u_3) - S^2 (4 u_0 - 6 u_1 - 6 u_3)] /4 . \qquad (4\text{-}21c)$$

(7) Find n_2 from Eq. (2-53b).

(8) Optical thicknesses are multiples of $\lambda_0/4$, where λ_0 appears in step (7) of §4.6.2.

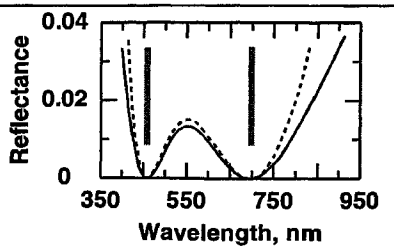


Fig. 4-40 $R(\lambda)$ of the antireflection coating
air Q_1 **D** Q_3 *glass* .
The optical thicknesses of Q_1, **D** and Q_3 are $\lambda_0/4$, $\lambda_0/2$ and $\lambda_0/4$, respectively, at λ_0 of 553 nm. Refractive indices of the layers and substrate are (*solid*)
air 1.35/1.74/1.52/1.60
and (*dashed*)
air 1.52/2.35/1.70/1.60.

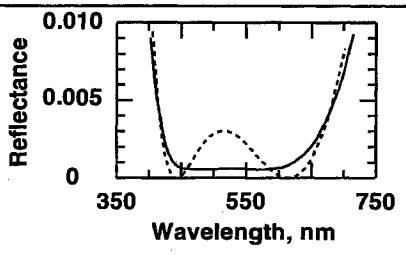


Fig. 4-41 $R(\lambda)$ of the antireflection coating
air Q_1 **D** Q_3 *glass* .
where optical thickness of Q_1, **D** and Q_3 are $\lambda_0/4$, $\lambda_0/2$ and $\lambda_0/4$, respectively, at λ_0 of 515 nm. The refractive indices of layers and substrate are (*solid*)
air 1.38/2.08/1.66/1.52
and (*dashed*)
air 1.38/2.08/1.61/1.52.

Example The antireflection coating in the example in §4.6.2 is emulated with a Q_1 **D** Q_3 design deposited on a substrate of index 1.60. Two solutions are found — each has a different index for the outer layer. Indices of layers in these designs are *air*/1.35/1.74/1.52/*sub* and *air*/1.52/2.35/1.70/*sub*. Although both designs manifest $R = 0$ at λ_1 and λ_2, R at λ_0 is different, as depicted in Fig. 4-40.

4.8.4 Simulation of the bottom layer with two materials

Consider an antireflection coating for the visible part of spectrum. Figure 4-41 illustrates the effect of varying the index n_3 of the quarterwave layer Q_3 adjacent to substrate. An objectionable reflectance peak in the green occurs when n_3 drops below 1.65. Rock [69] observed that there is a paucity of suitable materials with the correct index. He suggested that this layer be replaced with two layers of nonquarterwave optical thickness — an **L H** pair. Rock [69] explains how a circle diagram (see §10.1.2.3) produced his design. The procedure of §4.3 and concept of admittance are used here.

4.8.4.1 Design method

The objective is to find the optical thicknesses of layers **L** and **H**. Two questions are posed: (1) At λ_0, what is the normalized admittance of the bottom Q_3 layer (of optical thickness $\lambda_0/4$ and refractive index n_3) deposited on a bare substrate? (2) How is that admittance simulated with an **L H** pair?

Consider the Q_3 layer deposited on the substrate. Its normalized admittance Y is obtained from Eq. (2-48c)

$$Y = n_3^2 \, n_s^{-1} \ . \qquad (4\text{-}22a)$$

For example, with an index of 1.65 for layer Q_3 on BK-7 ($n_s = 1.52$),

$$Y = (1.65)^2 \, (1.52)^{-1} = 1.79 \ . \qquad (4\text{-}22b)$$

Thicknesses of films in the **L H** pair may be found by designing an antireflection coating that matches a BK-7 substrate to a fictitious medium of index 1.79. Following Rock [69], indices of 1.38 (magnesium fluoride) and 2.08 (zirconia) are used for **L** and **H** layers, respectively. Application of the procedure in §4.3.4.2.2 shows that optical thickness of

the H layer (next to the substrate) should be 0.054 waves and that the outer L layer should be slightly thicker — 0.070 waves. Fig. 4-42 shows measured $R(\lambda)$ of an antireflection coating produced by the assignee of the Rock [69] patent. Calculated $R(\lambda)$ is also depicted.

4.8.5 Design with an equivalent center layer

Ikeda [73] replaced the center D layer of the design in §4.8.3 with an equivalent layer (see §2.6.7) with a phase thickness of 540° at its center wavelength. This fictitious layer, with its optical thickness of $3\lambda_0/2$, provides an equivalent index in the range 1.85 to 1.95, as depicted in Fig. 4-43. The design of this equivalent layer is 2H' 2M' 2H' (where H' and L' are $\lambda_0/4$ in optical thickness).

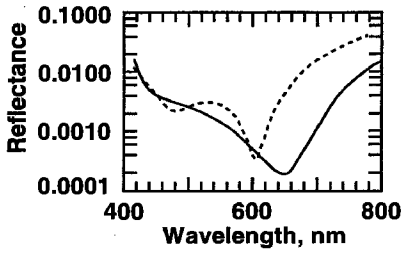


Fig. 4-42 $R(\lambda)$ of (*dashed*) *air* L 2H 0.28L 0.216H *glass* where optical thicknesses are $\lambda_0/4$ at λ_0 of 515 nm. Refractive indices of L, *glass* and H are 1.38, 1.52 and 2.08, respectively. (*Solid*) Measured $R(\lambda)$ of an antireflection coating of unknown design.

The patent of Ikeda [73] includes a design

air $Q_1\ D_2\ D_3\ D_4\ Q_5$ *substrate*

where Q layers and D layers have optical thicknesses of $\lambda_0/4$ and $\lambda_0/2$, respectively. This patent suggests the use of indices 1.38, 2.10, 1.63, 2.10 and 1.63 for the $Q_1\ D_2\ D_3\ D_4\ Q_5$ layers, respectively. The $R(\lambda)$ of this antireflection coating is depicted in Fig. 4-44. Reflectance peaks of nearly 1% in the red and blue are annoying. A coating with a nearly 2:1 bandwidth can be attained if the aforementioned indices are altered.

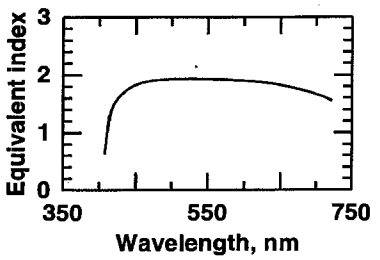


Fig. 4-43 Equivalent layer index of HH MM HH where optical thickness are $\lambda_0/4$ at λ_0 of 533 nm. The refractive indices of M and H are 1.63 and 2.10, respectively.

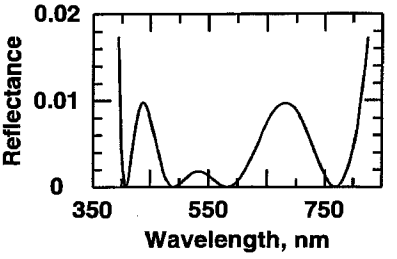


Fig. 4-44 $R(\lambda)$ of *air* L HH MM HH M *glass* where the layer thicknesses and refractive indices are captioned in Fig. 4-43. The refractive indices of L and *glass* are 1.38 and 1.52, respectively.

4.8.5.1 Use of refractive index optimization

The goal is to minimize $R(\lambda)$ over a 2:1 bandwidth. The design captioned in Fig. 4-44 was modified slightly by splitting each of the three halfwave layers into two

quarterwaves. Optimization then adjusted only refractive indices, with a restriction that no index could exceed 2.30 nor be less than 1.38.

The resulting design has eight layers of equal optical thickness, with a sequence of refractive indices *air*/1.38/2.18/2.28/1.64/1.54/2.03/2.20/1.78/*glass*. Figure 4-45 shows its $R(\lambda)$, which is less than 0.3% over a 2:1 bandwidth.

The refractive indices in the last paragraph are both conventional and esoteric. In the visible portion of the spectrum, indices of 1.38, 1.63 and 2.03 are close to those of magnesium fluoride, alumina and zirconia, respectively. A nonabsorbing film of index 1.78 could be produced in the visible part of the spectrum by mixing evaporants (see Tables 3-7, 3-8 and 3-9), which is difficult. Even if this approach were followed, this design would have to be adjusted for dispersion of coating materials, especially those whose indices exceed 2.00. Another problem is that few evaporation tanks contain evaporation sources for eight materials. Notwithstanding these problems, both Ganner [83] and Louderback [72] describe antireflection coatings in which a mixture of coating materials was used to synthesize a layer with a specified refractive index.

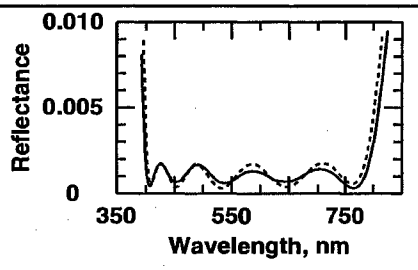


Fig. 4-45 $R(\lambda)$ of layers of equal optical thickness with indices (*dashed*) *air* 1.38/2.18/2.28/1.64/1.54/2.03/2.20/1.78/*glass* and (*solid*) *air* 0.56L 0.315H 0.074M 0.637H 0.613M 0.485L 0.115H 0.41M 0.25H 0.27M 0.12H 0.5M *glass* where optical thicknesses are λ_0 /4. Refractive indices of L, *glass*, M and H are 1.38, 1.52, 1.63 and 2.30, respectively. λ_0 is 533 nm for *dashed* and 1.0 μm for *solid*.

4.8.5.2 Equivalent layers and the computer optimization of layer thicknesses

Another approach is to select several indices and to synthesize the remaining indices with equivalent layers (see §2.6.7). Layer thicknesses are then adjusted to minimize R.

Three materials are selected: magnesium fluoride (n = 1.38), alumina (n = 1.63) and titania (n = 2.30). The foregoing indices are nominal, of course, since they change with wavelength. With reference to the design cited in §4.8.5.1, layers with indices 1.64 and 2.28 are close enough to indices of alumina and titania, respectively. The film of index 1.54 is replaced with an equivalent layer composed of magnesium fluoride and alumina. An equivalent layer consisting of alumina and titania is substituted for the index of 2.03.

After layer thicknesses are adjusted with an optimization computer code, the resulting $R(\lambda)$ is shown in Fig. 4-45 (*solid*). The design is not finished, however. Effects of dispersion should be considered and the exceedingly thin layers eliminated. That is, layer three should be discarded and the optimization process repeated.

4.9 Miscellaneous topics

This section covers topics that do not fit into the fabric of previous sections.

4.9.1 Simultaneous reduction of the reflectance in the visible and IR

The requirement is to dereflect a substrate at two widely separated wavelengths — where one is at least ten times larger than the other. An example is an antireflection coating that

dereflects in the visible and middle IR simultaneously. The following design procedure is used:

(1) Design an antireflection coating for the IR, using as few layers as possible. Materials in this coating must be transparent at both visible and IR wavelengths.

(2) Given the layers of the antireflection coating in step (1), insert an antireflection coating for the visible at all interfaces in the stack. Figure 4-46 shows an antireflection coating that uses two layers to dereflect in the IR. There are three interfaces, with a visible antireflection coating at each interface. The antireflection coating V_1 matches the upper layer IR1 to air. V_2 reduces R to zero (in the visible!) at the interface between the two IR layers. Proper design of coating V_3 ensures that there is no reflection between the substrate and layer IR2 contiguous to it.

(3) Adjust thicknesses of layers of the IR coating to account for perturbations caused by insertion of visible antireflection coatings.

Example: An optical parametric oscillator functions at wavelengths near 7.9 μm. An antireflection coating is coated onto a zinc selenide substrate that transmits at both 7.9 μm and 633 nm. Layers of zinc sulfide **H'** and a rare earth fluoride **L'** are used. Both materials are transparent at 633 nm, as well as in the IR. It is important to account for dispersion of both the substrate and films. IR indices of the rare earth fluoride, zinc sulfide and substrate are 1.40, 2.22 and 2.42, respectively. Corresponding indices in the visible are 1.50, 2.30 and 2.56, respectively.

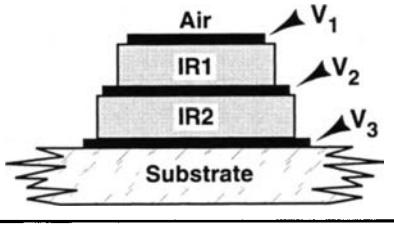


Fig. 4-46 Showing design of an antireflection coating with layers IR1 and IR2 that dereflect in IR and multilayers V_1, V_2 and V_3 that dereflect in the visible portion of the spectrum.

The procedure of §4.3.4.2.2 is used to concoct a two-layer antireflection coating for $\lambda_0 = 7.9$ μm

air **H' L'** *substrate*

where optical thicknesses of **H'** and **L'** are 0.219 λ_0 and 0.566 λ_0, respectively. These are layers IR1 and IR2 in Fig. 4-46.

For the visible antireflection coatings, an equivalent layer (see §2.6.7) is inserted at the surface of each IR layer. V_1 matches, for example, the surface of IR1 ($n = 2.30$ at 633 nm) to air. From Eq. (2-23j) its index should be $\sqrt{2.30} = 1.50$. This is close enough to the index of a rare earth fluoride, such as yttrium fluoride, so that a single layer of that material is all that is needed.

The V_2 layer matches a bulk material with index of IR1 to a bulk material with index of IR2. Its index is

$$n = \sqrt{2.30 \times 1.50} = 1.86 \qquad (4\text{-}23)$$

and its phase thickness is 90° at 633 nm. An equivalent layer is used to simulate it.

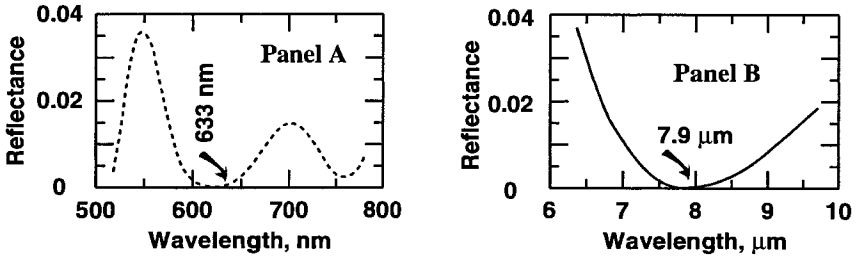


Fig. 4-47 $R(\lambda)$ of (*dashed, left panel*) of
air 0.1033**L** 0.1453**H** 0.0342**L** 0.0225**H** 0.738**L** 0.0297**H** 0.0287**L** *sub*
where thicknesses are metric in μm. The refractive indices of **L**, *sub* and **H** are 1.50, 2.56 and 2.30, respectively, for (*panel A — dashed*).
The corresponding indices are 1.40, 2.42 and 2.22 for (*panel B — solid*).

Finally, an equivalent layer V_3 matches index of IR2 to the substrate. Its optical thickness is a QWOT at 633 nm and its index should be

$$n = \sqrt{2.56 \times 1.50} = 1.66 \ . \qquad (4\text{-}24)$$

Insertion of these layers slightly alters the IR performance of the stack. Although there are no plots to illustrate it, the effect is to shift the minimum of $R(\lambda)$ (near 8 μm) to longer wavelengths by about 10%. Thicknesses of both IR layers are then re-adjusted to bring null reflectance back to the prescribed IR wavelength. Figure 4-47 depicts the $R(\lambda)$ of the final design in the visible and IR. A virtue of this design is that its band of low R in the visible is relatively insensitive to changes in thicknesses of the two thicker IR layers.

Berning [62] showed how equivalent layers (see §2.6.7) are used to simulate such a graded index profile. An advantage of this coating is that $R(\lambda)$ is quite insensitive to the thickness of any layer. A disadvantage is that $R(\lambda)$ is not particularly low, for reasons explained in §4.9.2.2.

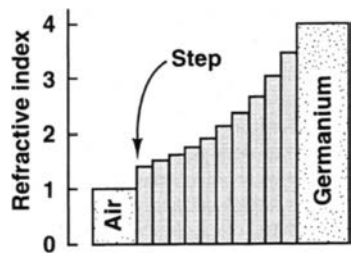


Fig. 4-48 Refractive index of ten layers that simulate a hyperbolic profile that grades from index of 1.39 to 4.00.

4.9.2 Simulated graded refractive index using equivalent layers

4.9.2.1 Introduction

It has been known since the early 1900s that a graded index layer dereflects a dielectric surface over a broad bandwidth, provided that refractive indices close enough to 1.00 are available. Figure 4-48 shows a graded profile coating (simulated by discrete layers which might be regarded as steps) that grades from a relatively high index germanium substrate to air. Such a graded index coating would function superbly on BK-7 if a coating material with an index as low as 1.06 were available. Since nature has not provided such films, graded index coatings are usually confined to higher index substrates.

4.9.2.2 Design example

A germanium substrate is dereflected in the spectral region from its cuton (near 2 μm) to 10 μm. Coating materials are germanium (n = 4.2), zinc sulfide (n = 2.2) and a rare earth

fluoride (n = 1.35). Although the dispersion of the rare earth fluoride should be taken into account, it is assumed to be nondispersive in this example.

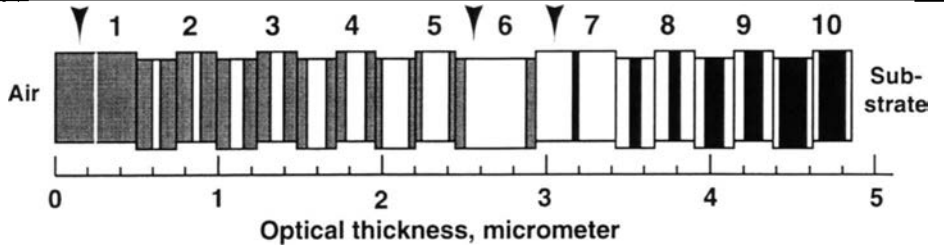


Fig. 4-49 Optical thickness of an antireflection coating that simulates a graded profile. It consists of layers of germanium (*solid black*), zinc sulfide (*white*) and yttrium fluoride (*shaded*).

The next questions are: (1) What refractive index profile should be used? (2) What is the total thickness of the antireflection coating? (3) How many steps should be used in its simulation with equivalent layers?

It really does not make much difference what index profile is used, provided that total optical thickness of the coating is at least a halfwave at the longest wavelength at which it is designed to function. The greater the number of steps, the lower the ripple in its $R(\lambda)$ curve. As an example, consider the sequence of indices

air/1.39/1.50/1.62/1.75/1.92/2.12/2.37/2.67/3.03/3.46/*Ge* .

With the exception of *air*, these indices follow a hyperbolic profile (see Jacobsson [75]) and are graphed in Fig. 4-48.

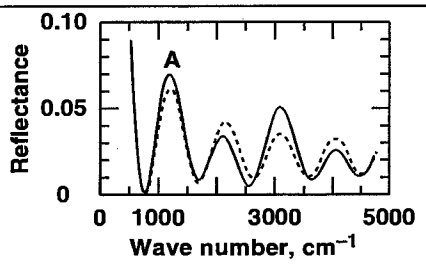


Fig. 4-50 $R(\sigma)$ of the antireflection coating with the simulated refractive index profile (*solid*) and $R(\lambda)$ for the ten-layer design (*dashed*). The coating design appears in §4.10.8.

The next step is to use equivalent layers to replace those layers in the last paragraph. Arrowed layers in Fig. 4-49 simulate a QWOT at 2.00 μm. All other layers simulate a QWOT at 1.00 μm. As depicted in Fig. 4-49, layers of yttrium fluoride (n = 1.35) and zinc sulfide (n = 2.20) simulate those layers whose indices are less than 2.12. Equivalent layers consisting of zinc sulfide and germanium simulate those layers whose equivalent indices exceed 2.20.

Figure 4-50 shows $R(\lambda)$ of a design with a stepped index profile (*dashed*) and its simulation with equivalent layers (*solid*). Since unprotected rare earth fluoride is the outer layer, there is no guarantee that this coating would pass the 24 h humidity environmental test. Sankur [84] emulated a design of Berning [62] by depositing upwards of 50 calcium fluoride and zinc sulfide films on a zinc selenide substrate.

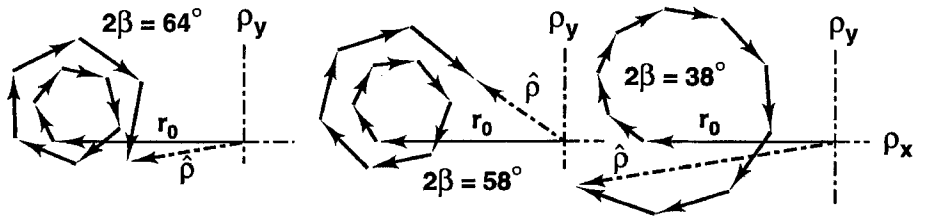


Fig. 4-51 Vector addition of amplitudes calculation of ρ of antireflection coating with ten layers captioned in Fig. 4-50.

4.9.2.3 Analysis using vector addition of amplitudes

Analysis by vector addition of amplitudes (see §2.8.3) offers insight into the graded profile antireflection coating and how it functions [Kossel 48, Jacobsson 66b]. Figure 4-51 shows such a diagram for the ten-layer antireflection coating captioned in Fig. 4-50. There is an index discontinuity between air and the outer layer of index 1.39 — identified as "step" in Fig. 4-48. The Fresnel amplitude reflection coefficient [from Eq. (2-8f)] at this interface is

$$r_0 = (1 - 1.39)/(1 + 1.39) = -0.161 \qquad (4\text{-}25)$$

and R is $(-0.163)^2 = 2.3\%$.

Figure 4-51 (*right panel*) shows a vector diagram when the double phase retardation 2β for each layer is 38°. Vectors form a spiral — the magnitude $|\hat{\rho}|$ is nearly a maximum. This corresponds to maximum R at point A in Fig. 4-50. Left and middle panels in Fig. 4-51 illustrate that the spiral tightens with increasing wave number. When 2β is 58°, for example, $|\hat{\rho}|$ is nearly a minimum. $|\hat{\rho}|$ is close to a maximum again when 2β has increased to 64°, although the maximum is now less than at $2\beta = 38°$ because the spiral has tightened.

The foregoing vector sums explain why $R(\lambda)$ in Fig. 4-50 eventually damps out to about 2% at wave numbers greater than 3000 cm^{-1}. This 2% is approximately r_0^2, where r_0 is obtained from Eq. (4-25). Although many gradient index profiles could be used, the main effect of using a different profile is to alter R at maxima and minima, as well as the wave numbers at which they occur. The 2% mean residual $R(\lambda)$ at high frequencies is essentially independent of the functional form of the index profile. This residual $R(\lambda)$ could be further reduced only by using a lower index for the outer layer.

4.9.3 Realization of coatings

The "moment of truth" eventually arrives when coatings designed in this chapter must be produced. A few designs could be taken from this book and deposited. In many other instances, however, there are problems to contend with such as adhesor layers, condition of substrate, dispersion, index control and regulation of thicknesses.

Adhesor layers are mentioned in §3.5.7 that "glue" a layer to either a substrate or to another layer. As an illustration, the design captioned in Fig. 4-5 will endure environmental tests — humidity and abrasion tests — if an adhesor layer of Y_2O_3 [Lubezsky 83] or MgO [Hacker 90] were deposited on the germanium substrate. The coating design would be re-optimized after addition of this layer.

The condition of the substrate affects the optical properties of an antireflection coating. §9.2.4 mentions that Günther [81] found that an acidic cleaning bath leached oxide constituents from the surface of ophthalmic glass — thus forming a surface layer. Designs of antireflection coatings had to be modified to account for optical properties of that layer, which was modeled to have a refractive index of 0.09 less than bulk glass and a thickness as large as 200 nm.

Dispersion is considered in only one coating — in Fig. 4-47. In other coatings that extend over a broad bandwidth, dispersion of both substrate and coating materials should be accounted for in the computation of $R(\lambda)$.

Two methods of controlling the thickness of an ultra-thin layer are: (1) quartz crystal (see §9.6.1.2); and (2) sector masking. Such masking greatly restricts the area of a rotating substrate being coated. This mask is similar to that shown in Fig. 9-44, but it obscures a much larger area of the substrates. This allows a layer thicker than quarterwave to be deposited on an optical monitor chip (see §9.6.4.2), while yet depositing a much thinner layer on the substrates. As an illustration, if the *control ratio* $\mathcal{M}$ [see Eq. (9-17e)] is 3, then deposition of a quarterwave layer on the monitor chip results in a film of "one twelfth-wave optical thickness" on the substrates.

4.9.4 Design using linear programming

Linear programming LP can be used to minimize the R / T polynomial for a stack with layers of equal optical thickness [Walsh 77]. Equation (2-51d) is composed of two terms: B_1 [in Eq. (2-51e)] and B_2 — Eq. (2-51f). Each term is squared. The LP defect function D_L is sum of absolute values

$$D_L \equiv |B_1| + |B_2| \ . \qquad (4\text{-}26)$$

The foregoing is termed the *Gershgorin norm* and is minimized by LP code at one or more (up to m) phase thicknesses $\beta_1, \beta_2, \beta_3, \ldots \beta_i \ldots \beta_m$. An advantage of LP is that it can be programmed to constrain refractive indices to attainable values. The amount that D_L is minimized at each β_i can be weighted, just as reflectances at different wavelengths are weighted in Eq. (10-2a).

4.9.4.1 Antireflection coating deposited on lithium niobate

An antireflection coating is designed for a material of relatively high index (n_s = 2.25), such as lithium niobate. It is widely used as a frequency doubler. The objective is to dereflect the substrate at its fundamental wave number λ_0^{-1} and its doubled wave number $2\lambda_0^{-1}$ with a band of low R surrounding each wave number. Due to symmetry of the stack (see §4.2.4), minimization of R at $\lambda / \lambda_0 = 0.667$ also minimizes R at $\lambda/\lambda_0 = 1.33$ — the doubled frequency. Its design is

air **Q Q Q D Q** *substrate* ,

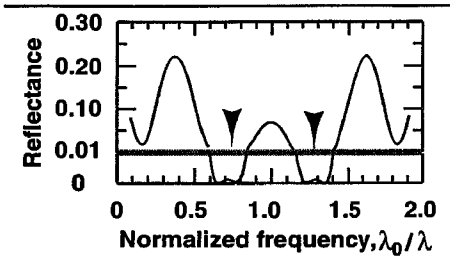


Fig. 4-52 $R(\lambda_0/\lambda)$ of the antireflection coating
air Q_1 Q_2 Q_3 D Q_5 *sub*
where indices of Q_1, Q_2, Q_3, D, Q_5 and *sub* are 1.38, 1.79, 2.03, 1.56, 1.79 and 2.25, respectively. Optical thicknesses of layers Q and D are $\lambda_0/4$ and $\lambda_0/2$, respectively. The ordinate changes scale at 0.01.

where Q and D represent layers of optical thickness $\lambda_0/4$ and $\lambda_0/2$, respectively. D_L was minimized at $\lambda/\lambda_0 = 0.55$ and 0.78 with weights of unity. In addition, it was minimized at $\lambda/\lambda_0 = 0.667$ with a weight of 2.5. Indices were constrained to lie in the range $1.38 \le n \le 2.15$. Results appear in Fig. 4-52. R is quite low at $\lambda/\lambda_0 = 0.667$ and at its double frequency, with a ± 10% bandwidth. It may be shown that LP produces a unique solution independent of starting design.

4.9.5 Antireflection coatings for nonnormal incidence

Any design method in Chap. 4 may be used at nonnormal incidence to produce an antireflection coating for a given polarization, provided effective indices (see §2.5.3.2) are substituted for indices that appear in the equations. The foregoing is unconditionally true when the substrate is nonabsorbing and layers are dielectric. §4.9.5.6, shows an approximation that may be used when the substrate is metallic.

4.9.5.1 An antireflection coating for both linear polarization states

This double-quarter antireflection coating reduces R_p and R_s to zero simultaneously at λ_0 [Baumeister 61]. This requires that the reduced admittance [see Eq. (2-48b)] is 1.00 for both polarizations

$$y_p = \frac{n_1^2\, n_s \cos^2\theta_2 \cos\theta_0}{n_2^2\, n_0 \cos^2\theta_1 \cos\theta_s} = 1.00 \qquad (4\text{-}27)$$

and

$$y_s = \frac{n_1^2\, n_s \cos^2\theta_1 \cos\theta_s}{n_2^2\, n_0 \cos^2\theta_2 \cos\theta_0} = 1.00 \qquad (4\text{-}28a)$$

where θ_0, θ_1, θ_2 and θ_s are incidence angles in the appropriate media (see Fig. 2-54). Since θ_1 and θ_2 are related to θ_0 by Eq. (2-4a), solution of the foregoing provides

$$n_1 = n_0 \sin\theta_0 \sqrt{(\cos\theta_s - \cos\theta_0\, n_0/n_s)/(\cos\theta_s - \cos\theta_0)} \quad . \qquad (4\text{-}28b)$$

Equation (4-2a) is used to find n_2. For air and BK-7 substrate, n_1 is 1.008 and n_2 is 1.357 at angles of incidence close to normal. Although the index of 1.357 is attainable, at angles close to grazing, the indices of 1.008 and 1.242 are both unattainable. Consider a high index substrate such as germanium (n_s = 4.0) and angle of incidence close to normal. Indices of 1.265 for n_1 and 2.53 for n_2 offer little hope, since the former is unattainable.

There is a situation when Eq. (4-28b) could be useful. Suppose BK-7 (n_0 = 1.52) is cemented to a flint glass substrate of index 1.75. At 60° angle of incidence (in BK-7) there is an annoying reflection of 0.5% at this interface and at 88° (near grazing incidence) R_s is 1.9%. At 88°, application of a coating in which n_1 is 1.527 and n_2 is

1.638 reduces R for both polarizations to below 0.07%. Optical thicknesses of the aforementioned layers should be *matched* (see §2.7.2.2) at the appropriate angle.

Table 4-7 Effective and actual refractive indices

	p polarization		*s* polarization	
Layer or medium	Actual index	Effective index	Actual index	Effective index
Incident medium	1.00	3.86	1.00	0.259
Layer*	2.56	2.76	1.11	0.551
Equivalent layer	1.422**H** 0.15**L** 1.422**H**		1.24**L** 0.422**H** 1.24**L**	
Substrate	1.52	1.97	1.52	1.174

At an incidence angle of 75° in air, effective indices for *s* polarization [Eq. (2-38a)] and *p* polarization [Eq. (2-38b)] for actual indices listed in column. *Index of layer is obtained from Eq. (4-29a) or Eq. (4-29b). Each equivalent layer produces an *effective* index of 2.76 or 0.551. Indices of **L** and **H** are 1.35 and 2.35, respectively.

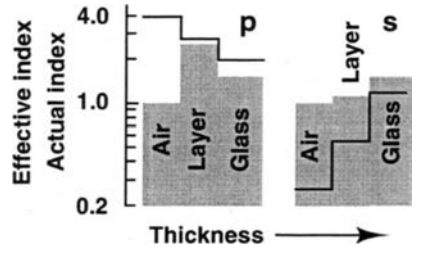


Fig. 4-53 At θ_0 of 75° in air, actual indices (*shaded*) and effective indices (*solid*) of a single-layer antireflection coating on glass of index of 1.52. (*Left*) *p* polarization and (*right*) *s* polarization.

4.9.5.2 Single-layer antireflection coatings at nonnormal incidence

The problem is to dereflect a BK-7 substrate when the ray is incident at 75° from air. Table 4-7 lists effective indices (see §2.5.3.2) of air, film and substrate. For *p* polarization, for example, effective indices of air and substrate are

$N_{0,p} = 3.86$ and $N_{sub,p} = 1.97$, respectively. In each case, the *effective* index of the layer $N_{e,p}$ or $N_{e,s}$ should satisfy Eq. (2-23j). For *p* polarization this is

$$N_{e,p} = (N_{0,p}\, N_{sub,p})^{1/2} = \sqrt{3.86 \times 1.97} = 2.76 \qquad (4\text{-}29a)$$

and for *s* polarization

$$N_{e,s} = (N_{0,s}\, N_{sub,s})^{1/2} = \sqrt{0.259 \times 1.174} = 0.551\ . \qquad (4\text{-}29b)$$

Given the effective index of 0.551, Eq. (2-38a) is used to find its actual index of 1.11 — a value that is nonattainable. Table 4-7 shows that the actual layer index for *p* polarization is 2.56 — an index that is difficult to attain in the visible portion of the spectrum.

There are separate antireflection coatings for each linear polarization. Inspection of effective indices and actual indices plotted in Fig. 4-53 shows that the antireflection coating for *p* polarization is a step-down (see §4.2.3.1) and a step-up for *s* polarization.

4.9.5.2.1 Use of equivalent indices

The synthesis procedure of §2.11.3 is used to produce the equivalent layers whose designs appear in Table 4-7. Solutions do not exist that have a 90° phase thickness at λ_0 of 633 nm. There is no alternative but to use 270° — a layer of three quarterwave

thickness. The center layer of the antireflection coating for p polarization has an optical thickness of 0.15, which is not producible. Step (4) of §2.11.3.1 points out that there is more than one solution when phase thickness is 270°. Figure 4-55 shows (*dashed*) a $R_p(\lambda)$ of such a solution.

4.9.5.3 Antireflection coatings that function in randomly polarized light

The *dashed* antireflection coating curves in Fig. 4-54 show that an average R [see Eq. (1-5a)] of about 13% is obtained. Multiple quarter designs (see §4.3.3) produce a slight improvement, as shown in Fig. 4-56. Figure 4-57 illustrates that computer optimization can reduce R of both polarizations to below 1%, albeit over a relatively narrow bandwidth, at an angle of 70°.

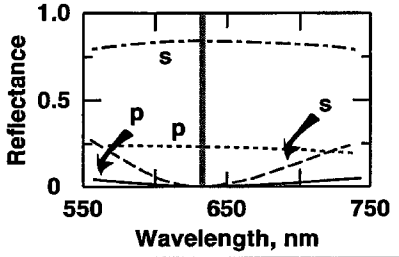


Fig. 4-54 At angle of incidence of 75°, $R_p(\lambda)$ (*solid*) and $R_s(\lambda)$ (*long-short*) of
air 1.422**H** 0.15**L** 1.422**H** *glass* .
$R_p(\lambda)$ (*long dash*) and $R_s(\lambda)$ (*short dash*) of
air 1.24**L** 0.422**H** 1.24**L** *glass* .
The refractive indices of **L**, *glass* and **H** are 1.35, 1.52 and 2.35, respectively. The optical thicknesses of **L** and **H** are $\lambda_0/4$ at λ_0 of 906 nm and 694 nm, respectively.

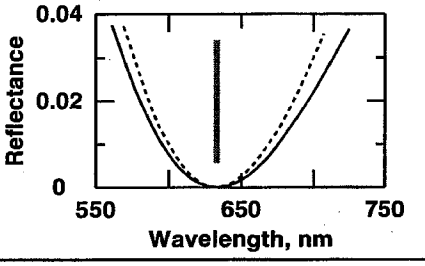


Fig. 4-55
At angle of incidence of 75°, $R_p(\lambda)$ (*solid*) of an antireflection coating of the design
air 1.422**H** 0.15**L** 1.422**H** *glass*
and (*dashed*)
air 0.578**H** 1.85**L** 0.578**H** *glass*
where the meaning of **L**, *glass* and **H** appear in the caption to Fig. 4-54.

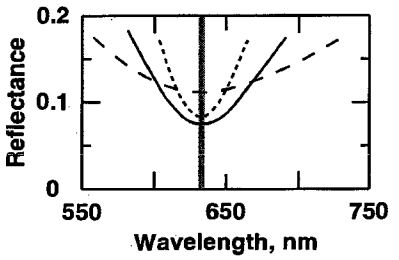


Fig. 4-56 At angle of incidence of 75°, reflectance $R_a(\lambda)$
air **L** (**H L**)^m *glass*
for $m = 1$ (*long dash*), $m = 2$ (*solid*) and $m = 3$ (*short dash*). Indices of **L**, *glass* and **H** are 1.35, 1.52 and 1.52, respectively. Optical thickness of **L** and **H** are $\lambda_0/4$ at λ_0 of 906 nm and 820 nm, respectively.

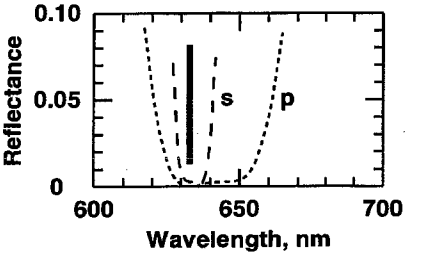


Fig. 4-57 At angle of incidence of 70°, $R_p(\lambda)$ (*short dash*) and $R_s(\lambda)$ (*long dash*) of
air 0.2384**L** 0.222**H** 0.091**L** 0.0455**H**
0.097**L** 0.0586**H** 0.1293**L** 0.073**H**
0.1434**L** 0.0667**H** 0.2283**L**
0.2757**H** 0.1596**L** 0.081**H** *glass* .
Indices of **L**, *glass* and **H** are 1.44, 1.52 and 2.30, respectively. Thicknesses are metric in μm. After DeBell [92].

4.9.5.4 Design with maximally flat spectral response

§4.5.1 mentions that even for a substrate of reasonably high index, such as germanium (n_s = 4.00), three-layer maximally flat coatings do not exist at normal incidence because of the lack of coating materials with sufficiently low refractive indices.

Consider a situation where such a coating is applied to a germanium substrate, which is to be dereflected at 45° incidence for *p* polarization. Effective index of air is sec(45°) = 1.414 and effective index of germanium is 4.06. It is possible to use a rare earth fluoride (index 1.35 at 10 μm) as the outer layer. From Table 4-3, effective index of the outer layer should be

$$n_1]_{effective} = (1.414)^{7/8}\ (4.06)^{1/8} = 1.61\ . \qquad (4\text{-}30)$$

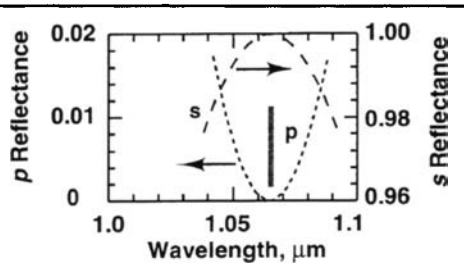


Fig. 4-58 At an incidence angle of 82.75°, $R_p(\lambda)$ (left ordinate, *short dash*) and $R_s(\lambda)$ (right ordinate, *long dash*) of the antireflection coating

air 1.114**H** 1.37**L** 1.114**H** 1.37**L** 1.114**H** *glass*

where indices of **H**, **L** and *sub* are 2.25, 1.45 and 1.508, respectively. The optical thickness of each layer is $\lambda_0/4$ at λ_0 of 1.065 μm.

At this angle of incidence, effective index of the rare earth fluoride is 1.58. Therefore, using either mixtures of materials (see §3.6.3) or equivalent layers (see §2.6.7), all indices for this three-layer maximally flat antireflection coating are attainable.

4.9.5.5 Antireflection coating for *p* polarization at 82° angle of incidence

This antireflection coating is the coating that appears in Fig. 1-176. It functions at an incidence angle of 82.75° in air and contains five quarterwave thickness (the thicknesses are *matched* — see §2.7.2.2) layers of high **H** and low **L** refractive indices of 2.25 and 1.45, respectively. This antireflection coating is designed by computing the reduced admittance — by substituting effective indices into Eq. (2-48c). A few trials shows that a five-layer stack **H L H L H** produces the requisite reduced admittance of 1.00.

Given the *glass* substrate index of 1.508, the effective indices (for *p* polarization) are 7.92, 2.00, 1.988 and 2.507 for *air*, *glass*, **L** and **H**, respectively. The reduced admittance is

$$y = \frac{(2.507)^2 \times (2.507)^2 \times (2.507)^2}{(1.988)^2 \times (1.988)^2 \times 7.92 \times 2.00} = 1.00. \qquad (4\text{-}31a)$$

Figure 4-58 shows that its *p* reflectance is quite low at the "tuned" wavelength and that the *s* reflectance is in excess of 99%. This antireflection coating is mentioned in §6.2.3.3 and is used as a reflector in the Ward polarizer (see Fig. 1-176, Fig. 1-177 and §7.9.7)

4.9.5.6 Dereflection of a metal at nonnormal incidence

Drude theory (see §2.2.4.2) states, in essence, that the normalized admittance of a metal at nonnormal incidence is approximated by its optical constant $n - j\,k$. Suppose a metal is to be dereflected at λ_0. The dielectric layer adjacent to it has an effective index (see §2.5.3.2) *N*. Use of the Drude approximation means that the following substitutions are made in Eqs. (2-46a) and (2-46b)

$n \rightarrow Y_r$, $-k \rightarrow Y_j$, $N \rightarrow n_1$. (4-31b)

Example: The metal mentioned in §4.4.3 is dereflected at 50° for *s* polarization, using dielectric layers of indices 1.45 and 2.30. Effective indices (see §2.5.3.2) of these materials are 1.231 and 2.169, respectively. It is assumed that a layer of index 2.30 is contiguous to the substrate. Solving for the phase thickness of this single layer, Eq. (2-46a) provides

$$\beta_1 = \tfrac{1}{2}\left[\pi + \arctan\left(\frac{-2\times 2.169\times(-3)}{(2.169)^2 - 2^2 - 3^2}\right)\right] \qquad (4\text{-}31c)$$

which evaluates to 61.3° or optical thickness (*matched* at 50°, see §2.7.2.2) of 0.17 λ_0. Values in Eq. (4-31b) are substituted into Eq. (2-46b) and the resulting normalized admittance is 0.62. Since the effective index of air is cos(50°) = 0.64, the match with this single layer is nearly perfect. An exact calculation shows that, notwithstanding use of the Drude approximation, application of a single **H** layer reduces R_s at λ_0 to less than 1%.

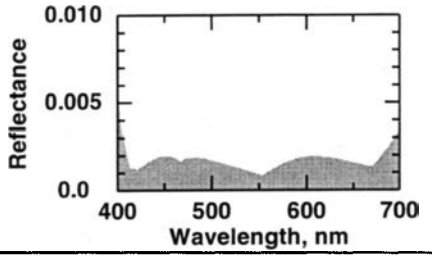


Fig. 4-59 Upper envelope of the $R(\lambda)$ curve for the antireflection

air 97.7**F** 30.09**H** 11**L** 64.93**H** 11**L** 19.1**H** 29.8**L** 11**H** *substrate*

for a *substrate* of refractive indices 1.52, 1.55 and 1.60. The refractive indices of **F**, **L** and **H** are 1.38, 1.45 and 2.30, respectively. The thicknesses of **F**, **L** and **H** are metric, in nm.

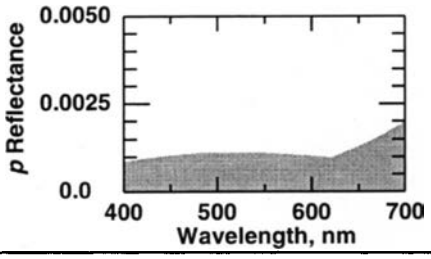


Fig. 4-60 Envelope of maximum *p* reflectance for angles of incidence (in incident medium of *glass*) in the range of 43° to 56° for the design

glass 24.7**M** 13.8**H** 29.5**M** 11.7**H** 167.6**M** *cement* .

Refractive indices of **H**, *glass*, **M** and *cement* are 2.10, 1.73, 1.60 and 1.55, respectively. Layer thicknesses are metric, in nm.

4.9.6 Antireflection coatings for multiple substrates

Suppose that the same antireflection is to be deposited on substrates of index 1.52, 1.55 and 1.60. Why would this be done? One reason is that different refractive index glasses could be coated in the same coating run, thereby economizing on the use of the coating equipment. Another possible reason is that the substrate is birefringent. This topic has been considered by several authors (Vvedenskii [82], Dobrowolski [96a], Dobrowolski [96b]). Modern computer optimization software, such as TFCalc [96], enables the user to optimize the layer thicknesses so that the reflectance is reduced simultaneously on three or more substrates of different refractive index. Figure 4-59 shows that a coating of eight layers effectively reduces the radiant reflectance to below 0.2% over the spectral region 420 nm to 680 nm.

4.9.7 Antireflection coating that matches optical cement to glass

This coating is useful when a glass prism (as in a MacNeille polarizer, see §6.3.3.5) is joined to an optical cement. An example is when the prism refractive index is 1.73 and the optical cement index is 1.55. Suppose that there is no coating. At an angle of 56° in the glass, the p reflectance is 2% at the glass-cement interface. Figure 4-60 shows that a five-layer antireflection coating on the glass reduces the p reflectance nine-fold over angles that range from 43° to 56°.

4.10 Appendix: Proofs, derivations and designs

This section focuses upon the mathematical details related to equations used in previous sections.

4.10.1 Two-layer unequal thickness

Both layers are nonabsorbing. There are two terms in Eq. (2-46d); both must be zero for R to be zero. This requires that

$$a_0 \cos\beta_1 \cos\beta_2 = a_2 \sin\beta_1 \sin\beta_2 \qquad (4\text{-}32a)$$

$$a_{01} \sin\beta_1 \cos\beta_2 = -a_{02} \cos\beta_1 \sin\beta_2 \ . \qquad (4\text{-}32b)$$

The foregoing reduce to

$$\tan\beta_1 \tan\beta_2 = a_0 \, a_2^{-1} \qquad (4\text{-}32c)$$

$$\tan\beta_1 \cot\beta_2 = -a_{02} \, a_{01}^{-1} \qquad (4\text{-}32d)$$

where the "a's" appear in Eqs. (2-46e) and (2-46f). Solution of Eqs. (4-32c) and (4-32d) leads to Eqs. (4-4), (4-5a) and (4-5b).

4.10.2 Maximally flat

Equations for indices of a three-layer maximally flat antireflection coating are derived. A general recursion formula is cited that applies to any number of layers. Coefficients a_1, a_2 and a_3 in Eq. (2-51c) must be zero if R / T is to take the form

$$F = (a_0 \cos^3\beta)^2 \ , \qquad (4\text{-}33a)$$

which represents a maximally flat design. Approximate values for coefficients in Eq. (2-52c) are used.. Extracting these from Table 2-15 provides

$$a_1 \approx 3u_0 - 2u_1 - 2u_2 - 2u_3 = 0 \qquad (4\text{-}33b)$$

$$a_2 \approx 3u_0 - 4u_1 + 4u_3 = 0 \qquad (4\text{-}33c)$$

$$a_3 \approx u_0 - 2u_1 + 2u_2 - 2u_3 = 0 \qquad (4\text{-}33d)$$

where u_0 is determined by the substrate index; u_1, u_2 and u_3 are unknowns. In addition, Eqs. (4-12d) and (4-13a) translate into

$$u_0 = u_1 + u_3 \text{ and } u_2 = u_0 / 2 \ , \qquad (4\text{-}33e)$$

respectively. Linear algebra produces

$$u_1 = 7u_0 / 8 \ . \qquad (4\text{-}33f)$$

Conversion of Eq. (4-33f) into n_0, n_s and n_1 using Eq. (2-53b) produces an entry in Table 4-3. Kard [57] shows that u_i for a maximally flat coating with ℓ layers is

$$u_{i+1} = u_i - u_0 \binom{\ell}{i} 2^{-\ell} \qquad (4\text{-}33g)$$

where brackets denote the binomial coefficient and i ranges from 0 to $\ell - 1$. An example is that with $\ell = 7$, $u_1 = 127/128 \, u_0$, $u_2 = 120/128 \, u_0$ and $u_3 = 99/128 \, u_0$.

4.10.3 Simultaneous and manifold solutions

The problem is to find refractive indices that produce zero R of a coating in which layers have identical optical thicknesses. In other words, the R / T polynomial in Eq. (2-51d) should be zero at one or more values of β. Regardless of the number of layers, however, R / T is a quadratic sum of two polynomials: the *real part* B_1 and the *imaginary part* B_2, as they appear in Eqs. (2-51e) and (2-51f), respectively. Both must be zero. Two classes of coatings satisfy these conditions: (a) simultaneous solutions; and (b) manifold solutions.

In the former class, the root of *real part* and the root of *imaginary part* occur simultaneously at a particular β. This requirement is quite restrictive. It is not surprising, therefore, that there are few combinations of refractive indices fulfilling these conditions. An example is the quarter-half coating described in §4.6.2 and §4.10.4.

Manifold solutions, on the other hand, require that either the *real part* or the *imaginary part* be identically zero, independent of β. Since the *real part* contains R / T of the uncoated substrate, at $\beta = 0$ it must be nonzero if $n_0 \neq n_s$. Therefore, if the incident medium and substrate are of different refractive index, a necessary (but not sufficient) condition for a manifold solution to exist is:

The imaginary part of R / T polynomial is identically zero, independent of β.

The foregoing is accomplished by constraining refractive indices. A consequence is that the roots are determined by the *real part* of the R / T polynomial. *If* roots exist, they occur as roots of a higher order polynomial. Thus, many roots exist, which justifies the name *manifold solution.* Constraints on refractive indices (see §2.7.4.4) produce manifold solutions.

4.10.4 Quarter-half coating

The layer next to the substrate has a halfwave optical thickness. The outer layer is a quarterwave. The phase thickness at which null R occurs is a *simultaneous solution* of the R / T polynomial, as mentioned in §4.10.3. This two-layer antireflection coating consists of three quarterwave layers in which the two bottom layers have identical refractive indices. Setting both terms in Eq. (2-51c) to zero produces the homogeneous equations

$$a_0 \cos^2\beta - a_2 \sin^2\beta = 0 \qquad (4\text{-}34a)$$

$$a_1 \cos^2\beta - a_3 \sin^2\beta = 0 \qquad (4\text{-}34b)$$

which requires that

$$a_0 a_3 - a_1 a_2 = 0 \ . \qquad (4\text{-}34c)$$

Setting $u_3 = u_2$ and using approximate coefficients in Table 2-15, Eq. (4-34c) becomes

$$u_0 (u_0 - 2u_1) - (3u_0 - 2u_1 - 4u_2)(3u_0 - 4u_1 + 4u_2) = 0 \ , \qquad (4\text{-}34d)$$

which is quadratic in u_2. Solving for u_2 results in Eq. (4-9b).

4.10.5 Reflectance zeros of QQ and QQQ coatings

The R / T polynomial of QQ design appears in Eq. (2-51b). a_1 is zero if indices satisfy Eq. (4-11a). Zero R requires that

$$a_0 \cos^2\beta = a_2 \sin^2\beta \ . \qquad (4\text{-}35a)$$

Substituting the approximate expression for a_2 from Table 2-15 results in

$$a_0 \cos^2\beta \, (\sin\beta)^{-2} \approx u_0 - 2\,u_1 + 2u_2 \ . \qquad (4\text{-}35b)$$

Conversion of Eq. (4-11a) into log index results in

$$u_0 = u_1 + u_2 \ . \tag{4-35c}$$

Using identity $\beta \equiv 90° - \xi$, solution of Eqs. (4-35b) and (4-35c) leads to Eq.(4-12a).

For the **Q Q Q** system, both a_1 and a_3 are zero in Eq. (2-51c) if indices satisfy Eqs. (2-52g) and (2-52h). Zero R requires that

$$a_0 \cos^2\beta = a_2 \sin^2\beta \ . \tag{4-36a}$$

Substituting the approximate expression for a_2 from Table 2-15 produces

$$a_0 \cos^2\beta \ (\sin\beta)^{-2} \approx u_0 - 4u_1 - 4u_3 \ . \tag{4-36b}$$

Using Eq. (4-33e) and identity $\beta \equiv 90° - \xi$, solution of Eq. (4-36b) leads to Eq. (4-13b).

4.10.6 Chebyshev antireflection coatings

Following Young [61], an antireflection coating containing ℓ layers with a Chebyshev response has the mathematical form

$$F = F_s \frac{\left[\mathcal{T}_\ell\left(\frac{C}{C_0}\right)\right]^2}{\left[\mathcal{T}_\ell\left(\frac{1}{C_0}\right)\right]^2} \tag{4-37a}$$

where $\mathcal{T}_\ell$ is a Chebyshev polynomial of the first kind of order ℓ, C_0 appears in Eq. (4-18b) and C is the cosine of β. F is the R / T ratio of the coating and F_s is the R / T of the uncoated substrate. The passband is the spectral region of low R, where $C \le C_0$. It follows from a property of the Chebyshev polynomial that R / T is bounded

$$0 \le F \le F_s \left[\mathcal{T}_\ell \left(1/C_0\right)\right]^{-2} \ , \tag{4-37b}$$

which is the ripple appearing in Eq. (4-18a).

As an illustration, equations for the indices for a four-layer antireflection coating are derived. Indices are found by equating B_1 in Eq. (2-51e) with Eq. (4-37a) and matching coefficients of $\cos^2\beta$ and $\cos^4\beta$ terms. Expansion of Eq. (2-51e) in $\cos\beta$ and equating it to Eq. (4-37a), with $\ell = 4$ provides

$$C^4(a_0 + a_2 + a_4) - C^2(a_2 + 2a_4) + a_4 =$$
$$C_0^4 \ a_0[\,8\,(C/C_0)^4 - 8\,(C/C_0)^2 + 1]\ P_0^{-1} \tag{4-37c}$$

where P_0 is

$$P_0 \equiv C_0^4 - 8\ C_0^2 + 8 \ . \tag{4-37d}$$

Table 2-15 provides approximate a values in Eq. (4-37c) in terms of the natural log of the index. Using these approximations and equating them to coefficients of powers of C in Eq. (4-37c) results in the linear equations

$$a_2 \approx u_0 P_0^{-1} \left(8C_0^2 - 2C_0^4\right) \approx 14\,u_0 - 12\ u_1 - 4\,u_2 \tag{4-37e}$$

$$a_4 \approx u_0 P_0^{-1}\ C_0^4 \approx u_0 - 4\,u_1 + 4\,u_2 \ . \tag{4-37f}$$

Solution of Eqs. (4-37e) and (4-37f) leads to

$$u_1 \approx \frac{u_0}{16}\left(15 - \frac{C_0^2\,(8 - C_0^2)}{C_0^4 - 8\,C_0^2 + 8}\right) \tag{4-38a}$$

and

$$u_2 \approx \frac{u_0}{16}\left(11 - \frac{C_0^2\,(8 - 5\,C_0^2)}{C_0^4 - 8\,C_0^2 + 8}\right) \ , \tag{4-38b}$$

which are identical to Eq. (4-19a) and entries in Table 4-6. A test that the algebra is correct is that $u_1 \rightarrow 15u_0/16$ and $u_2 \rightarrow 11u_0/16$, as $C_0 \rightarrow 0$, which corresponds to the entries in Table 4-3 for the maximally flat condition — i.e., zero bandwidth.

4.10.7 Zero reflectance of quarter-half-quarter antireflection coating

This is a degenerate four-quarter coating in which the pair of non-outer layers have identical refractive index. It is a system with a simultaneous solution of R / T polynomial (see §4.10.3). Zero R requires that both terms in Eq. (2-51d) are zero

$$u_0 C^4 - a_2 C^2 S^2 + a_4 S^4 = 0 \qquad (4\text{-}39a)$$

and

$$a_1 C^2 - a_3 S^2 = 0 \qquad (4\text{-}39b)$$

where

$$a_2 \approx 6u_0 - 6u_1 + 6u_3 \ , \qquad (4\text{-}39c)$$

$$a_4 \approx u_0 - 2u_1 + 2u_3 \ , \qquad (4\text{-}39d)$$

$$a_1 \approx 4u_0 - 2u_1 - 4u_2 - 2u_3 \ , \qquad (4\text{-}40a)$$

$$a_3 \approx 4u_0 - 6u_1 + 4u_2 - 6u_3 \qquad (4\text{-}40b)$$

and where C and S appear in Eq. (4-21a).

Equations (4-39c) and (4-39d) do not contain u_2. Given β at which the root occurs, Eqs. (4-39c) and (4-39d) are substituted into Eq. (4-39a) and log index u_3 is found, as in Eq. (4-20d). Then Eqs. (4-40a) and (4-40b) are substituted into Eq. (4-39b) and the remaining unknown, u_2, is found — this produces Eq. (4-21c). Nagendra [83] derived an explicit equation for R, but failed to derive an equation for β at which $R = 0$.

4.10.8 Appendix — designs of antireflection coatings

Listed here are the designs of the antireflection coatings whose spectral reflectance curves appear in Fig. 4-50. For the simulated design (*solid* curve):

air 0.479**L** 0.037**M** 0.678**L** 0.094**M** 0.397**L** 0.094**M** 0.359**L** 0.164**M** 0.322**L** 0.164**M** 0.284**L** 0.238**M** 0.247**L** 0.238**M** 0.2**L** 0.334**M** 0.153**L** 0.334**M** 0.193**L** 0.742**M** 0.116**L** 0.458**M** 0.069**H** 0.634 **M** 0.127**H** 0.352**M** 0.127**H** 0.305**M** 0.216**H** 0.258**M** 0.216**H** 0.21**M** 0.317**H** 0.159**M** 0.317**H** 0.08**M** *Ge*

The refractive indices of **L**, **M**, *Ge* and **H** are 1.35, 2.20, 4.0 and 4.2, respectively. The sequence of the indices in the ten-layer design (*dashed* curve in Fig. 4-50) are: ten layers with refractive indices

air/1.39/1.50/1.62/1.75/1.92/2.12/2.37/2.67/3.03/3.46/*Ge*

All optical thicknesses are multiples of λ_0 /4 at λ_0 of 2.0 μm.

4.11 Additional information about antireflection coatings

Review articles on antireflection coatings have been written by Cox and Hass [Cox 64b], as well as by Musset and Thelen [Musset 70]. Table 4-8 lists books and review articles where various topics may be found. The text of Macleod [86] analyzes antireflection coatings in terms of admittance diagrams, whereas the book of Thelen [89] eschews graphical methods. Mouchart [77a, 77b, 77c, 77d, 78b] authored a series of articles on antireflection coatings. His treatment of the buffer layer [Mouchart 78a] is especially helpful in the design of antireflection coatings at nonnormal incidence. Costich [87] describes antireflection coatings used in lasers and frequency doublers.

Readers with a mathematical bent should relish Chap. 6 of Knittl [76], which couches antireflection coating design in terms of electrical circuit theory. Chebyshev

antireflection coatings — similar to those appearing in §4.7 — are designed by positioning roots of impedance polynomial at their proper locations in the *s* plane.

Table 4-8 Cross reference of books and review articles

Topic	This book	C&H	M&T	M	T
Two-layer **Q Q** — zeros of *R*	§4.6.3.1	§2	§5	p. 96	§4.2
Two-layer unequal thickness	§4.3.4	§2	§3	p. 95	§4.1
Three-layer **Q Q Q** — zeros of *R*	§4.6.3.2	§3	§6	p. 125	
Quarter-half **Q D**	§4.5.2.1	§3	§3	p. 104	§4.
Quarter-half-quarter **Q D Q**	§4.8.3	§3	§7	p. 113	§4.3
Use of equivalent layers	§4.8.5.2	§5	§8	p. 118	§4.6
Graded index	§4.9.2	§6	§6		

References to specific antireflection coatings in this book (*This*) and other publications, such as Cox [64b] (*C&H*), Musset [70] (*M&T*), Macleod [86] (*M*) and Thelen [89] (*T*). **D** and **Q** represent nonabsorbing layers of optical thicknesses $\lambda_0/2$ and $\lambda_0/4$, respectively.

Chapter 5
Reflectors, edge filters and periodic structures

5.1 Introduction

Periodic structures are the building blocks of laser mirrors, bandpass filters, rejection filters and edge filters. The concept of such a structure furnishes insight into the functioning of such multilayers. An example is a stack of 20 layers

air $\mathbf{H}_1\,\mathbf{L}_1\,\mathbf{H}_2\,\mathbf{L}_2\,\mathbf{H}_3\,\mathbf{L}_3\,\mathbf{H}_3\,\mathbf{L}_4\,\mathbf{H}_5\,\mathbf{L}_5\,\ldots\,\mathbf{H}_{10}\,\mathbf{L}_{10}$ *substrate*

where each layer has a different thickness and possibly a different refractive index.

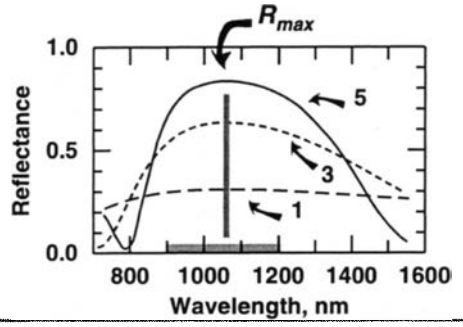


Fig. 5-1 $R(\lambda)$ of the stack
(1): *air* **H** *glass* (long dash),
(3) *air* **H L H** *glass* (short dash)
(5) *air* **H L H L H** *glass* (solid)
where the refractive indices of *glass*, **L** and **H** are 1.51, 1.45 and 2.30, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 1060 nm. The shaded area is the stop band.

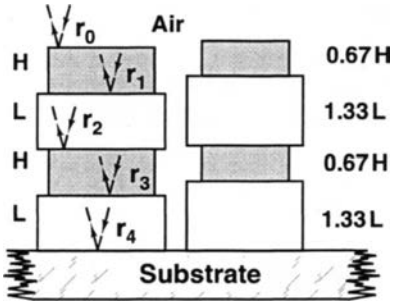


Fig. 5-2 For a quarterwave stack (*left panel*) and 2:1 stack (*right panel*), layers are drawn proportional to their optical thicknesses. The Fresnel amplitude reflection coefficients at each interface are shown as well as the substrate S. The optical thickness of **L** and **H** is $\lambda_0/4$.

Although the digital computer can readily compute its reflectance, it is difficult to gain insight into how that multilayer functions because of the overwhelming number of variables. If, however, the coating contains only two types of layers **H**, **L**

air **H L H L H L H L H L H L H L H L H L H L** *substrate* , which may be written

air $(\mathbf{H}\,\mathbf{L})^{10}$ *substrate* ,

then a good idea of the performance of the 20 layer stack is obtained by analyzing the properties of the **H L** basic period iterated 10 times. An additional advantage of treating this as a periodic structure is that efficient algorithms are used to compute its reflectance with improved computational efficiency (see §5.8.1).

Figure 5-1 shows the results of adding **HL** pairs to a stack. Each layer has a quarterwave optical thickness. Starting with a single **H** layer, the maximum reflectance is 31%. The addition of an **HL** pair to form **HLH** increases the maximum reflectance to 63%. Augmenting the stack with still another **HL** pair to produce **HLHLH** boosts the maximum reflectance to 85%.

§5.1.1 and subsequent sections explain the performance of the stacks in terms of a vector diagram. §5.2 develops the theory of the stop band of a periodic structure.

5.1.1 Examples — use of vector diagrams to calculate reflectance

The next few sections present examples of multilayer reflectors. They are analyzed using the vector addition of amplitudes (see §2.8.3). The goal is to gain insight. A warning: The net amplitude reflection coefficient in several instances exceeds 1.00. The numerical results are clearly incorrect.

5.1.1.1 Use of vector diagrams — goals

Given a stack of alternating high index and low index layers, the vector diagrams in §5.1 illustrate that:

- A quarterwave (QW) stack provides the high reflectance.
- Non-quarterwave stacks require more layers to provide a reflectance comparable to that of a quarterwave stack.
- The spectral region of high reflectance narrows as the ratio of high index to low index diminishes.
- As the ratio of high index to low index is reduced, a greater number of layers is needed to achieve a given level of reflectance.

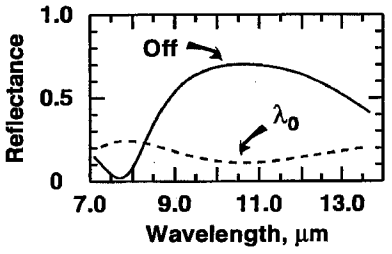


Fig. 5-3 $R(\lambda)$ of the quarterwave stack (*solid*)
air **H L H L** *substrate*
and (*dash*)
air **L H L H** *substrate*
where **L**, **H** and *substrate* represent media of refractive indices 1.40, 2.22 and 2.40, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 10.6 μm. The "off" arrow is the wavelength in panel *D* of Fig. 5-4 .

5.1.2 Example — output coupler for a carbon dioxide laser

The flux generated by a carbon dioxide laser is extracted from the cavity through a mirror with relatively low reflectance which is termed an *output coupler* OC. Although there are many wavelengths in the spectrum of the laser, the mirror is optimized at a wavelength of $\lambda_0 = 10.6$ μm.

5.1.2.1 Four-layer coating

Figure 5-2 depicts an OC consisting of four alternate layers of low index material (thorium fluoride, $n_L = 1.40$) and a high index substance (zinc sulfide, $n_H = 2.22$) coated on a zinc selenide substrate with an index n_s of 2.40. Representing the high and low index layers by **H** and **L**, respectively, its design is

air **H L H L** *substrate* .

Figure 5-3 shows that its reflectance peaks at 76% at the "tuned" wavelength λ_0 and drops to lower values at both longer and shorter wavelengths.

Caveat: At the date that this note was inserted (1999), thorium fluoride had largely been phased out in the USA as a coating material because of its radioactivity. It was still in use, however, by a few stalwarts, who have insisted on using it for mirrors that reflect the flux from high power carbon dioxide lasers.

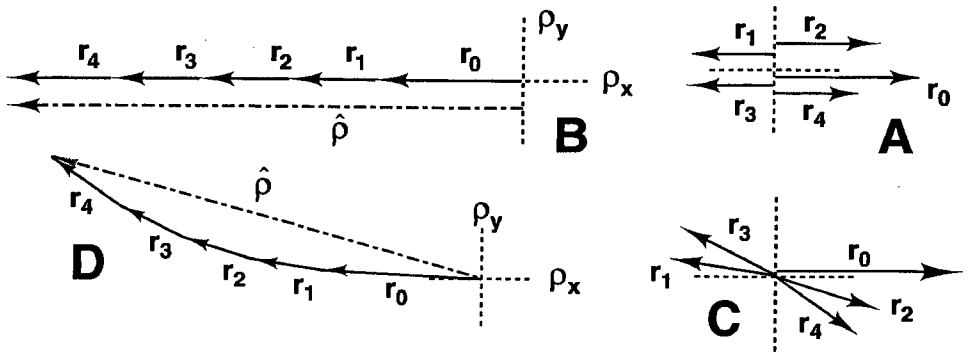


Fig. 5-4 Vector addition of amplitudes is used to compute the amplitude reflection coefficient $\hat{\rho}$ (*long-shorter dash*) of the quarterwave stack

air **H L H L** *substrate*

where the symbols are captioned in Fig. 5-3 . In panels *A* and *C*, the algebraic signs of the Fresnel amplitude reflection coefficients are ignored. For panels *A* and *B*, $\lambda = \lambda_0$. $\lambda = 0.95\ \lambda_0$ in the *C* and *D* panels.

The Fresnel amplitude reflection coefficients at each interface are shown (as vectors) in Fig. 5-2 . When these are added with the proper phase delays due to the layer thicknesses, the result is the vector diagram in Fig. 5-4 . Panels *A* and *B* obtain at $\lambda = \lambda_0$. The *A* panel depicts the layout of the vectors after execution of steps (*a*) to (*d*) in §2.8.3.2. The phase angle of each vector is correct but the signs of the amplitude reflection coefficients are ignored. The *B* panel shows the vectors after the aforementioned signs are taken into account. r_0 is a "rare-to-dense" reflection because the incident medium of air is of lower index than the outer layer of zinc sulfide. r_0 is negative—it lies along the negative real axis. When the signs and phases of the other vectors are taken into account, all vectors add co-linearly with r_0. A consequence is that the maximum length of the net amplitude reflection coefficient $\hat{\rho}$ occurs.

Panels *C* and *D* displays the vectors when the wavelength has shifted to 10.1 μm. In this instance, the 2β term (see step (*c*) of §2.8.3.2.1] is $-2\beta_1 = -180° \times 10.6/10.1 = -189°$. The net result is that the vector representing r_1 is 9° from the horizontal axis, as drawn in panel *C* of Fig. 5-4 . The 9° shift is cumulative with the other vectors. This is illustrated by the vector r_2, which is 18° from the horizontal axis. The net result is that the vectors are no longer collinear, as is the case when $\lambda = \lambda_0$, but rather curl upwards, so that the length of the net amplitude reflection coefficient $\hat{\rho}$ shrinks. This is manifested by the decrease in reflectance at the arrowed wavelength of the upper curve in Fig. 5-3 .

5.1.2.1.1 Optimal reflectance of a quarterwave stack

Given the substrate, available indices and number of layers, an *optimal* coating produces the maximum reflectance at $\lambda = \lambda_0$. The quarterwave stack in §5.1.2.1 is optimal. Several questions remain: (1) Why are the optical thicknesses quarterwaves, and (2) why are the layers ordered so that there is a high index next to air and a low index contiguous to the substrate? The vector diagram in Fig. 5-4 makes it clear that the maximum reflectance is obtained when the optical thicknesses are quarterwaves—when the vectors

are collinear. The four layers are arranged so that the maximum "kick" in the Fresnel amplitude reflection coefficients is obtained at each interface.

5.1.2.2 Non-optimal quarterwave reflector

Figure 5-3 shows the reflectance of the coating with the same layers but with a different sequence — an **L** is the outer layer

air **L H L H** *substrate* .

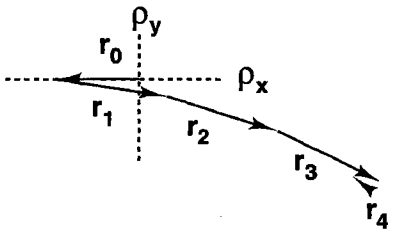


Fig. 5-5 The caption to Fig. 5-4 obtains, with the exceptions that the coating is

air **L HL H** *substrate*

and $\hat{\rho}$ is not shown. The wavelength is 0.95 λ_0 .

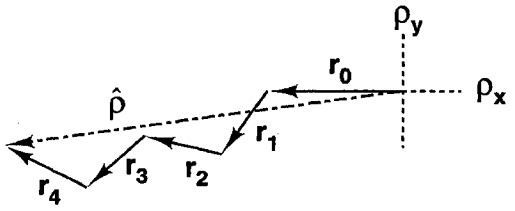


Fig. 5-6 Vector addition of amplitudes is used to compute the amplitude reflection coefficient $\hat{\rho}$ (*long-short dash*) of the reflector

air 0.67**H** 1.33**L** 0.67**H** 1.33**L** *substrate*

where the meaning of the symbols is the same as in Fig. 5-2 .

Its reflectance is substantially less than that of the coating of §5.1.2.1. The vector diagram in Fig. 5-5 provides insight as to why this occurs. The vectors r_0 and r_1 are in opposite directions. This is because *both* vectors represent rare-to-dense reflections and there is a 180° phase difference. This is in contrast to the coating in Fig. 5-4 , where the vectors are in the *same* directions because r_1 represents a dense-to-rare reflection. In addition, the index of the bottom layer of the stack nearly matches that of the substrate. The consequence is that not only is the magnitude of r_4 quite small, but in addition it is antiparallel to r_3. Figure 5-3 shows that its maximum reflectance is less than half that of the stack in which **H** is the outer layer.

Regardless whether the arrangement of the layers produces an optimal or non-optimal reflectance, the stacks mentioned in the previous sections are termed *quarterwave* stacks, or QW stacks for short. All layers have identical optical thicknesses.

5.1.3 Nonquarterwave coatings

Although the optimal reflectance is produced when the layer thicknesses are quarterwaves (see §5.1.2.1.1), nonquarterwave systems nevertheless produce a substantial reflectance.

5.1.3.1 Two-to-one stack

Consider the "two-to-one" stack

air 0.67**H** 1.33**L** 0.67**H** 1.33**L** *substrate*

The optical thickness of the 1.33**L** layer is twice that of the 0.67**H** layer, as depicted in Fig. 5-2 . The sum of the optical thickness of 0.67**H** and 1.33**L** is two QWOTs — a halfwave — at 10.6 μm. Compare the vector diagram for this coating (in Fig. 5-6) with the quarterwave stacks of Fig. 5-4 . Whereas at a wavelength near λ_0, the vectors are nearly collinear for the quarterwave stack, in this instance the mismatch of phase thickness

causes the vectors to zigzag. The result is that the peak reflectance of the two-to-one stack is less than the quarterwave stack, as illustrated in Fig. 5-7 .

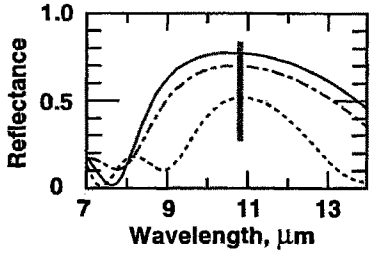


Fig. 5-7 $R(\lambda)$ of the quarterwave stack reflector (*solid*)
air **H L H L** *substrate* ,
2:1 stack (*long-short dash*)
air $(0.667\mathbf{H}\ 1.333\mathbf{L})^2$ *substrate*
and 6:1 stack (*short dash*)
air $(0.286\mathbf{H}\ 1.714\mathbf{L})^3$ *substrate*
where **L**, **H** and *substrate* represent media of refractive indices 1.40, 2.22 and 2.40, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 10.6 μm. 2:1 and 6:1 refer to the optical thickness ratio of the **H** and **L** layers.

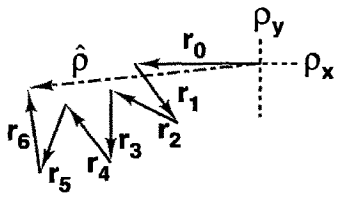


Fig. 5-8 Vector addition of amplitudes computation of the amplitude reflection coefficient $\hat{\rho}$ (*long-short dash*) for the reflector
air $(0.286\mathbf{H}\ 1.714\mathbf{L})^3$ *substrate*
where the symbols are captioned in Fig. 5-7 . The wavelength is 0.93 λ_0 . ρ_x and ρ_y are the Cartesian coordinates of $\hat{\rho}$.

5.1.3.2 Six-to-one stack

The final example is a six-layer stack

air $(0.286\mathbf{H}\ 1.714\mathbf{L})^3$ *substrate*

in which the optical thickness of the 1.714**L** layer is 6 times that of the 0.286**H** layer. The vector diagram appears in Fig. 5-8 . The zigzag of the vectors, so noticeable in Fig. 5-6 , is now quite accentuated because of the substantial mismatch of the phase thicknesses of layers. A consequence is that peak reflectance shown in Fig. 5-7 of the six (mismatched) layers is at least 20% less than the reflectance of the four-layer quarterwave stack.

5.1.3.3 Uses of such stacks

A logical question is that since the quarterwave stack produces an optimal reflectance, why should any other type of stack be used? Several reasons are: (1) Changing layer thicknesses alters the mean mechanical stress of a stack by adjusting the tension (see §9.7.3.2). (2) If one of the thin film materials has a high optical absorption, the absorption of the stack is minimized by using less of that material (see §5.3.10.5). (3) Harmonic reflection peaks (occurring at wavelengths of $\lambda_0/2$, $\lambda_0/3$, $\lambda_0/4$ and so on) are sometimes required (see §5.3.7). (4) It is necessary that one or more harmonic reflection peaks *not* be present (see §5.3.4.2 and §5.3.5). (5) A narrower reflectance bandwidth is needed (see §5.1.4).

5.1.4 Narrowing of the bandwidth

The bandwidth is that spectral region where the reflectance is "reasonably large." A more rigorous definition appears in §5.2.3.1. The bandwidth is relatively broad for the quarterwave stack shown in Fig. 5-3 ; it extends to wavelengths ±10% of λ_0 . Two methods of narrowing the bandwidth are: (1) The two materials used in the stack have refractive indices that are relatively close together (see §5.3.3.1); and (2) the ratio of the layer thicknesses is large.

5.1.4.1 Use of refractive index to narrow the reflecting spectral region

The goal is to develop a coating that reflects over a narrow spectral region centered at the 633 nm laser line and has low reflectance in other regions. Figure 5-9 shows the reflectance of the quarterwave stacks

air $(\mathbf{H}\,\mathbf{L})^q$ *glass* .

The lower curve is for 14 layers and the upper for 34. The indices of **H** and **L** are 1.46 and 1.38, respectively, which are relatively close to each other.

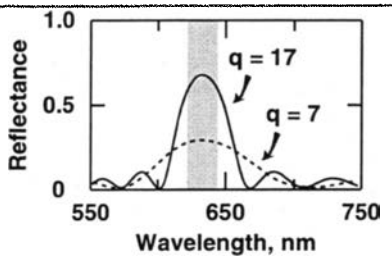


Fig. 5-9 $R(\lambda)$ of the reflectors *air* $(\mathbf{H}\,\mathbf{L})^q$ *glass* for $q = 7$ (*dashed*) and $q = 17$ (*solid*). **H** and **L** represent layers of optical thickness $\lambda_0/4$ and refractive indices of 1.46 and 1.38, respectively. λ_0 is 633 nm and the *glass* index is 1.52. The stop band spectral region is shaded.

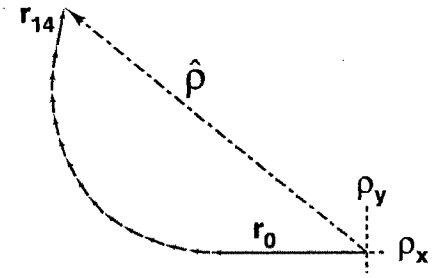


Fig. 5-10 Vector addition of amplitudes computation of the amplitude reflection coefficient $\hat{\rho}$ (*long-short dash*) for the 14 layer coating captioned in Fig. 5-9 . The wavelength is 609 nm.

The cause of the narrow bandwidth is suggested by the vector diagram for the 14 layer stack in Fig. 5-10. As compared to the vectors for the four-layer stack in Fig. 5-4 , there are many more interfaces (and, therefore, many more vectors) in the 14 layer coating. A consequence is that the cumulative effect of the small phase difference between the vectors causes the string of vectors to "curl." In Fig. 5-10, for example, the wavelength is 609 nm and the phase of the vector r_1 is $-2\beta_1 = -180° \lambda_0/\lambda = -180° \times 633/609 = -187°$. This means that there is a scant difference of $180° - 187° = 7°$ between the angles of each of the vectors. Yet the cumulative effect of this 7° with 14 vectors produces a phase difference of $14 \times 7° = 98°$ — clearly, the final vector is directed vertically upwards in Fig. 5-10.

5.1.4.2 Use of thickness mismatch

The quarterwave stack — in which the ratio of the optical thicknesses is 1:1 — manifests the broadest spectral region in which the reflectance is relatively large. Any stack with other thickness ratios has a narrower spectral region of high reflectance.

An example is the six-to-one stack mentioned in §5.1.3.2. Figure 5-11 shows that the use of 20 layers produces a peak reflectance in excess of 95%. This contrasts with the 10 layer quarterwave stack in the same Fig. It manifests a peak reflectance in excess of 98% and, of course, a broader region of high reflectance. The conclusion is that a narrowing of the reflectance band can be achieved by a thickness mismatch, but only at the expense of using more layers to achieve a reflectance comparable to that of a quarterwave stack. Further examples appear in §5.2.4.3.2.

5.2 Analysis of the basic period

5.2.1 Introduction

§5.1 mentions the concept of the *basic period.* This is a group of layers that is to be iterated many times. An example is the design of the six-to-one stack in Fig. 5-11, which is written as a two-layer basic period iterated ten times

air $(0.283\mathbf{H}\ 1.717\mathbf{L})^{10}$ *substrate.*

§5.2 introduces the following concepts: (1) Definitions of the stop band and passband; (2) the spectral width of a stop band and the reflectance envelope; and (3) the *order number* of a stop band. In addition, an appendix (see §5.8.1) develops algorithms for the exponentiation of a matrix of a basic period.

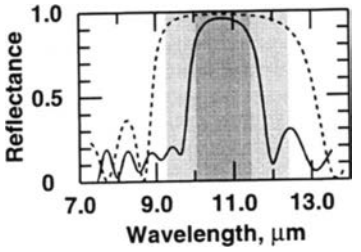


Fig. 5-11 $R(\lambda)$ of the coating (*solid* curve, darker shaded stop band spectral region)

air $(0.286\mathbf{H}\ 1.714\mathbf{L})^{10}$ *substrate*

and (*dashed,* lighter shaded stop band spectral region)

air $(\mathbf{H}\ \mathbf{L})^{5}$ *substrate*

where the refractive indices of **H**, *substrate* and **L** are 2.40, 2.22 and 1.40, respectively. The optical thicknesses of **H** and **L** are $\lambda_0/4$ at λ_0 of 10.6 μm.

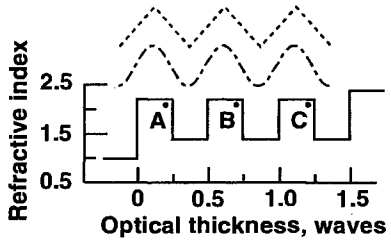


Fig. 5-12 Refractive index vs optical thickness of periodic media composed of homogeneous layers (*solid*), triangular index profile (*short dash*) and sinusoid (*long-short dash*).

5.2.2 Stop band and passband — conceptual

There are two ways of defining a stop band or passband. The first is intuitive and devoid of mathematical rigor. Consider a basic period surrounded on both sides by air

air (basic period)^q *air* .

How does the reflectance behave as more basic periods are added — as the exponent q progresses 1, 2, 3 ...? The stop band is that spectral region in which the reflectance *always* increases as more periods are added. Regions outside the stop bands are termed *passbands.*

The next definition is more rigorous. Suppose a steady state plane wave (wavelength λ) is incident upon a periodic structure. Imagine that it is actually possible to probe into the structure and measure the rms electric field E at the same relative position in each period, as, for example, at A, B and C, as shown in Fig. 5-12. As the probe is moved deeper into the multilayer, if the E field always (i.e., monotonically) decreases, then a stop band exists at that wavelength. If the E field oscillates — sometimes increasing and sometimes decreasing — then a passband exists.

One of the periodic structures in Fig. 5-12 consists of homogeneous layers — a "square wave" modulation of the refractive index. The reflector in Fig. 5-2 is this type of profile. The basic period could just as well be composed of inhomogeneous layers. Figure 5-12 depicts, for example, coatings with sinusoidal and triangular refractive index profiles. The former is of intense interest to holographers.

5.2.3 Quantitative definition of the stop band

An appendix (see §5.8.1) reviews the mathematical procedure for the exponentiation of a matrix representing a basic period. §5.2.3 explores methods of determining the location of a stop band or passband.

5.2.3.1 Matrix methods

The stop band spectral region of a nonabsorbing stack is found as follows: (1) Given the design of the basic period and wavelength, compute for this period its characteristic matrix [see Eq. (2-33a)]. (2) Find the trace x of the matrix [see Eq. (2-41e)].

When

$$|x| > 2 \tag{5-1a}$$

a stop band exists. A passband exists when $|x|$ is less than 2. The equation

$$|x| = 2 \ . \tag{5-1b}$$

establishes the demarcation between a stop band and a passband. In the parlance of solid state physics, this is the location of the Brillouin zone in a one-dimensional periodic structure.

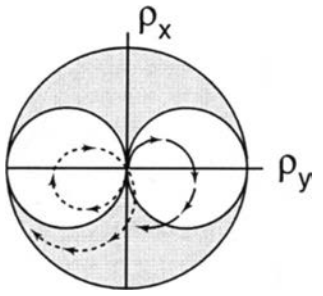


Fig. 5-13 Parametric plot of the amplitude reflection coefficient vs phase thickness of the equivalent layers (*solid*)
medium 0 5H' L' 0.5H' *medium*
and (*dashed*)
air 0 5HL 0.5H *air*
where the refractive indices of H', H, L', L and *medium* are 2.30, 4.20, 1.38, 1.40 and 3.00, respectively. The phase thickness of L and L' is in the range $0° \leq \beta \leq 90°$. For the layers 0.5H and 0.5H' it is $0° \leq \beta \leq 45°$. The shaded area is the stop band.

The foregoing procedure is also valid if the Kard matrix [see Eq. (2-29h)] is computed for the basic period. Equation (2-41e) then refers to the diagonal elements of the Kard matrix.

5.2.3.2 Use of the amplitude reflection coefficient

The basic period contains only nonabsorbing layers and is symmetrical. The basic period may be represented by an equivalent layer (see §2.6.7). The procedure follows from Dawirs [55]: (1) Surround the basic period by the same media so that $n_0 = n_s$. (2) Compute the amplitude reflection coefficient $\hat{\rho}$ of the basic period and plot it on the amplitude reflection plane, as illustrated in Fig. 5-13. The circle of radius $|\hat{\rho}| = 1$ — the unit circle — is shown.

In the first approach, Dawirs [55] has a procedure for finding n_e from the admittance of the equivalent layer deposited on a nonabsorbing substrate. Matthaei [64] reproduces the chart of Dawirs which, similar to the Smith chart, enables the user to plot $\hat{\rho}$ in the amplitude reflection plane and directly read n_e and β_e.

In the second approach [Epstein 51], given the complex amplitude reflection coefficient of the equivalent layer $\hat{\rho}$, its reciprocal is converted into Cartesian coordinates

$$\hat{\rho}^{-1} \equiv u + jv \ . \tag{5-1c}$$

It follows that

$$n_e = n_0 \sqrt{\frac{u-1}{u+1}} \tag{5-1d}$$

and β_e is

$$\beta_e = \arctan\left(\frac{2\, n_0}{v\,(n_e - n_0^2\, n_e^{-1})} \right) . \qquad (5\text{-}1e)$$

The following steps refer to the amplitude reflection plane mentioned above. Draw two circles, each with a radius of 0.5 and with the centers on the horizontal axis. The coordinates of one center are (–0.5, 0) and of the other (0.5, 0). Figure 5-13 illustrates that those two circles osculate the unit circle and also each other at the origin. They are termed the *passband circles* in the ensuing discussion.

At a given wavelength, a passband exists if the amplitude reflection coefficient of the basic period lies within either of the aforementioned *passband circles*. Otherwise, a stop band exists. As an example, Fig. 5-13 plots parametrically the amplitude reflection coefficient of the coating

air 0.5 **H L** 0.5 **H** *air* .

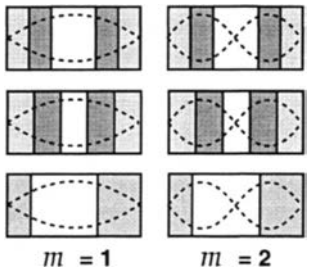


Fig. 5-14 In a basic period, the optical thickness of each layer of is shown in comparison with a half wavelength at the first-order (m = 1) and second-order (m = 2) stop bands. The designs are (*upper*) **A x 2B x A** , (*middle*) **A x B x A** and (*lower*) 0.35**L H** 0.65**L** .

Most of the points are inside the *passband circles* (*shaded*) and, therefore, are within the passband region.

5.2.3.3 Optical thickness rule

The following rule is useful in locating a stop band. First, compute the sum of the optical thicknesses OT of the ℓ layers in the basic period

$$\mathrm{OT} = \sum_{i=1}^{\ell} n_i h_i \ . \qquad (5\text{-}2)$$

A necessary (although not sufficient) condition for a stop band to exist at a wavelength λ_0 is that

$$\mathrm{OT} = m\, \lambda_0 / 2 \qquad (5\text{-}3)$$

where m is an integer — the order of the stop band.

A formal proof of the foregoing is omitted. It is essentially identical to the Bloch solution to the Schrödinger equation for a periodic potential. This appears in texts on the physics of solids (example: Kittel [86]). Equation (5-3) is represented graphically in Fig. 5-14, which plots the optical thickness of several basic periods with superimposed sine waves. At least it is *possible* that a stop band exists for each of the coatings in that Fig.

5.2.3.4 Order number of a stop band

Suppose that a halfwave of λ_0 "fits" into the basic period, as shown in the left panel of Fig. 5-14. According to Eq. (5-3), this is the *necessary* condition for the first-order stop band to exist. Shorter wavelengths of λ_0 / m also "fit " into the basic period. These are the higher order stop bands— m is termed the *order number*. Figure 5-14 shows the wavelength at both the first and second-order stop bands.

Consider the analogy between the harmonic of a vibrating system and the order number. One-half of an acoustic wavelength "fits" into an organ pipe at its resonant

frequency — the fundamental. Integral multiples of that half wavelength are the harmonics of that fundamental. The higher order stop bands are harmonics of the wavelength of the first-order stop band. §5.2.5 presents more examples.

5.2.4 Spectral width of a stop band

Many reflectors and rejection filters contain only two materials. There are relatively simple equations for the location and spectral width of the stop bands of such coatings.

5.2.4.1 Basic period containing two components

The basic period contains only two refractive indices—n_a and n_b. Another parameter is Λ — the ratio of the phase thicknesses [see Eq. (2-47b)]. The width of a stop band depends upon both the index and thickness ratios.

5.2.4.2 Equation for the stop band location

§5.2.3.1 states that a stop band *may* occur when Eq. (5-1b) is satisfied. For a nonabsorbing stack in which the phase thicknesses of the components are β_a and β_b the trace is:

$$x = 2\cos\beta_a \cos\beta_b - \left(\frac{n_a}{n_b} + \frac{n_b}{n_a}\right)\sin\beta_a \sin\beta_b \qquad (5\text{-}4a)$$

which is converted to

$$x = (1+\Upsilon)\cos\beta_t + (1-\Upsilon)\cos\left(\frac{1-\Lambda}{1+\Lambda}\beta_t\right) \qquad (5\text{-}4b)$$

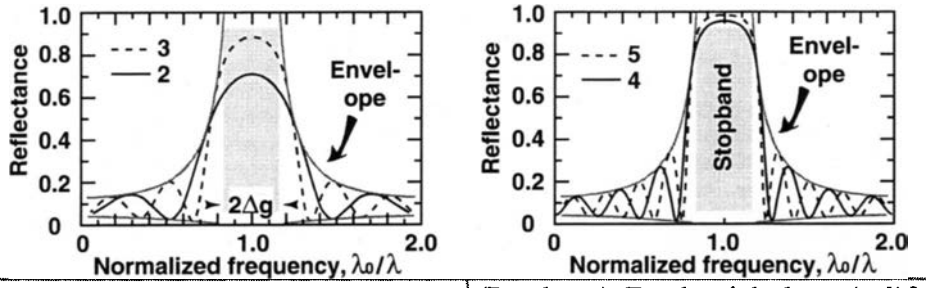


Fig. 5-15 (Panel one) For the four-layer (*solid*) and six-layer (*dashed*) stacks *air* $(\mathbf{H}\,\mathbf{L})^2$ *substrate* and *air* $(\mathbf{H}\,\mathbf{L})^3$ *substrate*, (Panel two) For the eight-layer (*solid*) and ten-layer (*dashed*) stacks *air* $(\mathbf{H}\,\mathbf{L})^4$ *substrate* and *air* $(\mathbf{H}\,\mathbf{L})^5$ *substrate*. $R(\lambda_0/\lambda)$ of the above designs where the refractive indices of **L**, **H** and *substrate* are 1.38, 2.31 and 1.52, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$. The shaded area is the stop band (total width $2\Delta g$). Shaded curves are the reflectance envelopes.

where Υ is related to the index ratio [see Eq. (2-47c)] and Λ is the thickness ratio

$$\Lambda \equiv \beta_a/\beta_b \qquad (5\text{-}4c)$$

which is equivalent to Eq. (2-47b). According to Eq. (5-3), a necessary (but not sufficient) condition for the occurrence of a stop band (of order m) is that the total phase thickness β_t satisfy

$$\beta_t \equiv \beta_a + \beta_b = m\,\pi\ . \qquad (5\text{-}4d)$$

Example: Consider the case of a quarterwave stack, where $\Lambda = 1$. With the definition of Υ in Eq. (2-79a), the extrema of x occur when Eq. (5-4d) is satisfied

$$|x| = |(1+\Upsilon) + (-1)^m (1-\Upsilon)| \qquad (5\text{-}5a)$$

and equals either 2 or 2Υ. When $|x| = 2$ at the *extrema*, the stop band is absent, since it has zero width. Thus the *even* numbered harmonic stop bands are absent. Because Υ always exceeds unity, all *odd* numbered stop bands exist.

5.2.4.3 Computation of the stop band spectral width

The β_t values at which the edges of the stop band occur are found by setting $x = \pm 2$ in Eq. (5-4b) and solving for β_t. The resulting equation is transcendental in most cases. Inspection of Eq. (5-4b) shows that the stop band width depends upon:

(1) The *ratio* of the phase thicknesses; and

(2) the *ratio* of the refractive indices.

Several examples illustrate the foregoing points.

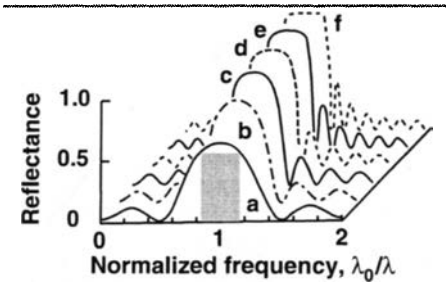


Fig. 5-16 Isometric plot of $R(\lambda_0/\lambda)$ for the quarterwave stacks

air **H** $(\mathbf{L}\,\mathbf{H})^q$ *glass*

where the indices of *glass*, **L** and **H** are 1.52, 1.38 and 2.30, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$. Curves *a, b, c, d, e* and *f* correspond to q = 1, 2, 3, 4, 5 and 6, respectively. The shaded area is the stop band.

5.2.4.3.1 Quarterwave stack

The phase thicknesses of the quarterwave stack are equal; $\beta_a = \beta_b$ and the stop bands are centered at β_t = 180°, 540°, 900° and so on. The reflectance in Fig. 5-15 is plotted vs normalized frequency g

$$g \equiv \lambda_0/\lambda \qquad (5\text{-}5b)$$

where $\lambda_0/2$ is the total optical thickness of the basic period at the lowest order stop band. There are two methods of defining the stop band width.

The *semiwidth* is the frequency difference Δg from the center of the stop band at $g = 1.0$ to its edge

$$\Delta g = \frac{2}{\pi} \arcsin\left(\frac{|1 - n_a/n_b|}{1 + n_a/n_b}\right) \qquad (5\text{-}5c)$$

where || denotes the absolute value; the principal value of the arcsin is used.

The other definition is to establish the *total width* at the ratio of g_2 to g_1, where g_1 and g_2 are the frequencies at the upper and lower edges of the stop band.

The QW stack is unique, in that it is symmetrical about $g = 1.0$. The total bandwidth is $2\Delta g$. As a numerical example, the quarterwave stack shown in Fig. 5-15 has $n_a = 1.38$ and $n_b = 2.31$. Substitution of n_a and n_b into Eq. (5-5c) provides for the semiwidth $\Delta g = 0.16$. The first-order stop band extends from a lower frequency of $g_1 = 1 - 0.16 = 0.84$ to an upper limit of $g_2 = 1 + 0.16 = 1.16$. The ratio total bandwidth is $g_2/g_1 = 1.16/0.84 = 1.39$ for this index ratio.

Figure 5-15 illustrates that the maxima and minima of the $R(\lambda)$ curve oscillate between the upper and lower limits—an envelope—except in the stop band, where R rises monotonically as more periods are added. Figure 5-16 plots the $R(\lambda_0/\lambda)$ of

air **H L H** *glass* and *air* **H L H L H L H L H L H** *glass* .

The reflectances of stacks with an intermediate number of layers are also depicted. The stop band spectral region is shaded. The rise of the reflectance in that region is quite evident as the number of periods progresses from 1 (three layers) to 6 (thirteen layers).

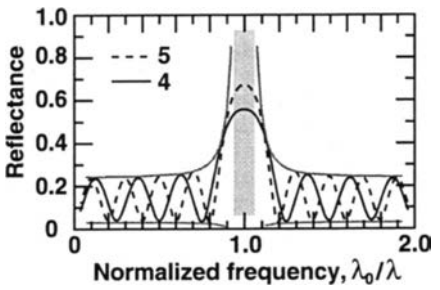


Fig. 5-17 $R(\lambda_0/\lambda)$ of the quarterwave stacks (*solid*)

air $(\mathbf{H}\,\mathbf{L})^4$ *substrate*

and (*dash*)

air $(\mathbf{H}\,\mathbf{L})^5$ *substrate*

where the refractive indices of **L**, **H** and *substrate* are 1.90, 2.30 and 1.52, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$. The shaded area is the stop band spectral region. The shaded curves are the reflectance envelopes.

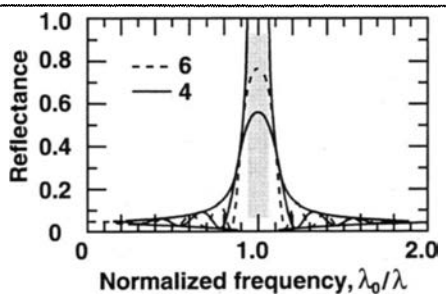


Fig. 5-18 $R(\lambda_0/\lambda)$ of the quarterwave stacks (*solid*)

air $(\mathbf{H}\,\mathbf{L})^4$ *substrate*

and (*dash*)

air $(\mathbf{H}\,\mathbf{L})^6$ *substrate*

where the refractive indices of **L**, **H** and *substrate* are 1.38, 1.67 and 1.52, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$. The shaded area is the stop band spectral region. The shaded curves are the reflectance envelopes.

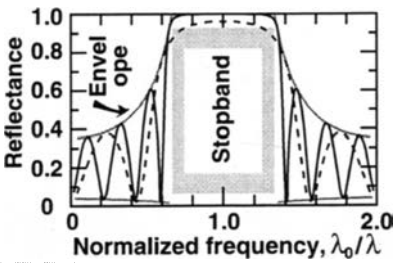


Fig. 5-19 $R(\lambda_0/\lambda)$ for the quarterwave stack (*dashed*)

air $(\mathbf{H}\,\mathbf{L})^2$ *substrate*

and (*solid*)

air $(\mathbf{H}\,\mathbf{L})^4$ *substrate*

where the refractive indices of **L**, **H** and *substrate* are 1.40, 4.2 and 1.50, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$. The shaded area is the stop band spectral region. The shaded curves are the reflectance envelopes.

The next example is in Fig. 5-17, which depicts the first-order stop band of quarterwave stacks in which the indices are relatively close together, with n_a = 1.90 and n_b = 2.30. The stop band is narrower than that of the stack in Fig. 5-15. Here, g_2/g_1 = 1.13, in contrast to the 1.39 width obtained previously. Although the indices of n_a = 1.38 and n_b = 1.67 are different for the quarterwave stacks shown in Fig. 5-18, the index *ratio* is the same and, hence, the total stop band width is quite close to that of Fig. 5-17.

§5.1.4.1 and §5.3.3.1 discuss the narrowing of the spectral bandwidth of the reflectance peak. This is achieved by using refractive indices that are relatively close together.

Figure 5-19 shows the reflectance envelopes and stop band of a stack in which the layers are germanium (n_H = 4.2) and a rare earth fluoride (n_L = 1.40). Only a few layers are needed to achieve a substantial reflectance at β_t = 180° because of the large index difference.

The ratios of the stop band edge frequencies g_2/g_1 are plotted in Fig. 5-20 for a quarterwave stack. The abscissa is the index ratio n_b/n_a. As an example, point *a* on the curve corresponds to the ratio 2.31/1.38 = 1.67 shown in Fig. 5-15. The "germanium plus rare earth fluoride" stack mentioned in the previous paragraph is at point *b* on the curve and covers a 2:1 bandwidth.

5.2.4.3.2 Stacks of unequal layer thickness

§5.1.3.3 lists reasons that stacks with unequal layer thickness ratios are used to achieve various design goals. §5.1.4.2 illustrates that stacks with unequal layer thicknesses manifest narrower bandwidth. The vector diagrams in §5.1.3.1 and §5.1.3.2 show that the narrowing is at the expense of a reduced peak reflectance, however. As an illustration, consider the coating

air $(\mathbf{H'}\ \mathbf{L''})^q$ *substrate* ,

where the indices of the **H'**, **L"** and *substrate* are the same as in §5.1.2.1 —namely zinc sulfide and rare earth fluoride layers deposited on a zinc selenide substrate. The question is: How much narrowing of the stop band occurs as the thickness ratio is varied? Figure 5-21 shows the narrowing of the stop band that occurs as the thickness ratio progresses from 1:1 (for a quarterwave stack) to 2:1 and finally 9:1. The total bandwidth of the stop band of a quarterwave stack [using Eq. (5-5c)] is slightly less than 30%. Figure 5-21 shows that bandwidth may be reduced by a factor of three—to less than 10% —by using a thickness ratio 9:1.

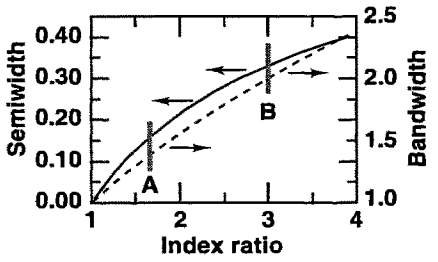


Fig. 5-20 Normalized semiwidth Δg and bandwidth

$$(1 + \Delta g)\ (1 - \Delta g)^{-1}$$

of the stop band of a quarterwave stack vs the index ratio of the layers n_a/n_b. The vertical line The vertical line *A* corresponds to an index ratio of 2.31/1.38 and line *B* to 4.2/1.40.

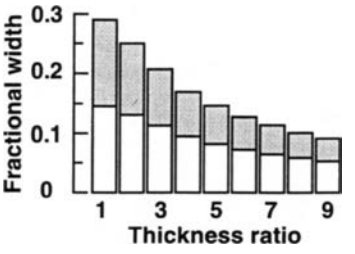


Fig. 5-21 For a stack centered at λ_0, total spectral width of the stop band (expressed as a fraction of λ_0) of the reflectors

$$(\mathbf{H'}\ \mathbf{L''})^q$$

where the indices of the **H'** and **L"** are 2.22 and 1.40, respectively. The abscissa is the ratio of the optical thickness of the **L"** to the **H'** layer.

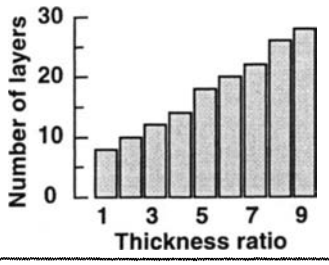


Fig. 5-22 For the stacks captioned in Fig. 5-21, the minimum number of layers needed to achieve a peak reflectance in excess of 95% at the first-order stop band.

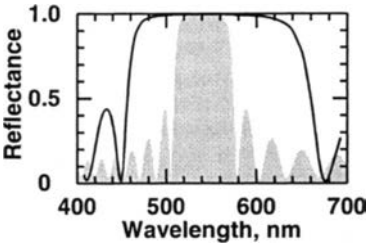


Fig. 5-23 $R(\lambda)$ of the quarterwave stack (*solid*)

air $(\mathbf{H}\,\mathbf{L})^6\ \mathbf{H}$ *glass*

and three-quarterwave stack (*shaded*)

air $(3\mathbf{H}\,3\mathbf{L})^6\ 3\mathbf{H}$ *glass*

where the indices of **L**, **H** and glass are 1.45, 2.32 and 1.52, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$, where λ_0 is 540 nm.

The next question is: How many layers are required to achieve a peak reflectance of a specified level? Figure 5-11 shows that doubling the number of the layers in the stack with the 6:1 thickness ratio failed to produce the same peak reflectance as the quarterwave stack. Figure 5-22 graphs the number needed to achieve a peak reflectance in excess of 96%. Whereas only eight quarterwave layers are needed, the 9:1 stack requires nearly 30 layers to achieve the 96% reflectance.

The next section shows that a higher order stop band produces a narrower bandwidth.

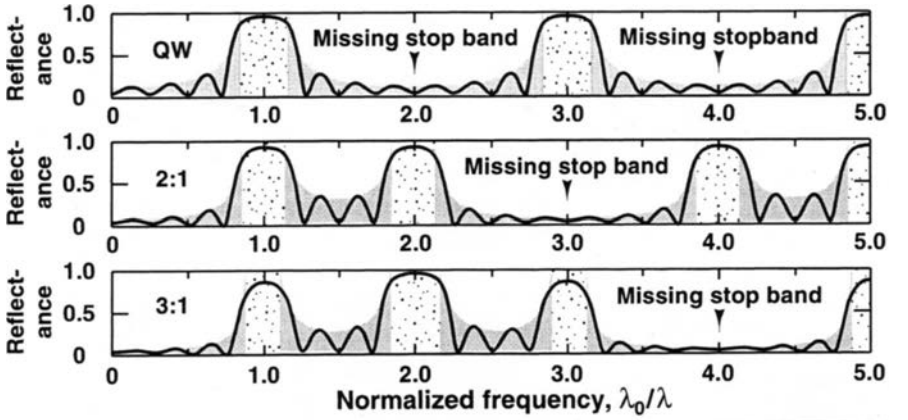


Fig. 5-24 Reflectance (*solid curves*), stop bands (*dotted area*) and reflectance envelope (*shaded region*) vs λ_0/λ of the reflectors

(*upper*) *air* $(\mathbf{H}\,\mathbf{L})^4$ *glass* ,

(*middle*) *air* $(0.667\mathbf{H}\,1.33\mathbf{L})^4$ *glass* and

(*bottom*) *air* $(0.5\mathbf{H}\,1.5\mathbf{L})^4$ *glass*

where the indices of **H**, **L** and *glass* are 2.31, 1.38 and 1.52, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$. Arrows designate missing stop bands.

5.2.5 Higher order stop bands

§5.2.3.4 establishes that the *order number* of a stop band is the number of halfwaves that "fit" into its basic period. The order number of a stop band is of concern to the designer for at least two reasons: (1) The *fractional* bandwidth of a higher order stop band is usually smaller than one of lower order; (2) it is often important that a higher order (or lower order) stop band be either absent or present.

Several examples illustrate the foregoing points. Figure 5-23 shows the reflectance of stacks that reflect in the green spectral region at $\lambda_0 = 540$ nm. The same materials are used in both quarterwave stacks. The stack with the narrower bandwidth, however, has the *third-order* stop band centered at λ_0. Because the absolute width of the stop band is the same for all orders, its *fractional* bandwidth is one-third smaller.

As another example, suppose a quarterwave stack heat reflector has its stop band centered at 1500 nm in the IR. Such a coating could not be used in an illuminator because its third-order stop band is centered at (1500 nm)/3 = 500 nm, thus reflecting strongly in the visible spectral region. §5.3.4.2 gives another example of such harmonics.

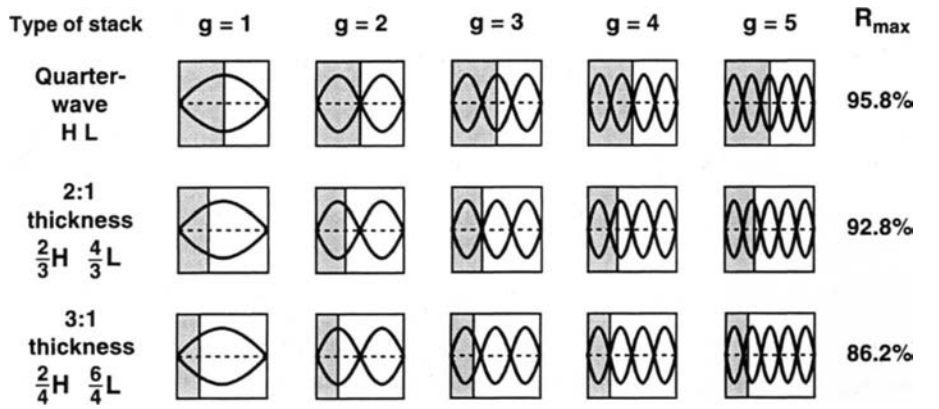


Fig. 5-25 At the stop bands of order numbers 1 through 5, the wavelengths at those stop bands is compared with the optical thickness of the basic periods appearing in the caption to Fig. 5-24. R_{max} is the maximum reflectance of the first-order stop band for the eight-layer stacks appearing in the caption to Fig. 5-24.

5.2.5.1 Quarterwave stack

The upper panel of Figure 5-24 depicts the reflectance vs normalized frequency g [see Eq. (5-5b)] of a quarterwave stack. At $g = 1.0$, the optical thickness of each layer is $\lambda_0/4$ and the first-order stop band occurs. The first-order stop band occurs when there is a resonance of the halfwave in the basic period, as shown in Fig. 5-25.

At double that frequency (i.e., half the wavelength of the fundamental), each layer is $\lambda_0/2$ in optical thickness. Hence, both layers are absentee (see §2.6.3.1) and the reflectance equals that of the uncoated substrate. The trace of the characteristic matrix equals ± 2 and hence no stopband exists.

The third-order stop band occurs at $g = 3.0$. Its absolute width (in terms of Δg) is the same as that of the first-order stop band. Its fractional width $\Delta g/g$, however, is reduced by one-third, since it is third order. Figure 5-25 shows that at $g = 4.0$ a full wave

resonates in each layer, which means that they are absentee. The fourth-order stop band is missing. The fifth-order stop band occurs at $g = 5.0$.

As noted in §5.2.4.2, all odd numbered stop bands are present and the even are missing.

5.2.5.2 Two-to-one stack

This type of coating is introduced in §5.1.3.1. The optical thickness of one of the components in the basic period is double that of the other. A first-order stop band exists at $g = 1.0$, where the total optical thickness of the basic period is $\lambda_0/2$, as shown in Fig. 5-25. The second-order stop band occurs at $g = 2.0$ when the total optical thickness of the basic period is a full wave at λ_0. It might be expected that a third harmonic stop band would occur at $g = 3.0$. The optical thickness of the 0.67H layer is a halfwave and the optical thickness of the 1.33L layer is a fullwave—hence, both are *absentee* (see §2.6.3.1). The reflectance is that of the uncoated substrate. At $g = 4.0$ and $g = 5.0$, however, stop bands occur because the optical thickness of the basic period equals four halfwaves and five halfwaves. In conclusion, stop bands with order numbers 1, 2, 4, 5, 7, 8 and so on are present. Stop bands with orders 3, 6 and 9 are missing.

The coating described above has a low index film that is twice the optical thickness of the high index. Suppose the converse is true. Inspection of Eq. (5-4a) shows that the stop band width is the same.

5.2.5.3 Three-to-one stack

The design of this stack is

$$air\ (0.5\mathrm{H}\ 1.5\mathrm{L})^q\ glass\ .$$

The thicknesses are chosen so that the total thickness of the basic period is $\lambda_0/2$—i.e., two QWOTs at λ_0. The bottom panel of Fig. 5-24 shows that a first-order stop band occurs at $g = 1.0$. Stop bands also occur at $g = 2.0$ and $g = 3.0$, where the optical thickness of the basic period is two halfwaves and three halfwaves of λ_0, respectively. When $g = 4.0$, however, the optical thickness of the 1.5L layer is three halfwaves and the optical thickness of the 0.5H layer is a single halfwave. Hence both are absentee (see §2.6.3.1) and the reflectance at $g = 4.0$ is identical to that of an uncoated substrate. At $g = 5.0$, the optical thickness of the basic period is five halfwaves and a stop band occurs. Thus, every fourth stop band is missing. In other words, stop bands occur when $m = 1, 2, 3, 5, 6, 7, 9, 10, 11, 13$ and so on. Stop bands with orders 4, 8 and 12 are missing.

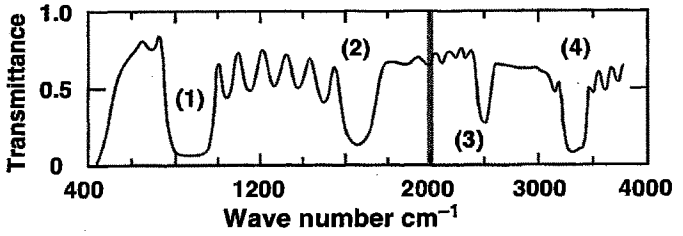


Fig. 5-26 Measured $T(\sigma)$ of an eleven alternating layers of lead fluoride and zinc selenide deposited on a cadmium telluride substrate. The numbers identify the order numbers of the stop bands. There is a scale change of wave number at 2000 cm^{-1}.

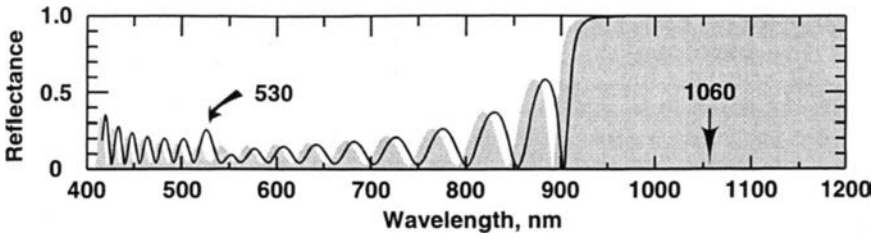


Fig. 5-27 $R(\lambda)$ of the YAG laser reflector

air $(\mathbf{H}\,\mathbf{L})^9\,\mathbf{H}$ *glass*

where the indices of **L** and *glass* are 1.45 and 1.52, respectively. The index of **H** is dispersive, with $A = 5.11$ and $B = 0.17$ in Eq. (2-2c). (*Solid*) Optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 1060 nm. (*Shaded*) optical thickness of **L** and **H** is $\lambda_0/2$ at λ_0 of 530 nm. Arrows designate 530 nm and 1060 nm.

5.2.5.4 Arbitrary thickness ratio

The higher order stop bands are absent *only* when the optical thickness ratio can be expressed as the quotient of integers, such as 1:2, 4:3 or 2:5. *All* higher order stop bands are present in every other type of stack. As an example, Fig. 5-26 shows the measured $T(\sigma)$ of an eleven-layer stack. The first-order stop band occurs at 850 cm^{-1}. The ratio of the optical thicknesses is such that the second, third and fourth harmonic stop bands are all present.

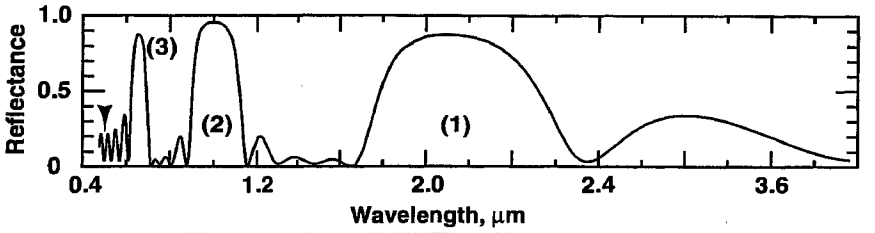
5.2.5.5 Effects of dispersion

A spurious stop band can occur due to the dispersion of the refractive index of a coating material. An example is a multilayer reflector for a YAG laser (operating at 1060 nm) that contains a nonlinear crystal. It must transmit the doubled frequency—at 530 nm. A quarterwave stack of titania and silica is used

air $(\mathbf{H}\,\mathbf{L})^9\,\mathbf{H}$ *glass* .

In the first instance, the optical thickness of each layer is *defined* as a QWOT at 1060 nm, where the refractive index of the titania is 2.29. At 530 nm the index of the titania layer has crept upwards to 2.39. The optical thickness [see Eq. (2-37a)] is increased by this change in index. The result is that the optical thickness of a titania layer is not a halfwave at 530 nm and, hence, it is not absentee (see §2.6.3.1). A stop band of rather narrow spectral width is created by the dispersion, as manifested by a narrow reflectance spike in the vicinity of 530 nm in Fig. 5-27.

The remedy is to define the layer thicknesses as *halfwaves* at 530 nm. The resulting reflectance is the shaded curve in Fig. 5-27. As compared to the stack in the foregoing paragraph, its reflectance at 1060 is slightly smaller. It is possible that computer optimization (see §10.2) could be used to increase the transmittance at 530 nm. But even so, it is desirable to start with an improved design, such as the "halfwaves at 530" in preference to the "QWOTs at 1060."



5-28 $R(\lambda)$ of the three-to-one stack

air $(1.5\mathbf{H}\,0.5\mathbf{L})^4$ *glass*

where the indices of **L**, **H** and *glass* are 1.38, 2.30 and 1.51, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 2.0 μm. The integers in parentheses are the order numbers of the stop bands. The arrow designates the missing fourth-order stop band.

5.2.5.6 Harmonic reflection peaks on a wavelength scale

Most of the reflectance curves in §5.2.5 are plotted on a frequency scale, thus avoiding the distortion produced by a wavelength scale (see §5.1.3.3). Figure 5-28 depicts $R(\lambda)$ of a three-to-one stack, the design of which is similar to that of the bottom panel Fig. 5-24. The first-order stop band — and its attendant reflection peak — is centered at $\lambda_0 = 2.0$ μm. A second-order peak occurs at $\lambda = 1.0\ \mu m = \lambda_0/2$. A third-order peak appears at $\lambda_0/3 = 0.667$ μm — one-third of the fundamental. A fourth-order peak does *not* occur at $\lambda_0/4 = 0.5$ μm because both the $0.5\mathbf{H}$ and $1.5\mathbf{L}$ layers are absentee. A fifth-order peak appears at $\lambda_0/5 = 0.400$ μm and so on.

If frequency were used for the abscissa of Fig. 5-24, the width of all of the stop bands of the same *relative* order would be the same. By way of illustration, the stop band widths of orders 1, 3, 5, 7 and so on are identical. The widths of stop bands at $m = 2$, 6, 10 and so on are also equal. This symmetry is lost on the wavelength scale.

5.2.6 Splitting of the layers in a basic period

Given a basic period such as **HL**, what happens when the **L** layer is split into two parts? The lowest panel of Fig. 5-14 shows a basic period

$$0.35\,\mathbf{L}\,\mathbf{H}\,0.65\,\mathbf{L}\ .$$

What is its stop band width? A calculation of its trace [see Eq. (2-41e)] shows that its stop band width is identical to that of an **LH** or **HL** pair. That is, given a coating

$$\mathbf{L'HL''}$$

where the optical thickness of **L'** is $f\lambda_0/4$ and the optical thickness of **L"** is $(1-f)\lambda_0/4$, its stop band width is independent of f, provided $0 \le f \le 1$.

A consequence is that an equivalent layer (see §2.6.7) may be used for the basic period, as shown in Table 5-1. That is, the stop band widths of the stacks $\frac{1}{2}\mathbf{H}\,\mathbf{L}\,\frac{1}{2}\mathbf{H}$ and $\frac{1}{2}\mathbf{L}\,\mathbf{H}\,\frac{1}{2}\mathbf{L}$ are identical to those of **HL** or **L H**.

Table 5-1 Equivalent layers consisting of two components

Type of stack	Λ	Design in terms of OT	Design in terms of QWOT
One-to-one or QW	1.0	$\frac{H}{8}\ \frac{L}{4}\ \frac{H}{8}$	$\frac{H}{2}\ L\ \frac{H}{2}$
One-to-two	0.5	$\frac{H}{12}\ \frac{4L}{12}\ \frac{H}{12}$	$\frac{H}{3}\ \frac{4L}{3}\ \frac{H}{3}$
Two-to-one	2.0	$\frac{H}{6}\ \frac{L}{6}\ \frac{H}{6}$	$\frac{2H}{3}\ \frac{2L}{3}\ \frac{2H}{3}$
One-to-three	0.333	$\frac{H}{16}\ \frac{6L}{16}\ \frac{H}{16}$	$\frac{H}{4}\ \frac{6L}{4}\ \frac{H}{4}$
Three-to-one	3.0	$\frac{3H}{16}\ \frac{2L}{16}\ \frac{3H}{16}$	$\frac{3H}{4}\ \frac{2L}{4}\ \frac{3H}{4}$
Two-to-three	0.667	$\frac{H}{10}\ \frac{3L}{10}\ \frac{H}{10}$	$\frac{2H}{5}\ \frac{6L}{5}\ \frac{2H}{5}$
Three-to-two	1.5	$\frac{3H}{20}\ \frac{4L}{20}\ \frac{3H}{20}$	$\frac{3H}{5}\ \frac{4L}{5}\ \frac{3H}{5}$

Equivalent layers that manifest a first-order stop band at λ_0. The optical thicknesses OT are multiples of λ_0. **H** and **L** represent QWOT layers of optical thickness $\lambda_0/4$. Λ appears in Eq. (2-47b).

5.3 "Single-stack" coatings

The "single stack" above distinguishes the coatings in this section from the "two-component" stacks discussed in §5.5.1. §5.3 focuses on multilayers containing one periodic structure, such as the YAG laser reflector in §5.2.5.5. In many instances, a few layers are added to either end of the stack to modify its performance in various ways. Such coatings are used as reflectors and rejection filters.

5.3.1 Decision rules that produce an optimal quarterwave reflector

Our goal is to produce, at a reference wavelength λ_0, a reflector with *optimal efficiency* — a term that is described below. The stack should be constructed as follows:

(1) Each layer of the stack is nonabsorbing.
(2) Each layer has an optical thickness of $\lambda_0/4$ — that is, a quarterwave.
(3) All refractive indices must alternate between a high index and a low index. The term "indices" in the previous sentence includes the refractive index of the incident medium, refractive index of the substrate as well as the refractive indices of the layers.
(4) The *ratio* of the refractive indices at each interface should be as large as possible.

The four statements above are regarded as *decision rules.*

5.3.1.1 Meaning of *optimal efficiency*

Optimal efficiency means simply that the highest reflectance is produced with the fewest layers. In quantitative terms, at a wavelength λ_0 the standing wave ratio V, or

alternatively the reflectance, is the largest for a given number of layers. An optimal coating is produced at the lowest cost.

5.3.1.2 Meaning of *high as possible* and *low as possible*

Consider rule (3) in §5.3.1. The *highest possible* index should be used. Although refractive indices as large as 2.25 could be used in WDM bandpasses, we settle for a high index of 2.065 (tantala) because this coating material is stable and its refractive index does not shift when exposed to humidity.

Although a refractive index as low as 1.38 could be considered as a material to be used in WDM devices, it is not used because of its lack of stability, as mentioned in the paragraph above. We have no choice but to use 1.47 (silica) as a low index material.

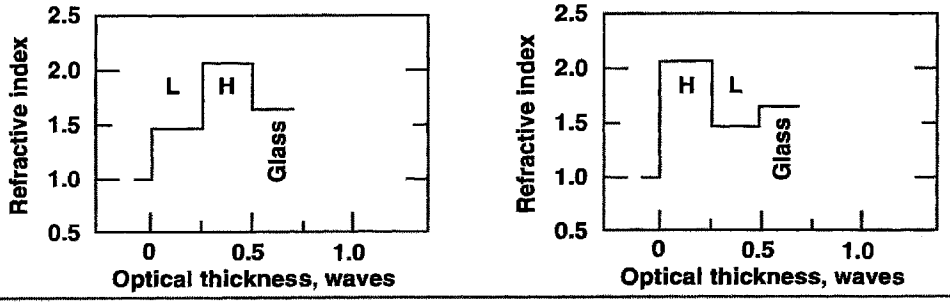


Fig. 5-29 Versus optical thickness, refractive index of the two-layer stacks *air* **L H** *glass* (*left panel*) and *air* **H L** *glass* (*right panel*).

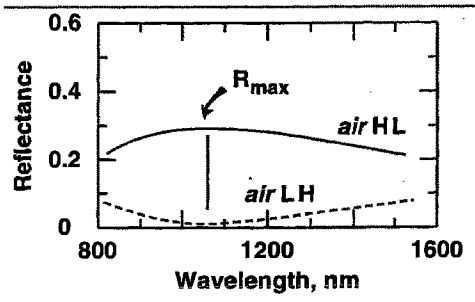


Fig. 5-30 Reflectance of the designs *air* **L H** *glass* (*dashed*) and air **H L** *glass* (*solid*) where the refractive indices of *glass*, **L** and **H** are 1.65, 1.47 and 2.065, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 1060 nm.

5.3.1.3 Two-layer coating — first example

These coatings are illustrative and are not necessarily intended for use in a WDM bandpass. The high index layer (its symbol is **H**) is tantala (index 2.065). Silica (index 1.47) is the low index material (its symbol is **L**). The glass substrate index is 1.65. Following decision rules (1) and (2) above, each layer is nonabsorbing and has an optical thickness of $\lambda_0/4$.

A first design is

air **L H** *glass* .

The left panel of Fig. 5-29 graphs the refractive index versus optical thickness. Figure 5-30 (*dashed curve*) and Table 5-2 show that the reflectance of this coating (at the "tuned" wavelength) is below 1%, which is less than the reflectance of the uncoated glass substrate.

Table 5-2 Standing wave ratios of quarterwave stacks

Design	§	n_H	n_L	n_F	n_s	V	R
air **L H** *glass*	5.3.1.3	2.065	1.47	-	1.65	1.2	0.008
air **H L** *glass*	5.3.1.4	2.065	1.47	-	1.65	3.25	0.281
air **L H L H** *glass*	5.3.1.5	2.065	1.47	-	1.65	2.3	0.164
air **H L H L** *glass*	5.3.1.6	2.065	1.47	-	1.65	6.4	0.534
air $(\mathbf{H\,L})^{16}$ **F** *glass*	5.3.1.7	2.065	1.47	1.86	1.50	311	0.9872

For the coating designs shown in the Design column, standing wave ratio V and radiant reflectance R of the stack at a wavelength λ_0. **L**, **H** and **F** represent layers of optical thickness $\lambda_0/4$. The refractive indices of **H**, **L**, **F** and *glass* appear in the n_H, n_L, n_F and n_s columns, respectively. The § column refers to the section where the coating is discussed.

In other words, this coating functions not as a reflector, but as an antireflection coating. The low reflectance may be attributed to two factors:

- Decision rule (3) in §5.3.1 is violated. The sequence of indices is

  air (*low*)
  L layer (higher)
  H layer (highest).

This "graded index" (from low, higher, highest) violates the rule that the indices should alternate from high to low.

- The left panel of Fig. 5-29 shows that rule (4) in §5.3.1 is violated because the optimum index ratio is *not* achieved at each interface. In fact, the coating below reverses the order of the layers and produces a higher reflectance at λ_0.

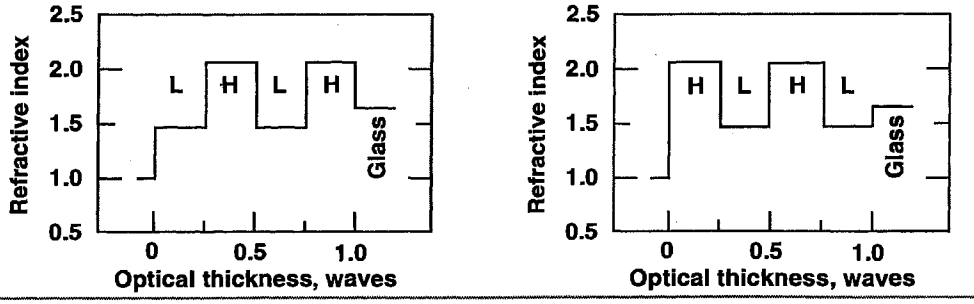


Fig. 5-31 Versus optical thickness, refractive index of the four-layer stacks *air* **L H L H** *glass* (*left panel*) and *air* **H L H L** *glass* (*right panel*).

5.3.1.4 Two-layer coating — second example

The order of the **H** and **L** layers in §5.3.1.2 is reversed, with the result

air **H L** *glass* .

The graph of the coating indices in the right panel of Fig. 5-29 shows that the ratio of the refractive indices at each interface is larger. This, in turn, produces a higher reflectance

of 28% [see Table 5-2 and Fig. 5-30 (*solid curve*)] because there is compliance with decision rule (4) cited in §5.3.1.

5.3.1.5 Four-layer coating — third example

The same refractive indices and symbols for the layers are used as in §5.3.1.3. Following decision rules (1) and (2) above, each layer is nonabsorbing and has an optical thickness of $\lambda_0/4$. A four-layer design is

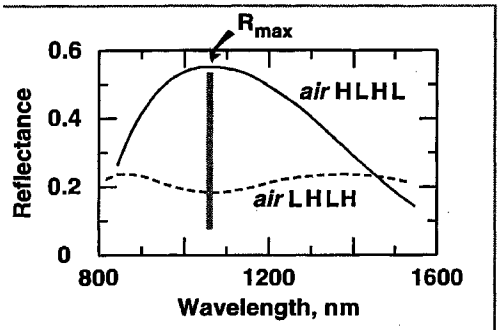


Fig. 5-32 Reflectance of the designs
air **L H L H** *glass* (*dashed*) and
air **H L H L** *glass* (*solid*)
where the refractive indices of *glass*, **L** and **H** are 1.65, 1.47 and 2.065, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 1060 nm.

air **L H L H** *glass* .

The left panel of Fig. 5-31 graphs the refractive index versus optical thickness. The same arguments apply as in §5.3.1, namely, that the refractive indices do not alternate between a high and a low. That is, the sequence "*air* **L H**" forms a graded index, which reduces the reflectance. Figure 5-31 shows that rule (4) (§5.3.1) is violated because the optimum index ratio is *not* provided at each interface. This coating of four layers produces a reflectance of only 16% (see Table 5-2 and the *dashed curve* in Fig. 5-32) in contrast to the two-layer coating in §5.3.1.4 that reflects 28%.

5.3.1.6 Four-layer coating — fourth example

The order of the **H** and **L** layers in §5.3.1.5 is reversed, with the result

air **H L H L** *glass* .

The graph of the coating indices in the right panel of Fig. 5-31 shows that the ratios of the refractive indices at each interface is larger than those in §5.3.1.5. This, in turn, produces a higher reflectance of 53% (see Fig. 5-32 and Table 5-2) because the teachings in decision rule (4) above (§5.3.1) are put into practice.

5.3.2 Reflectance of a quarterwave stack — derived from admittance

The question is, how many nonabsorbing layers (of indices n_L and n_H) are needed to produce a given reflectance? It is assumed that an **H** layer is contiguous to the incident medium. The extremum transmittance T_{ex} occurs at the quarterwave point, at $\lambda = \lambda_0$, where the phase thickness of each layer is 90°. The first step is to find the standing wave ratio V_{ex} using Eqs. (2-20b) and (2-19g)

$$V_{ex} \approx 4/T_{ex} . \tag{5-6a}$$

From Eq. (2-48b), the V of a stack with q periods and an even number of layers is

$$V = \frac{n_H^{2q} n_s}{n_L^{2q} n_0} \tag{5-6b}$$

and for an odd layer count

$$V = \frac{n_H^{2q+2}}{n_L^{2q} n_s n_0} . \tag{5-6c}$$

The foregoing equation is derived from admittance. One approach is to find a value of q such that V in Eqs. (5-6b) or (5-6c) exceeds V_{ex}.

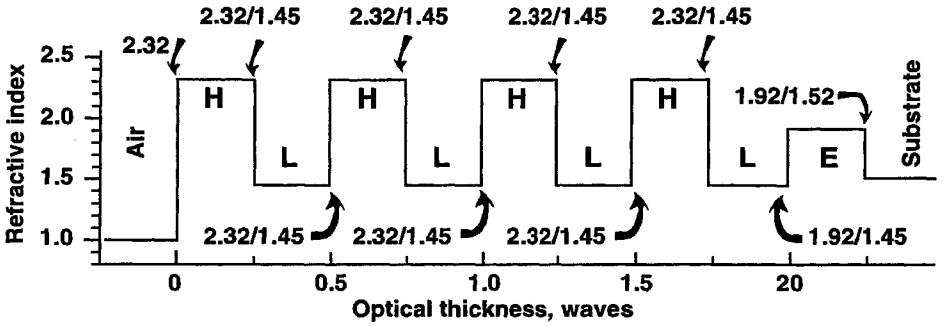


Fig. 5-33 Refractive index profile (vs optical thickness) of *air* $(\mathbf{HL})^4$ **E** *glass* where the indices of **H**, **E**, **L** and *glass* are 2.32, 1.924, 1.45 and 1.52, respectively. Each layer has an optical thickness of $\lambda_0/4$. The arrowed number and arrowed fractions designate the standing wave ratio at each of the ten interfaces.

5.3.2.1 Reflectance of a quarterwave stack — derived from interface standing wave ratios

The concept of "interface standing wave ratios" is derived from the theory of microwave transmission lines {see Eq. (6.03-4) in Matthaei [64]}.

> For a nonabsorbing stack consisting of ℓ layers of optical thickness $\lambda_0/4$ in which all refractive indices alternate between high and low, at λ_0 the standing wave ratio V of the stack is the product of the standing wave ratio V_i at each of the $\ell+1$ interfaces.

Here, *interface* refers to the boundary between layers, boundary between a layer and the incident medium or boundary between a layer and the emergent medium. We establish the standing wave ratio V_0 at the outermost layer of the stack (adjacent to the incident medium of refractive index n_0) as

$$V_0 \equiv [\, n_0 / n_1 \,]^{\pm 1} \,. \tag{5-6d}$$

At the layer-substrate interface the standing wave ratio is

$$V_\ell \equiv [\, n_\ell / n_s \,]^{\pm 1} \tag{5-6e}$$

and between the layers the standing wave ratio is

$$V_i \equiv [\, n_i / n_{i+1} \,]^{\pm 1} \,. \tag{5-6f}$$

The ± 1 exponent in Eqs. (5-6d), (5-6e) and (5-6f) is chosen so that the number within the brackets is not less than 1.00. The standing wave ratio V of the stack is the product

$$V = \prod_{i=0}^{\ell} V_i \,. \tag{5-6g}$$

5.3.2.1.1 Two-layer stack — first example

An example is the two-layer quarter-wave stack cited in §5.3.1.4

air **H L** *glass* .

It contains three interfaces. The first is between air (refractive index 1.0003) and the high index layer. The interface standing wave ratio V_0 is

$$V_0 = 2.065 \,/\, 1.0003 = 2.064 \,. \tag{5-6h}$$

At the interface between the **H** layer and the **L** layer, the standing wave ratio V_1 is

$$V_1 = 2.065 / 1.47 = 1.405 \ . \qquad (5\text{-}6i)$$

Finally, at the interface between the **L** layer and the substrate of refractive index 1.65, the standing wave ratio V_2 is

$$V_2 = 1.65 / 1.47 = 1.12 \ . \qquad (5\text{-}6j)$$

From Eq. (5-6g) the standing wave ratio V of the stack is the product

$$V = 2.064 \times 1.405 \times 1.12 = 3.25 \ . \qquad (5\text{-}6k)$$

From Eq. (2-8) the reflectance R is

$$R = (1 - V)^2 / (1 + V)^2 = (1 - 3.25)^2 / (1 + 3.25)^2 = 0.28 = 28\% \qquad (5\text{-}6m)$$

5.3.2.1.2 Two-layer stack — improper use of the formula

The final example is the quarterwave stack cited in §5.3.1.3. The outer layer is low index, as in the design

air **L H** *glass.*

It contains three interfaces. The standing wave ratio V_0 at the outer interface is

$$V_0 = 1.47 / 1.0003 = 1.47 \ . \qquad (5\text{-}6n)$$

At the interface between the **H** layer and the **L** layer, the standing wave ratio V_1 is

$$V_1 = 2.065 / 1.47 = 1.405 \ . \qquad (5\text{-}6p)$$

Finally, at the interface between the **H** layer and the substrate of refractive index 1.65, the standing wave ratio V_2 is

$$V_2 = 2.065 / 1.65 = 1.25 \ . \qquad (5\text{-}6q)$$

From Eq. (5-6g) the standing wave ratio V of the stack is the product

$$V = 1.47 \times 1.405 \times 1.25 = 2.58 \ , \qquad (5\text{-}6r)$$

which is the wrong answer. The actual V is 1.20. Why is this the wrong answer? Equation (5-6g) obtains if, and only if, the refractive indices *alternate* between a high and a low. As pointed out in §5.3.1.3, when the sequence of layers is "*air* **L H**", the indices do *not* alternate between high and low.

The correct way to find the standing wave ratio is to use equations (appearing in §2.7.3.2) that are derived from admittance theory.

A further example is the quarterwave stack of the design

air $(\mathbf{H}\cdot\mathbf{L})^4$ **E** *glass*

that is the focus of §5.3.2.2. The refractive indices of *glass*, **H**, **L** and **E** are 1.52, 2.32, 1.45 and 1.924, respectively. Figure 5-33 graphs the refractive index vs optical thickness for this nine-layer coating. The "step", i.e., the discontinuity in refractive index at each interface produces a standing wave, the ratio of which is computed. Since there are nine layers in the stack, there are ten such discontinuities.

Figure 5-33 shows that there are seven interfaces at which the standing wave ratio is 2.32/1.45. In addition, there are three remaining standing wave ratios of 2.32 (at the air-high index interface), 1.92/1.45 (at the interface between **E** and **L**) and 1.92/1.52 at the interface between the **E** layer and the substrate. The product standing wave ratio V for the entire stack [from Eq. (5-6g)] is

$$V = 2.32 \left[\frac{2.32}{1.45}\right]^7 \frac{1.924}{1.45}\frac{1.924}{1.52} = 104.5 \qquad (5\text{-}6s)$$

which is essentially identical to the value obtained in §5.3.2.2. The advantage of using Eq. (5-6g) is that, unlike the formulation in §2.7.3.1, there is only one equation, regardless of whether the stack contains an even number or an odd number of layers.

5.3.2.2 Output coupler with specified transmittance

The previous section assumes that it is permissible to produce a reflectance that *exceeds* the required amount. Such is not the case in the design of an output coupler OC for a laser. An example is an OC with a 3.8% transmittance at 488 nm. The coating materials have indices of $n_H = 2.32$ and $n_L = 1.45$ and are deposited upon a BK-7 substrate ($n_s = 1.52$). The substitution of $q = 3$ and $q = 4$ in Eq. (5-6c) shows that transmittances of 6.7% and 2.6% are obtainable. These values bracket the *desideratum* of 3.8%.

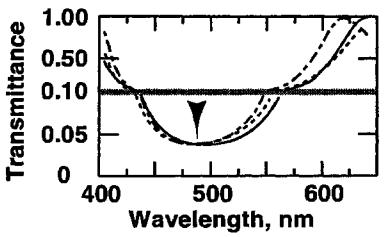


Fig. 5-34 $T(\lambda)$ of the output couplers (*solid*)
air (**H L**)^3 **H** 0.731**L** 1.61**H** *glass*, (*short dash*)
air (**H L**)^4 0.285**L** 0.397**H** 0.285**L** *glass*
and (*long-short dash*)
air (**H L**)^2 0.285**L** 0.397**H** 0.285**L** (**L H**)^2 *glass*
where the refractive indices of *glass*, **L** and **H** are 1.52, 1.45 and 2.32, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 488 nm. The ordinate changes scale at 0.10.

A coating with V_{ex} of 105 manifests the required transmittance of 3.8%. There are several ways to produce this V. The first is to include an equivalent layer and the other is to use an admittance matching coating composed of a pair of layers of unequal thickness.

Solution 1 One of the **H** layers is replaced by an equivalent layer **E** of 90° phase thickness. The resulting stack is

air (**H L**)^4 **E** *glass*

and its V is [see Eq. (2-48b)]

$$105 = V = \frac{n_H^{2q}\, n_E^2}{n_L^{2q}\, n_s\, n_0} . \qquad (5\text{-}6i)$$

Solving for n_E with $q = 4$ results in $n_E = 1.92$. The procedure of §2.11.3 is used to synthesize the equivalent layer. Figure 5-34 graphs the transmittance of the coating.

The equivalent layer is positioned next to the substrate in the foregoing coating. It is evident from Eq. (2-48b) that the standing wave ratio at λ_0 is unaltered, provided the **E** layer is located at any *odd* position in the stack. That is, the coatings

air **H L H L E H L H L** *glass* and *air* **E H L H L H L H L** *glass*

have identical reflectance at λ_0. Figure 5-34 shows that the offband transmittance is different, however. Although at λ_0 the reflectances of these multilayers are equal, the absorption is different. §5.3.10 demonstrates that the absorption of a quarterwave stack is relatively low—the inclusion of the nonquarterwave layers at the top surface should increase the absorption.

Solution 2 Another approach is to use a *pseudo-substrate*, as represented by the normalized admittance (see §2.2.5.2) of a two-layer stack. This admittance is often fictitious—that is, it is attainable using any known material. Suppose a seven layer stack is used. Its standing wave ratio (with $q = 3$) is

$$105 = V = \frac{n_H^{2q+2}}{n_L^{2q}\, n_p\, n_0} \qquad (5\text{-}6j)$$

where n_p is the refractive index of the pseudo-substrate. Solving for n_p results in 0.85. The procedure of §4.3.4.2.2 is followed to produce a two-layer coating

air **L' H"** *glass*

that transforms the normalized admittance of the substrate (1.52) *downwards* to 0.85. Its design is 0.731L 1.61H. The transmittance of the multilayer reflector is the solid curve in Fig. 5-34.

5.3.3 Rejection filters

A rejection filter attenuates over a narrow bandwidth and transmits other wavelengths, as illustrated in Fig. 1-49. Filters with either a homogeneous or inhomogeneous refractive index profile are examined.

5.3.3.1 Quarterwave stack

The narrow spectral bandwidth of the reflectance peak in Fig. 5-9 is attributed to its small index difference between the H and L layers. Equation (5-5c) shows that the bandwidth shrinks still further as the index difference diminishes. The fractional bandwidth B of the first-order stop band is approximated

$$B \equiv 2\Delta g/g \approx \frac{4}{\pi} \frac{|n_a - n_b|}{n_a + n_b} . \tag{5-7a}$$

The index difference Δn is defined

$$\Delta n \equiv |n_a - n_b| . \tag{5-7b}$$

If $n \approx n_a$, $n \approx n_b$ and $\Delta n/n_a << 1$, Eq. (5-7a) reduces to

$$B \approx \frac{2}{\pi} \frac{\Delta n}{n} \tag{5-7c}$$

and is directly proportional to the fractional index difference. Thus if a 1% bandwidth is required, a fractional index difference of about 1.6% is requisite.

Next, consider a stack containing an even number of layers in which the refractive indices of the incident medium and substrate are identical. Its maximum absorbance D_{max} is related to T_{ex} [see Eqs. (5-6a) and (1-6a)] by

$$D_{max} \equiv \log(1/T_{ex}) \tag{5-7d}$$

where the log is decadic. Combining Eqs. (5-6a), (5-6b), (5-7c) and (5-7d) results in

$$D_{max} \approx 0.43 \frac{\pi}{2} B \ell - 0.6 \tag{5-7e}$$

where $\ell \equiv 2q$ is the number of layers.

Equation (5-7e) establishes the trade-offs between the bandwidth and number of layers. For a given rejection D_{max}, the bandwidth is inversely proportional to the number of layers. Thus, if 300 layers are required to produce a stop band of 2% width, 600 layers are needed to produce the same level of rejection with a 1% width.

5.3.3.2 Rugate filters

A rugate (from Latin *rugare*, to wrinkle) filter has a sinusoidal refractive index profile. In other words, the layers are optically inhomogeneous (see §3.5.3.3). Its advantage is that the reflectances of all stop bands—other than the first order — are suppressed. The disadvantage is that the D_{max} of the first-order stop band [see Eq. (5-7d)] is less than that of a stack with homogeneous layers with the same number of periods. Its stop band width is also smaller than that of a quarterwave stack.

As an example, a periodic medium is designed in which the lowest index n_a is 1.45 and the highest n_b is 2.32. There are four possibilities: (1) The refractive index varies sinusoidally with metric thickness; (2) it is sinusoidal with optical thickness;

(3) the log index vs metric thickness is sinusoidal; and (4) the log index is sinusoidal with optical thickness. At distance z below the surface, the optical thickness OT(z) is

$$\mathrm{OT}(z) \equiv \int_0^z n(z')\,dz' \ . \tag{5-8a}$$

The index profiles as a function of the coordinate z are

$$n(z) = \exp\left\{ \tfrac{1}{2}[\ln(n_b)+\ln(n_a)] - \tfrac{1}{2}[\ln(n_b)-\ln(n_a)]\cos(2\pi\,\mathrm{OT}(z)/\lambda_0)\right\} \tag{5-8b}$$

and

$$n(z) = \tfrac{1}{2}(n_b+n_a) - \tfrac{1}{2}(n_b-n_a)\cos(2\pi\,\mathrm{OT}(z)/\lambda_0)) \tag{5-8c}$$

for the log and linear systems, respectively. Only the log index vs optical thickness profile truly suppresses the higher order stop bands [Carniglia 87, Bloom 88].

Table 5-3 Attributes of quarterwave stack and rugate filters

Attribute	Rugate	QW stack
Normalized stop band width	1.26	1.34
Maximum absorbance	4.52	6.11

Attributes of the stacks

air (basic period)^{16} *glass*

where the *glass* index is 1.52. The *basic period* of the quarterwave (QW) stack consists of **HL** — layers of optical thicknesses $\lambda_0/4$ and refractive indices of 2.32 and 1.45, respectively. The *rugate* design is captioned in Fig. 5-35.

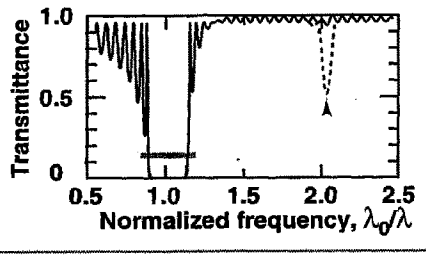


Fig. 5-35 $T(\lambda_0/\lambda)$ of *air* (basic period)^{16} *glass* (*glass* index = 1.52) where the continuously varying refractive index of the basic period is simulated with 53 finite steps. (*Solid*) Eq. (5-8b) is used, with n_a = 1.45 and n_b = 2.32. (*Dashed*) Eq. (5-8c) is used. An arrow designates a second harmonic reflection peak and the horizontal shaded line the width of a stop band of a quarterwave stack of homogeneous layers of the same indices.

The foregoing profiles were simulated by 53 finite steps of equal optical thickness and the *basic periods* were generated for the log and linear profiles. Figure 5-35 shows the transmittances of

air (basic period)^{16} *glass*

vs normalized frequency. The transmittances in the stop band spectral regions are quite similar. At the second harmonic, however, the linear index profile [see Eq. (5-8c)] manifests a second harmonic reflection peak.

Given the maximum refractive index difference of 1.45 and 2.32, it is possible to compare the performance of a quarterwave stack with a rugate. The horizontal shaded line in Fig. 5-35 is the spectral width of the stop band of a quarterwave stack consisting of indices n_a and n_b. Table 5-3 shows that the width of the rugate stop band is 6% below that of the quarterwave stack. For the 16 periods, the rugate manifests a maximum absorbance that is 1.5 density units less than that of the quarterwave stack. A rugate is simulated with layers of finite thicknesses in §5.3.5.4.

5.3.4 Blocking filters and reflectors

This section surveys several aspects of blocking filter and reflector design, such as bandwidth and harmonics of reflection peaks.

5.3.4.1 Attenuation and bandwidth

The OC in §5.3.2.2 is simple to design because it need attain only a given transmittance over a narrow range of wavelengths. This section addresses the problem of a single stack blocking filter that manifests a transmittance of less than 0.1% over as wide a spectral band as possible. The coating is a quarterwave stack with layers of indices $n_H = 2.32$ and $n_L = 1.45$.

Given that the transmittance should be less than 0.001, several questions arise: (1) Does the stop band width provide any inkling of the width of the spectral region over which the attenuation attains a predetermined level? (2) Over what bandwidth is the transmittance below 0.1% and how many layers are required to produce this attenuation?

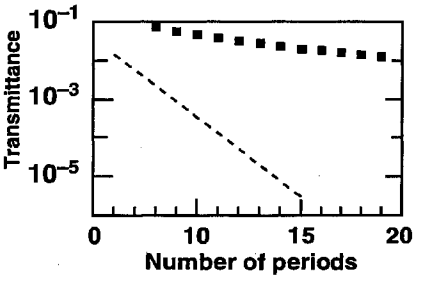


Fig. 5-36 Transmittance vs number of periods q of the blocking filter

glass $(\mathbf{HL})^q$ *glass.*

where the indices of **L**, **H** and glass are 1.52, 1.45 and 2.32, respectively. All layers are $\lambda_0/4$ in optical thickness. (*Squares*) T at stop band edge. (*Dashed*) T at quarterwave point.

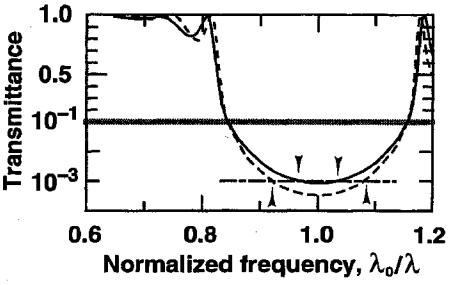


Fig. 5-37 $T(\lambda_0/\lambda)$ of

glass $(0.5\mathbf{H}\,\mathbf{L}\,0.5\mathbf{H})^q$ *glass*

where the indices of *glass*, **H** and **L** are 1.52, 2.32 and 1.45, respectively. Optical thickness of **L** and **H** is $\lambda_0/4$. (*Solid*) $q = 9$; (*dashed*) $q = 10$. The ordinate changes from linear to log at 0.1.

5.3.4.1.1 Attenuation at the stop band edge and elsewhere

From Eq. (5-5c) the semiwidth of the stop band is

$$\Delta g = \frac{2}{\pi} \arcsin\left[\frac{(2.32/1.45)-1}{(2.32/1.45)+1}\right] = 0.148 \ . \qquad (5\text{-}8d)$$

The total bandwidth is $2\ \Delta g = 0.296$. An appendix (see §5.8.2.1) shows that an approximate value of the transmittance at the stop band edge is found as follows: (1) Assume that the refractive indices of the incident medium and substrate are the same — as, for example, the configurations "air-to-air" or "glass-to-glass." (2) Compute the reflectance R_1 and transmittance T_1 of the basic period at the *stop band edge.* (3) The transmittance T_q of q periods is

$$T_q = (1 + q^2\, R_1/T_1\,)^{-1} \qquad (5\text{-}9a)$$

which shows that the transmittance decreases at a rather slow pace as more layers are added to the stack. The conclusion is that although it is useful to compute the wavelengths at which the stop band is located— to the extent that it predicts a region of

"low" transmittance or "high" reflectance—the attenuation *at the stop band edge* itself is rather poor.

This is illustrated in Fig. 5-36, which shows the transmittance (on an absorbance scale) of the quarterwave stack at its stop band edge

glass $(\mathbf{HL})^q$ *glass* .

Because the incident medium and substrate are of the same index, Eq. (5-9a) may be used to find the transmittance at the edge of the stop band—shown as the squares in Fig. 5-36. Depicted in the same Fig. is the minimum transmittance T_{min}

$$T_{min} \approx 4 \frac{n_L^{2q}}{n_H^{2q}} \qquad (5\text{-}9b)$$

which is a straight line on a semilog plot. This occurs when the layer thicknesses are all quarterwaves. Whereas the transmittance at the stop band edge barely approaches 1% with the use of 20 periods, the minimum transmittance is five decades lower—below 10^{-7}.

5.3.4.1.2 Spectral width of the blocking region

This section determines the spectral width over which transmittance is below 0.1%. The data were obtained from a digital computer. The basic period is an "equivalent layer" form of a quarterwave stack

glass $(\frac{\mathbf{H}}{2}\mathbf{L}\frac{\mathbf{H}}{2})^q$ *glass* .

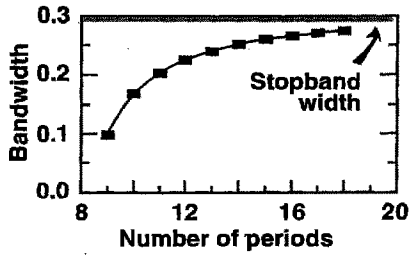


Fig. 5-38 Fractional bandwidth vs number of periods q for the basic period captioned in Fig. 5-37. The bandwidth is the spectral region where the transmittance is below 0.1%.

Figure 5-37 shows that its minimum transmittance is less than 0.1% when more than nine periods are used. This Fig. also depicts the bandwidth (distance between the arrows), which is the frequency span in which the transmittance does not exceed 0.1%.

Figure 5-38 plots the bandwidth vs number of periods. The width of the stop band is the shaded line at the top of the graph. Although the bandwidth doubles as the number of periods progresses from nine to eleven, the incremental gains are progressively smaller as more than twelve periods are added.

5.3.4.1.3 Determination of the number of periods

§5.3.4.1 poses the question: What bandwidth can be attained? The answer: Although Fig. 5-38 shows that as many as 18 periods (37 layers) could be used, the incremental gain of bandwidth over 14 periods is marginal. The number of periods is limited by the cost of depositing the layers, the mechanical stress and absorption.

5.3.4.2 Harmonic reflection peaks

§5.2.5 illustrates that although a coating may provide the required attenuation at λ_0, it may be unacceptable because of undesirable harmonic reflection peaks. This topic is explored further below.

5.3.4.2.1 Example — 633 nm reflector

The specification for this coating is that it manifest a narrow reflection peak at λ_0 = 633 nm, with 40% < R_{max} < 50%. The remainder of the visible spectrum should be transmitted. The bandwidth is defined arbitrarily as the spectral width at the 20% reflectance level. The coating contains layers of silica (n_L = 1.45) and alumina (n_H = 1.63). The reflection bandwidth of the quarterwave stack with a first-order stop band at λ_0 is simply too broad, which impels the designer to consider higher order stop bands. Figure 5-39 shows several coatings and the bandwidth of each. They include a 2:1 stack, a 3:1, a 5:1 and a 1:1. The order (see §5.2.3.4) at λ_0 is three for both the 1:1 and 5:1. This means that it is possible that a reflectance peak exists at the fourth order — at a wavelength of 633×3/4 = 474 nm. Figure 5-40 shows that the 5:1 stack exhibits just such a peak. The fifth-order stop band (and reflectance peak) of the 1:1 stack is missing because the layers are absentee (see §5.2.4.3.1).

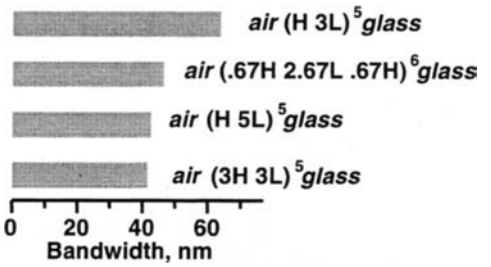


Fig. 5-39 Spectral bandwidth of the design that appears at the right of the shaded bar. The indices of *glass*, **H** and **L** are 1.52, 1.63 and 1.45, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$ at λ_0 of 633 nm. The *bandwidth* is the spectral region where the reflectance exceeds 20%.

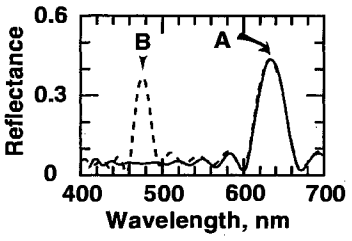


Fig. 5-40 $R(\lambda)$ of the three-quarterwave stack (*solid*) *air* $(3\mathbf{H}\ 3\mathbf{L})^6$ *glass* and five-to-one stack (*dashed*) *air* $(\mathbf{H}\ 5\mathbf{L})^6$ *glass* . The symbols appear in the caption to Fig. 5-39.

5.3.5 Suppression of reflection peaks of higher order stop bands

From the discussion of §5.2.5 it is evident that a *single* reflection peak of a higher order stop band may be suppressed in a two-component coating by choosing the layer thicknesses properly. For a quarterwave stack, for example, Fig. 5-24 shows that the reflectances of all stop bands with even order numbers are suppressed. Epstein [55] and Thelen [63, 73] have shown that at least three materials must be used in a stack to suppress two or more *contiguous* stop bands. The approach in §5.3.5 is based on the theory of Kard (see §2.4.3.4). Mathematical details are outlined in several sections in an appendix — starting in §5.8.4. §5.3.5.2 presents a design procedure for multicomponent coatings containing as many as five different coating materials. It is often possible to reduce the number of coating materials to two, as shown in §5.3.5.3.

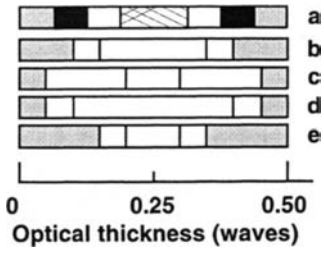


Fig. 5-41 Optical thicknesses of layers in basic periods that suppress reflectance bands at order numbers $m = 2, 3, 4, 5$ and 6 (design *a*) and at $m = 4, 5$ and 6 (designs *b*, *c*, *d* and *e*). The indices of the shaded and cross-hatched layers are 1.45 and 2.30, respectively.

5.3.5.1 Overview of the problem and definition of the terminology

This section defines the terminology and provides examples of reflectors with suppressed reflection peaks.

5.3.5.1.1 Definition of the basic period

The basic period is symmetrical about the center layer and consists of five or more layers, designs such as **A x B x A** (see Fig. 5-14), **A x y B y x A** or **x y z B z y x A**. A multilayer reflector results when such a basic period is exponentiated q times:

$(\mathbf{AxBxA})^q$ $(\mathbf{AxyByxA})^q$ or $(\mathbf{AxyzBzyxA})^q$.

If the center layer is a halfwave, then the reflector design is written

$(\mathbf{AxBBxA})^q$ $(\mathbf{AxyBByxA})^q$ or $(\mathbf{AxyzBBzyxA})^q$

Reflectance peaks at stop bands of higher order number are suppressed if the refractive indices of the **x**, **y** and **z** layers are chosen properly. The order number of the stop band is defined in §5.2.3.4. As discussed in §5.3.5.2, the refractive indices of the **A** and **B** layers are constrained. The **x**, **y** and **z** layers are termed "index to be found," since their refractive indices are expressed in terms of the innermost **A** layer and outermost **B** layer.

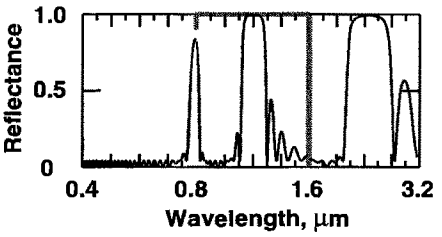


Fig. 5-42

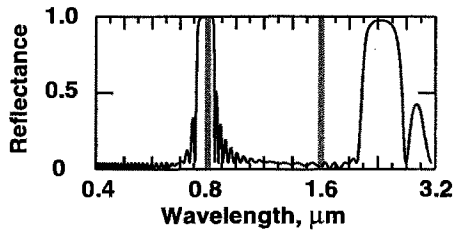


Fig. 5-43

$R(\lambda)$ of IR reflectors of the design *air* **(Basic period)**^{9} *glass* where the refractive index of *glass* is 1.52. The designs appear in Table 5-4. The abscissa changes scale at 0.8 μm and 1.6 μm.

Table 5-4
Reflectors with suppressed reflectance at higher order stop bands

Fig.	Basic period design	x index	y index	ID	q	R_{max} order 1	R_{max} order 2
-	**AAA x BB x AAA**	1.729	-	*e*	9	0.99108	0.99565
5-42	**A x BBBBBB x A**	1.929	-	*d*	9	0.98254	0.99569
5-43	**A x BB xxx A**	1.826	-	*c*	9	0.97453	-
-	**AA x BBBB x AA**	1.826	-	*b*	9	0.99863	-
5-44	**A x y BB y x A**	1.660	2.010	*a*	7	0.9832	-

Designs of the basic periods that suppress the reflectance at stop bands of order numbers $m = 4$, 5 and 6 (ID = *b*, *c*, *d* and *e*) and $m = 2$, 3, 4, 5 and 6 (ID = *a*). The refractive indices of **A** and **B** are 1.45 and 2.30, respectively. The refractive indices of **x** and **y** appear in the *Index* columns. R_{max} is the peak reflectance of stop bands of order numbers one or two of

air (Basic period)^q *glass*

where *glass* has an index of 1.52 and q appears in the above table. A computed curve of its spectral reflectance appears in the Fig. listed in the first column.

5.3.5.1.2 Examples of basic periods

As an example, Fig. 5-41 depicts the optical thicknesses of four basic periods, each of which manifests a reflectance peak at λ_0 of 2.4 μm. In each case, all harmonic reflectance peaks are suppressed in the visible portion of the spectrum, at order numbers $m = 4$, 5 and 6. Inspection of the $R(\lambda)$ curves depicted in Figs. 5-42 and 5-43 shows that the reflectance bands in the visible have been suppressed. Each of the basic periods designated as *b*, *c*, *d* and *e* in Table 5-4 contains three materials.

Figure 5-41 also shows a basic period containing four materials. The advantage of using four materials is that the reflectance of *all* stop bands from $m = 2$ to $m = 6$ may be suppressed, as shown in Fig. 5-44.

5.3.5.1.3 Number of suppressed stop bands

The first question is, how many stop bands may be suppressed? This depends upon whether the center of the basic period is occupied by a layer of halfwave optical thickness. If p is the total number of "index to be found layers" and the center layer is a halfwave, then $p+1$ higher order stop bands may be suppressed. Otherwise, it is possible to suppress the reflectance of only p stop bands.

Consider the coating captioned in Fig. 5-44. It contains four "index to be found" layers — its center layer is a halfwave. It can, therefore, suppress as many as five contiguous stop bands. As another example, consider the basic period **A x y z B z y x A**. The center layer is not a halfwave and there is a total of six "index to be found" layers. This coating can suppress the reflectance of six or fewer contiguous stop bands.

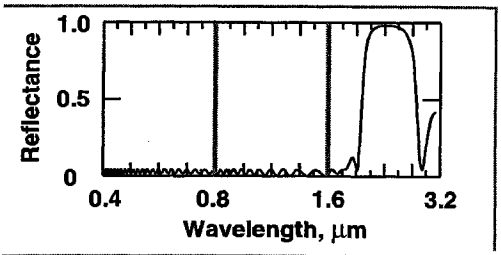


Fig. 5-44 $R(\lambda)$ of the hot mirror *air* (A x y BB y x A)^7 *glass* where the refractive indices of A, x, y, B and *glass* are 1.45, 1.66, 2.01, 2.30 and 1.52, respectively. The optical thickness of A, x, y and B is $\lambda_0/16$ at $\lambda_0 = 2.4$ μm. The scale of the abscissa changes at 0.8 μm and 1.6 μm.

5.3.5.2 Design procedure

Figure 5-45 delineates basic periods and the order numbers of the stop bands that each suppresses. The following design procedure is used:

(a) Select the wavelength λ_0 of the first-order stop band.

(b) Choose the order numbers of the higher order stop bands that should be suppressed.

(c) Select a design from the possible basic periods listed in Table 5-5.

(d) Given that an optimum reflectance at λ_0 is to be achieved, refractive index of the A layer is chosen as the lowest possible and the index of the B layer as the highest attainable.

(e) The index (or indices) of the "index to be found" layer(s) is (are) found from equations designated in Table 5-5. These equations appear below.

(f) The optical thicknesses of the layers in the basic period are adjusted such that the total optical thickness of the period is $\lambda_0/2$.

(g) The exponent q is chosen such that an adequate reflectance is achieved at λ_0.

5.3.5.2.1 Refractive indices

Given the constants a and b appearing in Table 5-5, equations for the refractive indices of the x, y and z layers appear below. n_a and n_b are the refractive indices of the A and B layers, respectively.

$$n_x \approx \exp[a\ \ln(n_a) + (1-a)\ln(n_b)\ ] \quad (5\text{-}9a)$$

$$n_y \approx \exp[b\ln(n_a) + (1-b)\ln(n_b)\ ] \quad (5\text{-}9b)$$

$$n_z \approx \exp[0.121\ln(n_a) + 0.879\ln(n_b)\ ] \quad (5\text{-}9c)$$

$$n_x = \sqrt{n_a\, n_b} \quad (5\text{-}9d)$$

$$n_y = \sqrt{n_a\, n_b} \quad (5\text{-}9e)$$

$$n_y = n_a\, n_b / n_x \quad (5\text{-}9f)$$

$$n_z = n_a\, n_b / n_x \quad (5\text{-}9g)$$

5.3.5.2.2 Examples

In the first example, the *desideratum* is a reflector with a first-order stop band at $\lambda_0 =$ 2400 nm. It should be highly transmissive in the visible spectral region (400 nm to 700 nm). The third-order reflectance peak is situated at 2400/3 = 800 nm, which is outside of the visible. The seventh order peak is at 2400/7 = 343 nm in the UV. The reflectance peaks with order numbers $m = 4$, 5 and 6 are to be suppressed. Inspection of Fig. 5-45 (in conjunction with Table 5-5) shows that there are eleven possible basic periods that would accomplish this. That eleven shrinks to nine if fewer than five materials are used and reduces to four if the coating is restricted to three materials. The latter option is chosen.

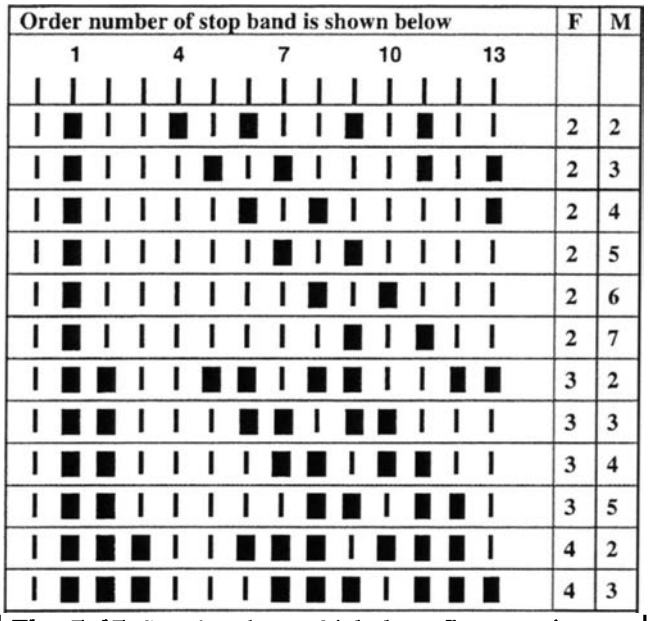


Fig. 5-45 Stop bands at which the reflectance is not suppressed (*broad rectangles*) and is suppressed (*narrow rectangles*) by the basic periods listed in Table 5-5. *F* is the lowest order number suppressed stop band and *M* is the number of contiguous suppressed stop bands, starting at *F*.

Assuming that oxide materials are used in the multilayer, the outer **A** layer index is chosen as 1.45 and the inner **B** layer as 2.30. The index of the **x** layer (of three quarterwave optical thickness) is obtained from Table 5-5 in the design **A xxx BB xxx A**. It is the geometric mean of the indices of the inner and outer layers

$$n_x = \sqrt{1.45\ 2.30} = 1.826\ . \qquad (5\text{-}9h)$$

The total optical thickness of the basic period should be 1200 nm, which is a halfwave at λ_0. The optical thickness of **A**, **x** and **B** should be 1200/10 = 120 nm. In other words, the optical thickness of the layer of index 1.826 (represented as **xxx**) should be 360 nm, since it is three times as thick as the outer **A** layer.

In the second example, four materials are used in a multilayer that is highly reflective at $\lambda_0 = 2400$ nm. Its design is **A x y BB y x A**. The indices of the **A** and **B** layers are the same as in an earlier part of this section. According to Table 5-5, Eqs. (5-9a) and (5-9f) provide the "index to be found" values. Numerical evaluation provides

$$n_x \approx \exp[0.7071\ \ln(1.45) + 0.2929\ \ln(2.30)] = 1.66 \qquad (5\text{-}9i)$$

$$n_y \approx 1.45 \times 2.30/1.66 = 2.01 \qquad (5\text{-}9j)$$

The foregoing appear in the caption to Fig. 5-44. The optical thickness of **A**, **x**, and **y** is a sixteenth of a wave at λ_0 and the optical thickness of the center layer **BB** is an eighth of a wave. Fig. 5-44 depicts its spectral reflectance.

Table 5-5
Multilayers with suppressed reflectance at higher order stop bands

F	*M*	*a*	*b*	OT	Design	Eqs.
2	2	0.3820	-	0.1000	**A x B x A**	(5-9a)
2	3	-	-	0.0833	**A x BB x A**	(5-9d)
2	4	0.6431	0.1981	0.0714	**A x y B y x A**	(5-9a) (5-9b)
2	5	0.7071	-	0.0625	**A x y BB y x A**	(5-9a) (5-9f)
2	6	0.7733	0.4260	0.0556	**A x y z B z y x A**	(5-9a) (5-9b) (5-9c)
2	7	0.8090	-	0.0500	**A x y z BB z y x A**	(5-9a) (5-9e) (5-9g)
3	2	0.3569	-	0.0714	**A x BBB x A**	(5-9a)
3	2	0.4450	-	0.0714	**AA x B x AA**	(5-9a)
3	3	0.4142	-	0.0625	**A x BBBB x A**	(5-9a)
3	3	0.5858	-	0.0625	**AA x BB x AA**	(5-9a)
3	4	0.5739	0.1479	0.0556	**A x y BBB y x A**	(5-9a) (5-9b)
3	4	0.7422	0.2578	0.0556	**AA x y B y x AA**	(5-9a) (5-9b)
3	5	0.6180	0.1910	0.0500	**A x y BBBB y x A**	(5-9a) (5-9b)
3	5	0.8090	0.3820	0.0500	**AA x y BB y x AA**	(5-9a) (5-9b)
4	2	0.3473	-	0.0556	**A x BBBBB x A**	(5-9a)
4	2	0.4679	-	0.0556	**AAA x B x AAA**	(5-9a)
4	2	0.2578	-	0.0556	**A xxx B xxx A**	(5-9a)
4	2	0.4260	-	0.0556	**AA x BBB x AA**	(5-9a)
4	3	0.6180	-	0.0500	**AAA x BB x AAA**	(5-9a)
4	3	0.3819	-	0.0500	**A x BBBBBB x A**	(5-9a)
4	3	-	-	0.0500	**A xxx BB xxx A**	(5-9d)
4	3	-	-	0.0500	**AA x BBBB x AA**	(5-9d)

Designs of multilayer reflectors that suppress the reflectance of the stop bands listed in Fig. 5-45. *F* is the lowest order number suppressed stop band. *M* is the number of contiguous suppressed stop bands, starting at *F*. The constants *a* and *b* refer to equations (in Eqs. column) for the refractive indices of the **x** and **y** layers, respectively. OT is the optical thickness (in waves at λ_0) of **A**, **B**, **x**, **y** and **z** such that the first-order stop band is located at λ_0.

5.3.5.3 Heat reflector with two components

Lotz [87] demonstrated that an equivalent layer may be used in lieu of the actual material for an "index to be found" layer in Table 5-5. In this example, a heat reflector at λ_0 =2250 nm is designed using the basic period **AA x BBB x AA**. Table 5-5 shows that it suppresses the reflectances of stop bands with order numbers m = 4, 5. With indices of 1.45 for the **A** layer and 2.30 for the **B** layer, Eq. (5-9a) shows that the index of the **x** layer should be 1.89. The goal is to replace it with an equivalent layer composed of materials of indices 1.45 and 2.30.

The next question: What is the phase thickness β_e of the equivalent layer and at what wavelength should it manifest this thickness? According to Table 5-5, the optical thickness of the x layer is $\lambda_0/18$. It is important that it have the proper phase thickness in the wavelength range $\lambda_0/5 \leq \lambda \leq \lambda_0/4$. At a wavelength λ_1 midway in frequency between orders 4 and 5, the phase thickness of the x layer is a quarterwave—i.e., 90°. A calculation shows that λ_1 is 500 nm. The equivalent layer should manifest a 90° phase thickness and index of 1.89 at that wavelength. The procedure of §2.11.3 is used to synthesize the equivalent layers.

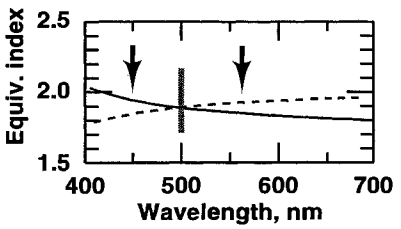


Fig. 5-46 Index of the equivalent layer of the design
0.295**L** 0.379**H** 0.295**L** 0.352**H** 0.271**L** 0.352**H** (*solid*), and (*dashed*).

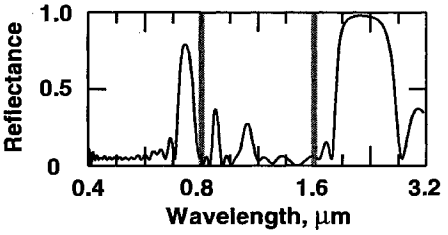


Fig. 5-47 $R(\lambda)$ of
air (2**L** 0.295**L** 0.379**H** 0.295**L** 3**H** 0.295**L** 0.379**H** 0.295**L** 4**L** 0.352**H** 0.271**L** 0.352**H** 3**H** 0.352**H** 0.271**L** 0.352**H** 2**L**)^3 *glass* .

For both foregoing designs, the refractive indices of *glass*, **L** and **H** are 1.52, 1.45 and 2.30, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 500 nm. In Fig. 5-47, the abscissa changes scale at 0.8 μm and 1.6 μm.

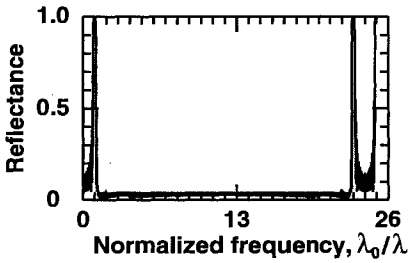


Fig. 5-48 $R(\lambda_0/\lambda)$ of the coating listed below. The refractive indices of the layers appear in §5.8.4.4. The optical thickness of **A**, **p**, **q**, **r**, **t**, **u**, **v**, **w**, **x**, **y**, **z** and **B** is $\lambda_0/48$.

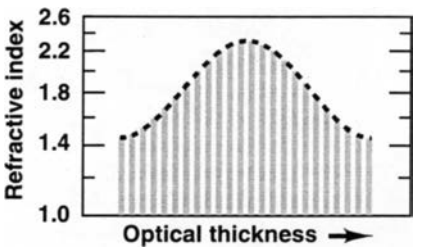


Fig. 5-49 Refractive index (on a log scale) vs optical thickness of the design cited below and in §5.8.4.4. The dashed curve is a sinusoid.

Given the refractive indices listed in §5.8.4.4, the reflector design is
air (**A p q r t u v w x y z BB z y x w v u t r q p A**)^{10} *glass* .

As explained below, two slightly different basic periods are cascaded
air (**A A x' B B B x' A A A A x" B B B x" A A**)^3 *glass* .
The difference is that the **x'** layers are simulated with an *LHL* type of equivalent layer and the **x"** layers with an *HLH* type. As explained below, this is done to correct for the dispersion of the index of the equivalent layers.

The advantage of using two different equivalent layers is that each compensates for the dispersion of the other. Both equivalent layers have an index of 1.89 and a phase thickness of 90° at the central wavelength of 500 nm in Fig. 5-46. At shorter

wavelengths, the index of the *LHL* simulation increases by 0.06 to 1.95 at 450 nm ($=\lambda_0/5$). The index of the *HLH* equivalent layer drops by nearly the same amount. The converse is true at 553 nm, where the fourth harmonic stop band is suppressed. The $R(\lambda)$ is shown in Fig. 5-47. The transmittance in the visible part of the spectrum could be improved still further with computer optimization (see §10.2).

5.3.5.4 Simulation of a rugate

As discussed in §5.3.3.2, a rugate filter has the ability to reflect strongly at λ_0, yet to manifest no reflection peaks at harmonics of λ_0. The performance of a rugate can be approximated if a sufficient number of "index to be found" layers are added to a basic period. An example is the reflector

$$(\mathbf{A\,p\,q\,r\,t\,u\,v\,w\,x\,y\,z\,BB\,z\,y\,x\,w\,v\,u\,t\,r\,q\,p\,A.})^{10}$$

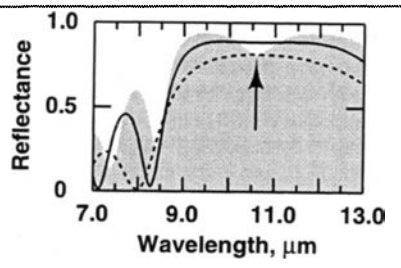


Fig. 5-50 $R(\lambda)$ of the reflector
(*dashed*)
air **L** $(\mathbf{H\,L})^3$ *substrate* ,
(*shaded*)
air **H LL** $(\mathbf{HL})^4$ *substrate*
and (*solid*)
air $(\mathbf{H\,L})^3$ **H LL** *substrate*
where the refractive indices of **L**, **H** and *substrate* are 1.40, 2.22 and 2.40, respectively. The optical thicknesses of **L** and **H** are $\lambda_0/4$ at λ_0 of 10.6 μm.

that contains twenty "index to be found" layers. Equations for the refractive indices of the layers appear in an appendix—§5.8.4.4. The center of the basic period is occupied by a layer of halfwave optical thickness. The rule from §5.3.5.1.3 predicts that the reflectance of a total of twenty one contiguous higher order stop bands may be suppressed. This is depicted in Fig. 5-48. A plot of the log of the refractive indices vs optical thickness in Fig. 5-49 shows that the profile approximates a sinusoid.

5.3.6 Spectral region of flattened reflectance

5.3.6.1 Introduction

The reflectance (*long dash*) in Fig. 5-50 is similar to that of the output coupler in Fig. 5-2 , in that both contain only quarterwave thickness layers. The reflectance attains a maximum at the "tuned" wavelength of $\lambda_0 = 10.6$ μm and drops at both the long wavelength and short wavelength sides of λ_0. Several methods are used to "flatten" this spectral region, as shown by the other curves in Fig. 5-50.

The reflectance is raised at the *edges* of the stop band and, hence, this technique is often termed *edge boost.* Two physical models provide insight into such coatings: (1) Vector addition of amplitudes; and (2) combination of a Fabry-Perot and quarterwave stack.

5.3.6.1.1 Vector addition of amplitudes

The vector addition of amplitudes in Figs. 5-4 (panel *D*) and 5-10 presents a physical picture that explains that the net amplitude reflection coefficient $\hat{\rho}$ decreases when the wavelength is less than the "tuned" value of λ_0 for a quarterwave stack. The vectors curl upwards at an ever increasing angle, thus driving $\hat{\rho}$ closer to the origin—with a resultant drop in reflectance.

Figure 5-51 depicts the vector diagram for the coating

air $(\mathbf{H}\,\mathbf{L})^3\ \mathbf{H}\,\mathbf{L}\,\mathbf{L}$ *substrate* .

The outer seven layers are quarterwaves in optical thickness. The effect of the halfwave next to the substrate is to reverse the direction of the vector r_8 that occurs at the bottom interface — between an **L** layer and the substrate. The upper panel of Fig. 5-51 illustrates how this reversal increases the net amplitude reflection coefficient $\hat{\rho}$. This Fig. shows that although vector r_7 is turning *towards* the origin with increasing wave number, r_8 is driving $\hat{\rho}$ *away* from the origin, thus increasing $\hat{\rho}$.

5.3.6.1.2 Combination of a quarterwave stack and a Fabry-Perot

Fabry-Perot is another name for a single-cavity filter (see §2.11.4), which consists of a layer of halfwave optical thickness sandwiched between two reflectors. In the simplest form, the reflectors are single quarterwave films. The design of the Fabry-Perot is

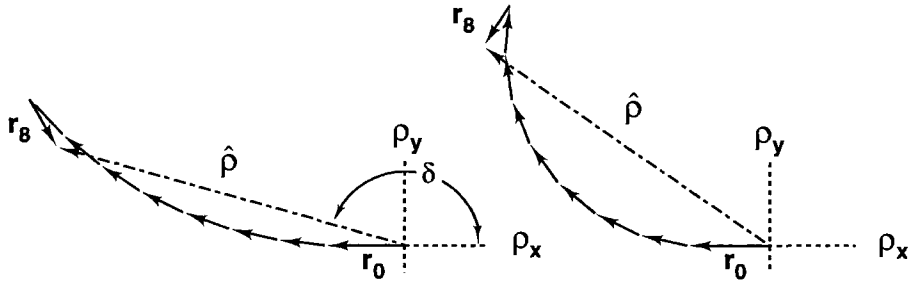


Fig. 5-51 Vector addition of amplitudes computation of the amplitude reflection coefficient $\hat{\rho}$ (*long-short dash*) for the reflector

air $(\mathbf{H}\,\mathbf{L})^3\ \mathbf{H}\,\mathbf{L}\,\mathbf{L}$ *substrate*

where the symbols are captioned in Fig. 5-50. (*Left panel*) $\lambda = 10.21\ \mu m$. (*Right panel*) $\lambda = 9.85\ \mu m$.

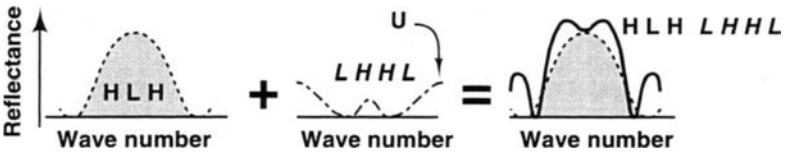


Fig. 5-52 Concatenation of a quarterwave stack (*dashed* and *shaded*) **H L H** and a three-layer Fabry-Perot (*long-short dash*) **L H H L** produces the broadband reflector (*solid*) **H L H L H H L** . The curves display $R(\sigma)$. The optical thickness of **H** and **L** is $\lambda_0/4$.

L H H L .

Figure 5-52 shows that its reflectance has a small peak at its center frequency, dips to a minimum and then increases, as at point *U* in that Fig. The reflectance of the quarterwave stack decreases at frequencies other than the center wavelength. The combination of the quarterwave stack and the Fabry-Perot produces a coating with a "W" shape in reflectance, as shown by the solid curve in Fig. 5-52.

The "plus" sign in Fig. 5-52 indicates symbolically that the two coatings — the Fabry-Perot and quarterwave stack — are "added" together. The "addition" is more than

a mere superposition of the two curves. The partitioning method (see §2.6.4.3) must be used to combine the coatings.

Figure 5-52 shows that the reflectances of the stacks

H L H L H H L and **H L H**

are the same at the "tuned" wavelength, when $\lambda = \lambda_0$. This is proven as follows. At the tuned wavelength, the **HH** layer (of halfwave optical thickness) is absentee (see §2.6.3.1) and hence may be removed from the stack. After its removal, the resulting coating

H L H L L

contains two contiguous layers of quarterwave optical thickness, which coalesce to form a halfwave thick **LL** layer. This layer is also absentee. The remaining layers are **H L H**.

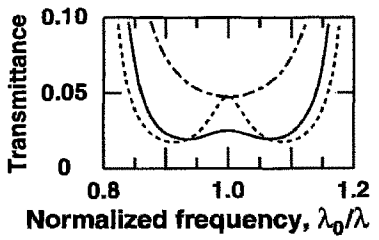


Fig. 5-53 $T(\lambda_0/\lambda)$ of the output coupler (*long-short dash*)
air $(\mathbf{L\,H})^5$ *glass*,
(*short dash*)
air **L HH L H LL H (L H)**^{5} *glass*
and (*solid*)
air **L H"H" L H'LL H (L H)**^{5} *glass*
where the refractive indices of **L**, **H"**, **H'**, **H** and *glass* are 1.45, 2.00, 1.70, 2.35 and 1.52, respectively. The optical thickness of **L**, **H"**, **H'** and **H** is $\lambda_0/4$.

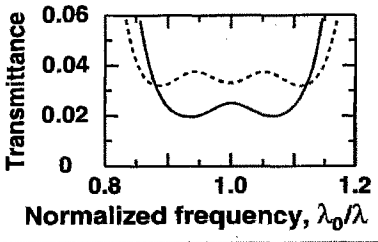


Fig. 5-54 $T(\lambda_0/\lambda)$ of the output coupler (*dashed*)
air 1.404**L** 0.871**H** 0.429**L** 0.504**H** 1.322**L**
0.796**H** 2.187**L** 1.05**H** 0.989**L** 1.002**H**
0.995**L** 0.997**H** 0.998**L** 0.999**H** 1.002**L**
1.006**H** 1.026**L** 1.417**H** *glass*
and (*solid*)
air 1.256**L** 0.452**H** 0.512**L** 0.452**H** 1.642**L**
0.206**H** 2.386**L** **(H L)**^{5} **H** *glass*
where the refractive indices of *glass*, **L** and **H** are 1.52, 1.45 and 2.35, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$.

5.3.6.2 Design examples

The coatings may also be designed with the aid of an optimization computer code (see §10.2). The coatings of §5.3.6.1.2 may be used as starting designs.

5.3.6.2.1 Simulated halfwaves

The design procedure is more intuitive than analytical. $T(\lambda_0/\lambda)$ of the starting design

air **L H" H" L H' L L H (L H)**^{5} *glass*

appears as the *solid* curve in Fig. 5-53—the indices of **H"**, **H'** and **H** are the same. The goal is to reduce this objectionable ripple of ±1% in the spectral region of its minimum transmittance.

The reflectance of the Fabry-Perot in the offband spectral region—point *U* in Fig. 5-52—contributes to the ripple. A lowering of this reflectance reduces the ripple. This is achieved by lowering the index of the spacer layer so that in the coating

L H"H" L

the **H"** layer has an index of 2.00, rather than 2.35.

The ripple is also reduced if the transmittance at the "tuned" frequency is lowered. This is accomplished by making one section of the Fabry-Perot asymmetrical by decreasing the index of the **H'** layer

air **H' L L H (L H)**^{5} *glass* .

The effect of this asymmetry is evident when the reduced admittance y (at $\lambda = \lambda_0$) is written for the foregoing stack [see Eq. (2-48b)]

$$y = \frac{n_{H'}^2 n_L^2 n_L^{10} n_s}{n_L^2 n_H^2 n_H^{10} n_0} \quad . \tag{5-10b}$$

y becomes smaller as the index of the **H'** layer is reduced, which, in turn, decreases the transmittance at λ_0.

The solid curve in Fig. 5-53 shows the transmittance of the modified stack containing four refractive indices. Each intermediate index layer is replaced with an equivalent layer—the *solid* curve in Fig. 5-54. The *dashed* curve is the result of an optimization computer code (see §10.2) that adjusted the optical thicknesses to flatten the transmittance. There are a few thin layers of eighth-wave optical thickness that could make this coating difficult to produce.

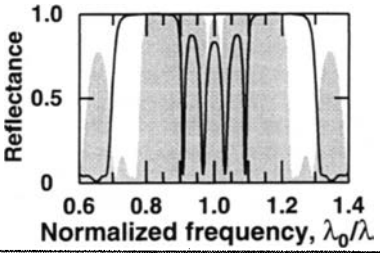


Fig. 5-55 $R(\lambda_0/\lambda)$ of the dual-band reflector (*shaded*)

air **(H L H L H)**^{5} *glass*

and (*solid*)

air **(H L H L H L H L H)**^{3} *glass* .

The refractive indices of *glass*, **L** and **H** are 1.52, 1.45 and 2.30, respectively. **H** and **L** represent layers of optical thickness $\lambda_0/4$.

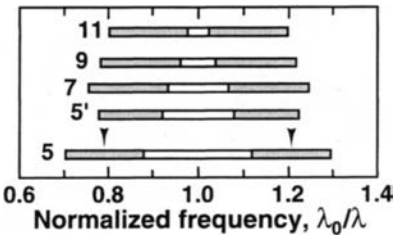


Fig. 5-56 Stop bands (*shaded*) and passbands (*nonshaded*) of the designs captioned in Table 5-6. The arrows indicate stop band centers, the ratios of which are listed in Table 5-6.

5.3.7 Dual-band reflectors

Dual-band describes a coating with a strong reflectance at just two wavelengths—λ_1 and λ_2. There is little concern if the reflectance fluctuates wildly at other wavelengths. Examples of such coatings are

air **(H L H L H)**^{5} *glass* and *air* **(H L H L H L H L H)**^{3} *glass* . .

The cascading of such basic periods causes the outer layers of each period to coalesce—producing halfwave layers, as

air **H L H L HH L H L HH L H L HH L H L HH L H L H** *glass* .

The *solid* curve in Fig. 5-55 shows that the foregoing coating reflects at frequencies of $g_1 = 0.8$ and $g_2 = 1.2$. The frequency ratio v is

$$v = g_2/g_1 = 1.2/0.8 = 1.5 \quad . \tag{5-10c}$$

The frequency ratio v for the basic period with nine layers is smaller: 1.29. Table 5-6 lists the v ratios for basic periods in which the layer count increases from five and eleven. Figure 5-56 layers out the positions of the stop bands. The frequency used in computing v is the stop band center, which is slightly displaced from the point of highest reflectance. The frequency ratios in Table 5-6 change if the refractive indices of the **H** and **L** layers are altered.

5.3.7.1 Design procedure

The goal is a coating with high reflectance at two wavelengths: λ_1 and λ_2. The design procedure is: (1) Select the indices of the two components—the **H** and **L** layers. (2) Create a table similar to Table 5-6 for several basic periods. (3) Select a stack with a v [see Eq. (5-10c)] close to λ_2/λ_1. (4) If the **L** layers are less absorbing than the **H** layers, the outer layers of the basic period should be **L**, as in **LHLHLHL**. (5) Cascade a sufficient number of periods to attain the required reflectance at λ_1 and λ_2.

Table 5-6 Optical properties of dual-band reflectors

Basic period	Lyr	Ratio v	Basic period	Lyr	Ratio v
HLHLH	5	1.53	**HLHLHLHLH**	9	1.29
HHLHLHLHH	5'	1.36	**HLHLHLHLHLH**	11	1.25
HLHLHLH	7	1.37			

For the two stop bands contiguous to $\lambda = \lambda_0$, the frequency ratio v of the frequencies of the stop band centers. *Lyr* is the number of layers. **L** and **H** represent layers of optical thickness $\lambda_0/4$, with indices of 1.45 and 2.30, respectively.

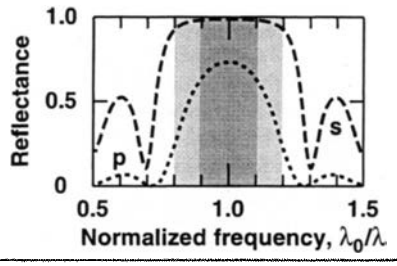


Fig. 5-57 The incidence angle is 60° in air and the design is

air **HLHLHLH** *glass*

where the refractive indices of *glass*, **L** and **H** are 1.52, 1.45 and 2.35, respectively. **H** and **L** represent layers of optical thickness $\lambda_0/4$, matched at 60°. For *p* polarization, $R(\lambda_0/\lambda)$ is *short dash* and the stop band region is *dark shaded*. For *s* polarization, $R(\lambda_0/\lambda)$ is *long dashed* and the stop band region is *light shaded*.

Suppose that a basic period with seven layers has the desired frequency ratio v. Six periods are needed to obtain the required reflectance. The multilayer is

air **HLHLHL(LHLHLHL)**^{5} *glass*

where the outermost **L** layer has been discarded to enhance the reflectance. The **L** layer next to the substrate should also be dropped.

The foregoing are also used as *a periodic medium stack* in a bandpass filter (see §7.3).

5.3.8 Performance of stacks at nonnormal incidence

Spectral transmittance curves at the beginning of Chap. 6 and especially §6.1.2.5 emphasize that the *polarization splitting* at nonnormal incidence can be quite severe. The splitting is desirable in a linear polarizer (see Fig. 5-89), but is objectionable when an edge filter is illuminated with randomly polarized light.

Most of the previously mentioned concepts and equations apply at nonnormal incidence, provided effective indices (see §2.5.3.1) are used. Consider, for example, Eq. (5-5c) for the bandwidth and Eq. (5-6b) for the transmittance at an extremum of a quarterwave stack. They are valid at nonnormal incidence when effective indices are used. Table 5-7 summarizes the results. Figure 2-60 shows that with increasing angle of incidence the effective indices for *p* polarization get closer together and farther apart for *s*. The consequence is that both the stop band width and reflectance are greater for *s* than for *p*, as depicted in Fig. 5-57.

Table 5-7 Effective indices

Pol	N_H	N_L	N_s	N_0	y	R_{max}	Δg
p	2.528	1.808	1.850	2.0	12.9	0.733	0.106
s	2.185	1.163	1.250	0.5	336	0.988	0.194

At an angle in air of 60°, effective indices (see §2.5.3.1) N_0, N_H, N_L and N_s of *air*, **H** ($n_H = 2.35$), **L** ($n_L = 1.45$) and *glass* ($n_s = 1.52$), respectively, for the *p* and *s* polarizations. For the coating *air* **H L H L H L H** *glass*, R_{max} and *y* are, respectively, the reflectance and reduced admittance (see §2.3.2) at λ_0. Δg is the semiwidth of the stop band (see §5.2.4.3.1). **L** and **H** represent layers of optical thickness $\lambda_0/4$, *matched* at 60°.

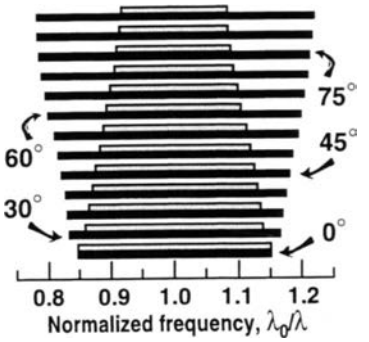


Fig. 5-58 Stop band width vs λ_0/λ of two layers of equal phase thickness, *matched at all angles*, for *p* polarization (*shaded*) and *s* polarization (*solid*). The indices are 1.45 and 2.35. The vertical scale is the incidence angle (in degrees) in air. The angle steps in five degree increments at angles greater than 30°.

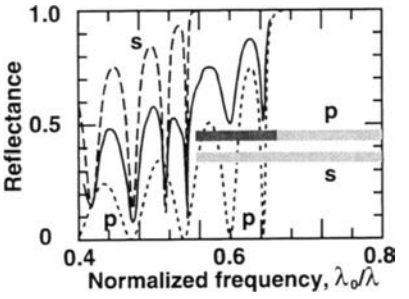


Fig. 5-59 At 60° incidence in air, $R(\lambda_0/\lambda)$ for *p* polarization (*short dash*), *s* polarization (*long dash*) and average (*solid*) of

air $(\mathbf{H}\,\mathbf{L})^8$ *substrate*

where the refractive indices of *substrate*, **L** and **H** are 1.54, 1.35 and 4.20, respectively. **H** and **L** represent layers of optical thickness $\lambda_0/4$, *matched* at 60°. Stop band widths for each polarization are shown; the non-overlap region is dark shaded.

A consequence is that if the objective is to reflect randomly polarized flux at nearly the 100% level, the multilayer must be designed for the "worst case," which is the *p* polarization. Table 5-7 shows, for example, that seven layers produce a reflectance of nearly 99% for the *s* polarization and a mere 73% for *p*. So, more layers must be added to the stack just to increase the *p* reflectance. Figure 5-58 shows the splitting of the stop

band that occurs when quarterwave stacks are designed so that the layers are matched — the phase thicknesses are equal (see §2.7.2.2). At large angles, the spectral width of the stop band is 32% greater for *s* than for *p*. This difference in stop band spectral width is exploited in the edge filter polarizer (§5.4.3.5.4).

This splitting of the polarization at the edges of the stop bands adds complexity to edge filter design. An extreme example is used—Fig. 5-59 depicts the reflectance of a germanium-chiolite stack deposited upon an alkali halide substrate. The shaded area is the gap between the stop bands for the two linear polarizations. The average reflectance [see Eq. (1-5a)] always exceeds 50% in this spectral region because, even though the *p* reflectance is often close to zero, the *s* reflectance is nearly 100%.

5.3.8.1 Reflector in a cemented cube

The goal is the attainment of a reflectance of at least 99% for the *p* polarized flux for a coating immersed in a "cemented cube," as shown in Figs. 1-180 or 1-196. Suppose that layers of **L** (cryolite, $n_L = 1.35$) and **H** (zinc sulfide, $n_H = 2.326$) are to be used. The glass prism index is 1.62, angle of incidence is 45° in the glass and the layer thicknesses are matched quarterwaves (see §2.7.2.2). Figure 6-21 shows that a T_s of 0.01 (equals R_s of 0.99) is easily attained.

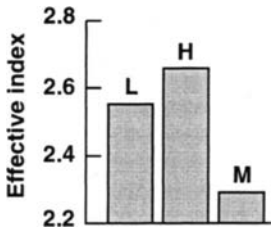


Fig. 5-60 At 45° incidence in a medium of index 1.62, the effective indices (see §2.5.3.2) for *p* polarized flux for indices of (*L*) 1.35, (*H*) 2.307 and (*M*) 1.62.

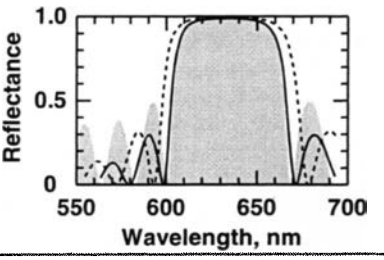


Fig. 5-61 At 45° incidence angle in *glass* of index 1.62, reflectance for *p* polarization (*solid*) and *s* polarization (*shaded*) of

glass $(\mathbf{L}\,\mathbf{M}\,\mathbf{H}\,\mathbf{M})^{12}$ *glass*

The *short-dash* curve is for *p* polarization of

glass $(\mathbf{H}\,\mathbf{M}')^{20}\,\mathbf{H}$ *glass*.

The refractive indices of **L**, **M**, **M'** and **H** are 1.35, 1.558, 1.62 and 2.307, respectively. All layers have an optical thickness of $\lambda_0/4$, *matched* at 60°, where λ_0 is 633 nm.

Although it is tempting to use the design *glass* $(\mathbf{H}\mathbf{L})^q\,\mathbf{H}$ *glass*, a computation shows that even a 101 layer stack ($q = 50$) fails to produce the needed *p* reflectance. Why? Figure 5-60 plots for *p* polarization the effective indices (see §2.5.3) of **L** and **H**. The problem is that they are quite close together, differing by only 4%. The effective index of the *M* layer in Fig. 5-60 corresponds to a material whose index is close to that of the incident medium. Inspection of Eq. (2-39b) indicates that such an index should be used at 45° incidence in order to attain the maximum spread between the effective indices for *p* polarization. Substitution of a **M** layer of index 1.62 for the **L** layer in the foregoing coating attains the sought-after reflectance for *p* polarized flux with 20 periods, as

depicted in Fig. 5-61. The spectral region of high reflectance of the *s* polarization is quite broad—R_s is not plotted in that Fig.

Another approach is to use the teachings of §6.4.4 to design a nonpolarizing reflector. The basic period contains three components

glass $(\mathbf{H}\,\mathbf{M}\,\mathbf{L}\,\mathbf{M})^q$ *glass* .

The procedure is to select a trial value for the number of periods q and to constrain one of the indices, such as that of the **L** layer. The indices of the other two layers—**M** and **H** — are then found for a given reflectance. Another value of q is used if the indices of **M** and **H** are unattainable. Figure 5-61 shows that the requisite reflectance is provided by a coating with 12 periods containing indices of (n_L = 1.35), (n_M = 1.558) and (n_H = 2.307). The index of the **M** layer must be held to quite tight tolerances if the coating is to be truly nonpolarizing. It is surprising that the nonpolarizing coating contains just seven layers more than the two-component design containing indices of 1.62 and 2.307.

5.3.9 Minus filter

The term *minus filter*, as coined by Thelen [71], designates a coating that rejects over a prescribed bandwidth and transmits on both the long-wave and short-wave sides of that rejection region. Although the spectral transmittance of numerous rejection filters are displayed in the Figures in previous sections of this chapter, the transmittances in the passband regions are often quite poor. The reflectance in Figs. 5-27 and 5-61 displays a strong ripple structure, for example.

The chapter on minus filters in Thelen [89] mentions several design methods. Several are considered here: (1) The graded index; and (2) optimization.

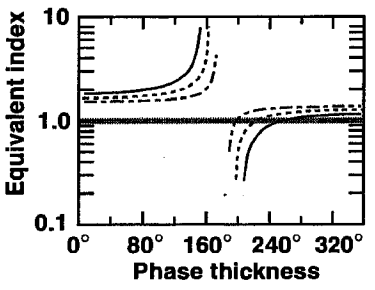


Fig. 5-62 Index of the equivalent layer vs total phase thickness of the basic period [see Eqs. (5-4d) and (2-46g)] of

(*solid*) $\frac{1}{2}\mathbf{L}\,\mathbf{H}\frac{1}{2}\mathbf{L}$,

(*short dash*) $\frac{1}{2}\mathbf{L}\,\mathbf{C}\frac{1}{2}\mathbf{L}$

and (*long-short dash*) $\frac{1}{2}\mathbf{L}\,\mathbf{B}\frac{1}{2}\mathbf{L}$

where **L**, **H**, **C** and **B** represent layers of the same phase thickness and indices of 1.45, 2.30, 1.90 and 1.62, respectively.

5.3.9.1 Graded index coating

This design method assumes that the filter may be constructed with at least three different materials. It is also helpful, although not imperative, that the coating is completely symmetrical, so that the indices of the incident medium and substrate are the same.

Four different coating materials are used in the example. They are designated as **H** (the highest index, n_H = 2.30), **L** (the lowest index, n_L = 1.45), **C** (intermediate of higher index, n_C = 1.90) and **B** (intermediate of lower index n_B = 1.62). The foregoing represent materials that approximate the indices of titania, silica, yttria and alumina, respectively, in the visible spectral region.

The first step is to form an equivalent layer basic period composed of **H** and **L**. This is termed the *core stack* and contains a sufficient number of periods to provide the needed rejection. Antireflection coatings are added to improve the passband transmittance,

cement $AR_B\ AR_C$ **core** $AR_C\ AR_B$ *glass*

or

$$\textit{cement}\ (\tfrac{1}{2}\mathbf{L}\,\mathbf{B}\,\tfrac{1}{2}\mathbf{L})^2\ (\tfrac{1}{2}\mathbf{L}\,\mathbf{C}\,\tfrac{1}{2}\mathbf{L})^2\ (\tfrac{1}{2}\mathbf{L}\,\mathbf{H}\tfrac{1}{2}\mathbf{L})^6$$
$$(\tfrac{1}{2}\mathbf{L}\,\mathbf{C}\,\tfrac{1}{2}\mathbf{L})^2\ (\tfrac{1}{2}\mathbf{L}\,\mathbf{B}\,\tfrac{1}{2}\mathbf{L})^2\ \textit{glass}\ .$$

Figure 5-62 graphs the equivalent indices of the *core*, AR_C and AR_B stacks. On the low frequency side of the stop band, the indices of the antireflection stacks grade down from that of the core to the glass, which is 1.52. The converse is true on the high frequency side—the core index is quite low and the indices of the antireflection coatings grade upwards. The equivalent-layer antireflection coating sections function most effectively when their phase thicknesses are 90° or an odd multiple thereof, which is the motivation for using two periods. That is, when β_e is 135° for a single period, the phase thickness of two cascaded periods is 270°. Figure 5-63 shows the transmittance of the final coating.

As Thelen [89] points out, the suppression of the transmission spikes in the passband spectral region is improved when the index of the incident medium and substrate equals that of the L layer.

5.3.9.2 Optimization

It is proverbial that a computer optimization code provides a solution, but little insight. This is certainly true of the rejection filter, the transmittance of which is depicted in Fig. 5-64. The purpose of the filter, as described by its designer [Lange 91], is to attenuate the green and blue lines of an argon ion laser by at least 3 absorbance units. Of the fifty layers in the stack, the thicknesses of a *core stack* of nineteen layers were constrained as QWOTs at 486 nm. The thicknesses of the outer layers were adjusted to enhance the transmittance in the passband spectral region. Although this coating contains several rather thin layers, it has been successfully manufactured on a routine basis.

5.3.10 Loss and absorption in a quarterwave stack

This section examines reflectors that attain the ultimate reflectance—where absorption and scatter losses become the limiting factors in their performance. The losses may be due to radiant absorption (that converts flux into joule heat), scattering or both. Both mechanisms remove energy from a plane wave. Any scatter of a laser mirror is loss, provided it is outside of the angular acceptance cone of the mode of the laser. §5.3.10.1 reviews the absorption qualitatively and §5.3.10.2 quantifies the absorption. §5.3.10.3 explains the absorption in terms of the standing wave electric fields and §5.3.10.3.2 uses precise analytical expressions to find the absorption.

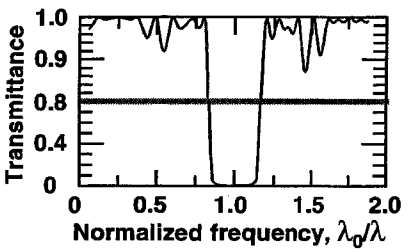


Fig. 5-63 $T(\lambda_0/\lambda)$ of the optical minus filter of the design

cement $(\frac{1}{2}\,\mathbf{L}\,\mathbf{B}\,\frac{1}{2}\,\mathbf{L})^2\ (\frac{1}{2}\,\mathbf{L}\,\mathbf{C}\,\frac{1}{2}\,\mathbf{L})^2\ (\frac{1}{2}\,\mathbf{L}\,\mathbf{H}\,\frac{1}{2}\,\mathbf{L})^6$

$(\frac{1}{2}\,\mathbf{L}\,\mathbf{C}\,\frac{1}{2}\,\mathbf{L})^2\ (\frac{1}{2}\,\mathbf{L}\,\mathbf{B}\,\frac{1}{2}\,\mathbf{L})^2$ *glass* .

The indices of *cement*, *glass*, **L**, **H**, **C** and **B** are 1.52, 1.52, 1.45, 2.30, 1.90 and 1.62, respectively. **L**, **H**, **C** and **B** represent layers of optical thickness $\lambda_0/4$. The ordinate changes scale at 0.8.

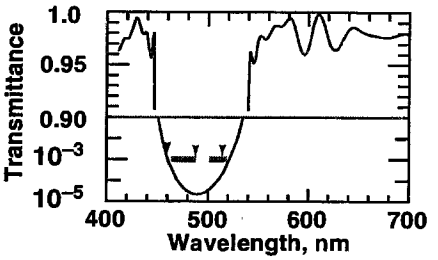


Fig. 5-64 $T(\lambda)$ of a rejection filter (containing 50 layers) whose design is listed in §5.8.5.1. The ordinate changes from linear to logarithmic at 0.9. After Lange [91]. The arrows designate the emission lines of an argon ion laser that were attenuated.

5.3.10.1 Overview

The goal of the semi-quantitative overview of this section is to present lucidly the physical model. Subsequent sections are more quantitative. The concepts of loss, absorption and so on are introduced by analyzing a reflector for an UV argon ion laser. Its performance is evaluated at wavelengths of 333 nm, 351 nm and 364 nm. The multilayer consists of quarterwaves of hafnia and silica

air $\mathbf{H}\,(\mathbf{L}\,\mathbf{H})^q$ *substrate* .

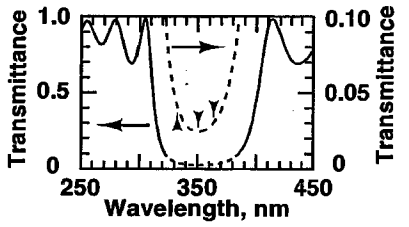


Fig. 5-65 $T(\lambda)$ of the laser output coupler reflector

air $\mathbf{H}\,(\mathbf{L}\,\mathbf{H})^7$ *substrate*

where *substrate*, **H** and **L** represent media of refractive index 1.49, 1.95 and 1.45, respectively. The extinction coefficient of **H** is 0.0003. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 351 nm. (*Solid*) Ordinate ranges from 0 to 1.0 and (*dashed*) from 0 to 0.10. Vertical arrows identify 333 nm, 351 nm and 364 nm.

The optical constants of the **H** and **L** layers are modeled as $\hat{n}_H = 1.95 - j\ 0.0003$ and $\hat{n}_L = 1.45 + j\ 0$. In other words, **H** is absorbing and **L** is not. The optical thicknesses are QWOTs at the center wavelength — 351 nm. Figure 5-65 depicts $R(\lambda)$ of a fifteen-layer (q = 7) output coupler OC for the aforementioned laser. The multilayer permits approximately 2% to 3% of the flux to leak out of the cavity.

The coating with 12 periods is a high reflector HR. Its reflectance at those three wavelengths should be as high as possible. The number of periods depends principally upon the extinction coefficient $k_H = 0.0003$ of the high index layer.

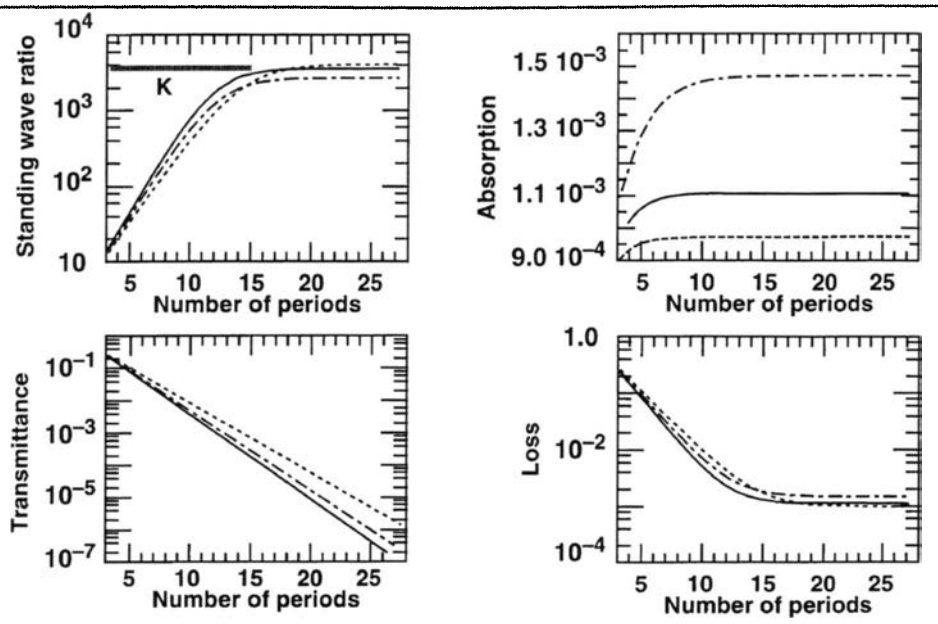


Fig. 5-66 At wavelengths of (*short dash*) 333 nm, (*solid*) 351 nm and (*long-short dash*) 365 nm, standing wave ratio, radiant absorption, radiant transmittance and loss [see Eq. (1-4b)] vs number of periods q of

air **H** (**L H**)^q *substrate*

where *substrate*, **H** and **L** represent refractive indices of 1.49, 1.95 and 1.45, respectively. The extinction coefficient of the **H** layer is 0.0003. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 351 nm.
The shaded line K is the Koppelmann limit [see Eq. (5-10d)], which was converted from a radiant reflectance into a standing wave ratio.

Figure 5-66 plots the standing wave ratio V, total absorption, transmittance and loss $\mathcal{L}$ [see Eq. (1-4b)] at the three wavelengths captioned in Fig. 5-65. As the number of periods increases from $q = 5$ (eleven layers) to $q = 25$ (51 layers), the following is observed:

(a) The transmittance decreases by at least five decades.
(b) The reflectance of the stack — which is related to the standing wave ratio — does not increase appreciably after 31 layers ($q = 10$) are deposited.
(c) The absorption saturates at $q = 10$ and does not increase any further.
(d) The loss steadily drops until it saturates at approximately 0.1%, at which point it does not decrease any further.
(e) The reflectance at the shortest wavelength (333 nm) eventually exceeds the R at both of the longer wavelengths.

Each of these points is discussed below.

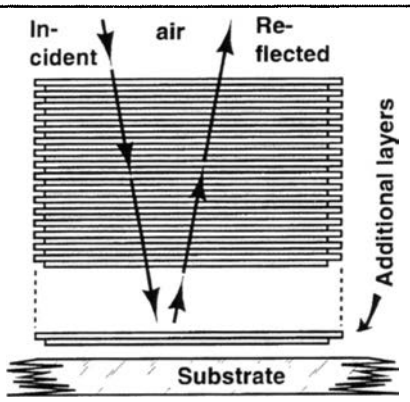


Fig. 5-67 An existing multilayer is augmented by two additional layers at the bottom of the stack. The incident flux penetrates the existing stack, reflects from the additional layers and is directed to the incident medium, where it emerges.

5.3.10.1.1 Transmittance

The transmittance may be viewed in terms of the Fresnel reflection occurring at each interface, as shown in Fig. 5-2 . Since each interface "kicks" a fraction of the flux back towards the incident medium, the addition of more interfaces decreases the amount of flux traveling towards the substrate.

From a mathematical point of view, Eq. (2-34g) shows that the transmittance depends inversely upon the square of the elements of the characteristic matrix. These matrices are exponentiated (see §5.8.1) as q increases. Hence, several of the matrix elements become quite large, which decreases the transmittance.

5.3.10.1.2 Reflectance and standing wave ratio

The essential difference between the reflectance and transmittance is illustrated in Fig. 5-67, which shows an existing stack of eleven layers. Two layers are added to the bottom of the stack, thus increasing the total number to 13 — these are the *additional layers*. Figure 5-67 shows that Fresnel reflections occur at the interfaces of the *additional layers*, thus directing flux towards the top of the stack. If the amplitude of the wave reflected from those *additional layers* is to increase the net amplitude reflection coefficient of the stack, that flux reflected at the bottom must travel through the entire stack of eleven layers above it. A portion of that flux is absorbed as it makes its journey upwards. Also, part of the flux is absorbed on its way downwards to reach those *additional layers*.

Eventually, there is a point where the small gain in the amplitude reflection coefficient produced by those *additional layers* is canceled by the absorption. Figure 5-66 shows that the standing wave ratio (and, hence, the reflectance) saturates and levels off as more periods are added to the multilayer.

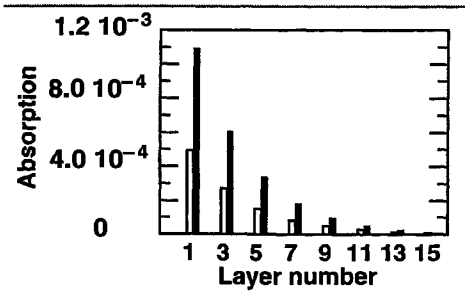


Fig. 5-68 At a wavelength λ_0, absorption in each layer (*nonshaded*) and total cumulative absorption (*shaded*) of

air **H** (**L H**)^7 *substrate*

where *substrate*, **H** and **L** represent media of refractive indices 1.49, 1.95 and 1.45, respectively. The extinction coefficient of the **H** layer is 0.0003. **L** and **H** represent layers of optical thickness of $\lambda_0/4$.

The section (§10.1.3) on graphical calculation methods shows the effects of absorption in limiting the reflectance of a stack.

5.3.10.1.3 Absorption

Figure 5-68 depicts the absorption in each layer (*nonshaded*) and the total absorption (*solid black*) in each layer of the output coupler described in §5.3.10.1. The absorption, which is related to the standing wave ratio (see §2.9.6.1), is concentrated in the top ten layers of the stack. The standing wave ratio levels off—and the absorption does likewise. That is, the absorption shown in Fig. 5-68 for 15 layers is not appreciably different than that of a 37 layer stack.

5.3.10.1.4 Loss

The loss is the sum of two components: (1) Transmittance, i.e., "leakage"; and (2) absorption. Figure 5-69 shows the loss (for 9 and 15 periods) of the coating introduced in §5.3.10.1. The transmissive component of the loss dominates when nine periods are used and loss grows vanishingly small as more layers are added. The loss is dominated by absorption when the stack contains in excess of 30 layers.

Because the absorption saturates as the number of layers increases, the loss does also when the transmittance becomes much less than the absorption. This saturation is depicted in Fig. 5-66.

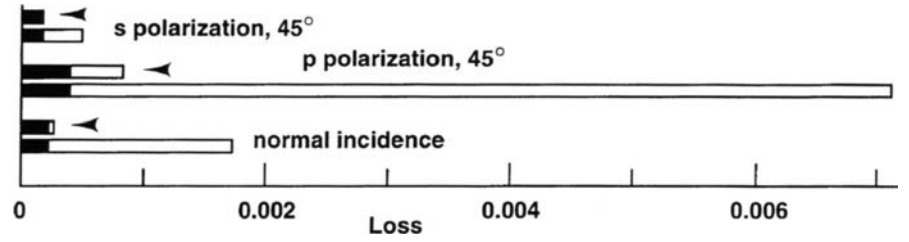


Fig. 5-69 Loss, including absorption (*dotted*) and transmittance (*shaded*), at normal incidence (*lower*) and at 45° for *p* and *s* polarizations of

air **H** (**L H**)^q *substrate*

where the symbols have the same meaning as in the caption to Fig. 5-68. (*Arrowed*) $q = 15$. (*Nonarrowed*) $q = 9$.

5.3.10.1.5 A parable

The year is 1270. News of the journey to Cathay of Niccolò and Maffeo Polo has reached Padua. Giovanni Pellicola, a renowned (if not fictional) merchant of that city, was determined to follow the footsteps of the Polo family. He desired to trade gold for

spices in Cathay, which he intended to sell upon his return to Italy. He returned empty handed.

Why? The problem is that each fiefdom that he traversed on his journey extracted tribute. The result was that one-half of his gold was depleted by the time he reached Cathay. Nevertheless, he purchased the spices and headed for home. The same fiefdoms extracted tribute again (as spices) on his homeward journey. His spices were completely gone when he returned to his native city.

The analogy is that the absorption depletes the amplitude of a wave as it travels from the incident medium to the bottom of a stack (see Fig. 5-67). After it reflects from those bottom layers, it is depleted still again as it travels towards the outer surface.

Naturally, this intrepid merchant never attempted the journey again. Had they extracted a smaller toll, the fiefdoms would have been able to extract a toll on each of his many trips—yes, extracted it again, again and again. "You can shear a sheep many times, but only slaughter him once."

5.3.10.2 Koppelmann's equation

Koppelmann [60] derived an expression for the limiting loss $\mathcal{L}_a$ of a stack. When $R \approx 1$,

$$\mathcal{L}_a \approx \frac{2\,\pi\, n_0\,(|k_H| + |k_L|)}{n_H^2 - n_L^2}\ . \tag{5-10d}$$

The equation is simple and hence a hand-held calculator (rather than a larger computer) may be used to evaluate it. Its features are: First, it is independent of the substrate index. The rationale is that if many absorbing layers are used, the radiant flux does not penetrate to the substrate and hence it has no influence. $\mathcal{L}_a$ also depends inversely on the difference of the squares of the refractive indices. This means that the largest index difference should be used in order to minimize $\mathcal{L}_a$. Finally, $\mathcal{L}_a$ is proportional to the sum of k_H and k_L. Thus, it makes no difference whether the loss is concentrated in the high index or low index layer, since their contributions to $\mathcal{L}_a$ are linear.

As an example, consider the loss of a quarterwave stack of hafnia and silica

$$\mathcal{L}_a \approx \frac{2\,\pi\, n_0\,(|0.0003| + 0)}{(1.95)^2 - (1.45)^2} = 0.0011. \tag{5-11a}$$

From Eq. (2-19f) the loss is translated into the maximum standing wave ratio V_m by

$$V_m \approx 4/\mathcal{L}_a\ . \tag{5-11b}$$

The foregoing value is plotted as the shaded line on the standing wave panel of Fig. 5-66.

Koppelmann used matrix theory to derive Eq. (5-10d) and later Sparks [79] used standing waves. An appendix (§5.8.6) presents a short derivation. Koppelmann's equation can be applied at nonnormal incidence for TE polarization by substituting the appropriate effective indices [see Eq. (2-38a)]. The layer thicknesses must be matched (see §2.7.2.2)—the β of each layer is 90°—for Eq. (5-33a) to apply.

5.3.10.3 Absorption and standing waves

Equation (2-66a) shows that the absorption is proportional to the square of the electric field and refractive index of that layer. An easy way to account for the index n_i is to plot $|E_i(z)|^2$ vs *optical thickness*. A graph of this field enables the designer to visualize its relative magnitude and how the absorption is partitioned amongst the various layers.

5.3.10.3.1 Electric fields

Although the matrix formalism of §2.9.6.4 may be used to find the fields in any stack, it is useful to examine the quarterwave stack because the equations are quite simple. The electric field at each interface is either a maximum or a minimum at the tuned wavelength.

Assume that the fields $\hat{E}_B$ and $\hat{H}_B$ of a *preexisting stack* containing quarterwave thickness layers have already been computed. Two additional quarterwave layers are added, as

air **layer1 layer2 stack** *substrate* .

The characteristic matrix of the *preexisting stack* is either diagonal

$$\begin{bmatrix} \hat{E}_B \\ \mathcal{Z}_0 \hat{H}_B \end{bmatrix} = \begin{bmatrix} m''_{11} & 0 \\ 0 & m''_{22} \end{bmatrix} \begin{bmatrix} \hat{E}_s \\ \mathcal{Z}_0 \hat{H}_s \end{bmatrix} \tag{5-11c}$$

or antidiagonal

$$\begin{bmatrix} \hat{E}_B \\ \mathcal{Z}_0 \hat{H}_B \end{bmatrix} = \begin{bmatrix} 0 & jm''_{12} \\ jm''_{21} & 0 \end{bmatrix} \begin{bmatrix} \hat{E}_s \\ \mathcal{Z}_0 \hat{H}_s \end{bmatrix} \tag{5-11d}$$

depending upon whether the *preexisting stack* contains an even or odd number of layers, respectively. $\hat{E}_B$ and $\hat{H}_B$ are the total electric and magnetic fields at the outer layer of the *preexisting stack*. In either case the *preexisting stack* matrix is multiplied by the matrix representing the two additional layers of quarterwave optical thickness

$$\begin{bmatrix} \hat{E}_A \\ \mathcal{Z}_0 \hat{H}_A \end{bmatrix} = \begin{bmatrix} \frac{-n_2}{n_1} & 0 \\ 0 & \frac{-n_1}{n_2} \end{bmatrix} \begin{bmatrix} \hat{E}_B \\ \mathcal{Z}_0 \hat{H}_B \end{bmatrix} \tag{5-11e}$$

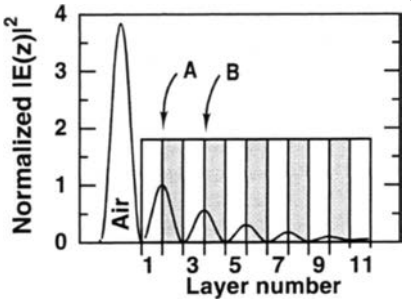


Fig. 5-70 At a wavelength of λ_0, normalized square electric field [see Eq. (2-63g)] vs optical thickness of

air **H** (**L H**)^5 *substrate*

where *substrate*, **H** and **L** represent media of refractive indices 1.49, 1.95 and 1.45, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$.

where $\hat{E}_A$ and $\hat{H}_A$ are the total electric and magnetic fields at the outermost interface of the entire multilayer, including the two additional layers. The effect of those layers is to multiply the matrix elements such as m_{11}^2 or m_{12}^2 appearing in the numerator of Eq. (2-64d) by a factor $(n_1/n_2)^2$ at normal incidence or by $(N_1/N_2)^2$ at nonnormal incidence, where N_1 and N_2 are the effective indices for the TE polarized flux. Because the matrix elements at point *B* [appearing in the numerator of Eq. (2-64d)] are unaltered, the square electric field at that point decreases.

Figure 5-70 shows the aforementioned decrease of the normalized square electric field of five periods

air **H** (**L H**)^5 *glass* or

air $\mathbf{H}_1$ $\mathbf{L}_1$ (**H L**)^4 **H** *glass*

where **H** and **L** refer to the hafnia and silica layers of §5.3.10.1. The former method of expressing the design emphasizes that the square electric field of the *preexisting stack* $(\mathbf{H}\,\mathbf{L})^4$ — represented in Fig. 5-70 as point B— increases by an amount $(n_H/n_L)^2$ with the addition of the $\mathbf{H}_1\,\mathbf{L}_1$ pair. The converse is also true; the field decreases by an amount $(n_L/n_H)^2$ as it penetrates more deeply into the stack through each $\mathbf{H}_1\,\mathbf{L}_1$ pair.

There are several consequences. The first is that most of the scattering and absorption occurs in the outer layers of the quarterwave stack. The outer layers are most susceptible to laser damage because the electric field is the highest.

Finally, the normalized square electric field of a quarterwave stack can be plotted easily if the amplitude reflection coefficient $\hat{\rho}$ is known. As discussed in §2.3.2, the phase shift upon reflection of 180° for this odd-layered stack places the node of the standing wave at the surface. The maximum of the electric field is, therefore, located at a distance of $\lambda/4$ in front of the outermost layer and has a maximum amplitude of

$$|\hat{E}_0|^2 = |\hat{E}_0^{(+)}|^2\ (1 + |\hat{\rho}|)^2\ . \tag{5-12a}$$

It follows that the maximum normalized square electric field is close to 4 when the reflectance exceeds 99%, which would be the case if a few more **H** and **L** layers were added to the coating captioned in Fig. 5-70.

Absorption of the ith layer of the stack

air $(\mathbf{H}\,\mathbf{L})^7\ \mathbf{H}$ *substrate*

where **H** and **L** have a phase thickness of 90° and optical constants $1.95 - j\,0.0003$ and $1.45 - j\,0$, respectively. The index of the *substrate* is 1.49. Ψ_i is the net flux ratio for that layer. $\prod \Psi_{i-1}$ is the final term in Eq. (2-61f) and A_i' is the absorption of the ith layer using that equation. V_i is the standing wave ratio of the stack on the emergent side of the ith layer, assuming that the stack is nonabsorbing. G_i appears in Eq. (2-63f). A_i'' is computed from Eq. (2-63e), assuming R is that of a nonabsorbing stack.

Given that the normalized square electric field is 4.00, the field at point A in Fig. 5-70 is computed by considering that the outermost "layer" is fictitious and composed of air — thus $n_1 = 1.00$ in Eq. (5-11e). The next layer is the outer **H** layer of the stack. Because the field amplitudes are pure real, the normalized square electric field at point A is

$$\left(\frac{E_A}{E_0^{(+)}}\right)^2 = 4\frac{(1.00)^2}{(1.95)^2} = 1.05\ . \tag{5-12b}$$

5.3.10.3.2 Absorption — exact and approximate

The absorption is computed by an approximate method in §2.9.6 and by an exact method in §2.9.5.1. Both methods are used to find the absorption of

air $\mathbf{H}(\mathbf{L}\,\mathbf{H})^7$ *glass*

where **H** and **L** represent layers of optical thickness $\lambda_0/4$ and refractive indices of 1.95 and 1.45, respectively. The k value of the high index layer is 0.0003. Table 5-8 summarizes the absorption at $\lambda = \lambda_0/4$. The net flux ratio formulation [see Eq. (2-61f)] is used in one instance and the square electric field [using Eq. (2-63e)] is also employed. Table 5-8 shows that the amount of absorption computed by the two methods is approximately the same — both for each layer and the total absorption for the stack.

Table 5-8 Absorption in each layer of a quarterwave stack

i	Ψ_i	$\prod\Psi_{i-1}$	$1-\Psi_i$	A_i'	V_i	G_i	A_i''
1	0.98083	1.0	0.01917	0.0004894	82.8	41.4	0.000490
2	1.0	0.98083	0	0	61.6	30.8	0
3	0.98919	0.98083	0.01081	0.0002707	45.8	22.9	0.000271
4	1.0	0.97022	0	0	34.1	17.0	0
5	0.99395	0.97022	0.00605	0.0001498	25.3	12.7	0.000150
6	1.0	0.96435	0	0	18.8	9.4	0
7	0.99662	0.96435	0.00338	0.0000831	14.0	7.0	0.000083
8	1.0	0.96109	0	0	10.4	5.3	0
9	0.99810	0.96109	0.00190	0.0000465	7.7	3.9	0.000047
10	1.0	0.95927	0	0	5.8	3.0	0
11	0.99891	0.95927	0.00109	0.0000267	4.3	2.3	0.000027
12	1.0	0.95823	0	0	3.2	1.7	0
13	0.99933	0.95823	0.00067	0.0000165	2.4	1.4	0.000016
14	1.0	0.95758	0	0	1.8	1.2	0
15	0.99950	0.95758	0.00050	0.0000122	1.3	1.0	0.000012
			Sum	0.001095		Sum	0.001096

Reflectance R of absorbing stack: 0.974477 $1-R = 0.024428$	R of nonabsorbing stack: 0.9755384 $1-R = 0.024462$

5.3.10.4 Nonnormal incidence

Figure 5-69 shows the loss at the "tuned wavelength" at 45° angle of incidence. The thicknesses of the layers are *matched* (see §2.7.2.2) so that the 90° phase thickness of each is preserved. The transmittance has the same general features as those observed in §5.3.8—namely that T_p always exceeds T_s. An illustration is that the T_s of the nine-period stack is less than the T_p of a fifteen-period stack.

Effective indices (see §2.5.3) may be substituted into Eq. (5-10d) to find the limiting loss for TE polarized flux. Consider the coating in §5.3.10. The effective indices at an 45° incidence angle appear in Table 5-9. The limiting loss $\mathcal{L}_a$ for s polarization is

$$\mathcal{L}_a \approx \frac{2\pi N_0 (|k_H| + |k_L|)}{N_H^2 - N_L^2} \approx \frac{2\pi\, 0.707 \times 0.0003}{(1.817)^2 - (1.266)^2} = 0.00078. \qquad (5\text{-}12c)$$

which is 1.4 times smaller than at normal incidence, as depicted in Fig. 5-69. For the 15 period stack, the transmittance loss in that Fig. is so small that it could not be plotted.

5.3.10.4.1 Quarterwave stack — nonnormal incidence, TE

The calculation of the fields for the TE polarized wave (at $\theta_0 = 56.1°$) follows directly from the previous section. §2.4.4 shows that for a given layer, its characteristic matrix at normal incidence is converted to nonnormal incidence by the substitution of effective indices. It follows that since the equations in §2.9.6.5 are obtained from this matrix, they may be used at nonnormal incidence, provided appropriate effective indices are used.

The thicknesses of the layers are adjusted so that the *matched* condition obtains (see §2.7.2.2). Table 5-9 lists the normalized square field in the incident medium as 3.93. The foregoing value is close to 4.00, which occurs when $R \approx 1.00$.

Using the effective indices in Table 5-9, the ratio in Eq. (5-12b) for the TE wave is $3.93 \times (0.557)^2 (1.764)^{-2} = 0.392$. The normalized square electric field (appearing in Fig. 5-71) at the electric outermost interior interface (point *A*) is attenuated by nearly a factor of 10. The field all but vanishes a few periods deeper into the stack.

5.3.10.4.2 Quarterwave stack — nonnormal incidence, TM

The computation of the square electric field for the TM polarization is more complicated, since the projections of the field along both the y and z axes must be considered. There is an equivalent of Eq. (5-12a) at nonnormal incidence. At the maxima, the electric field in the incident medium is proportional to

Table 5-9 Effective indices and ratios of electric fields

Angle	N_H	N_L	N_0	$N_0^2 N_H^{-2}$	$N_L^2 N_H^{-2}$	R	Max $\lvert E \rvert^2$
$\theta = 0°$	1.95*	1.45*	1.00*	0.263	0.553	0.921	3.71
$\theta = 45°$ TE	1.817	1.266	0.707				
$\theta = 45°$ TM	2.092	1.661	1.414				
$\theta = 56.1°$ TE	1.764	1.189	0.557	0.10	0.454	0.982	3.93
$\theta = 56.1°$ TM	2.155	1.768	1.794			0.679	2.30

*Actual index of this medium.

Effective index N [see Eqs. (2-38a) and (2-38b)] of media at incidence angles of 0°, 45° and 56.1° for TE and TM polarized flux. In the incident medium, the reflectance R and maximum normalized square electric field $|E|^2$ [see Eq. (2-63g)] of the coating

air $(\mathbf{H}\,\mathbf{L})^5\ \mathbf{H}$ *substrate*

where the indices of **H**, **L** and *substrate* are 1.95, 1.45 and 1.49, respectively. The phase thickness of each layer is 90°, regardless of the incidence angle.

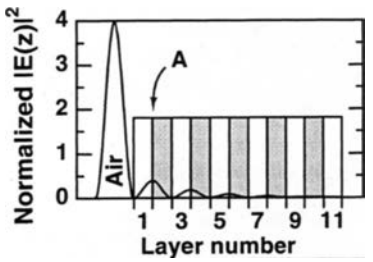


Fig. 5-71 Plotted for *s* polarization.

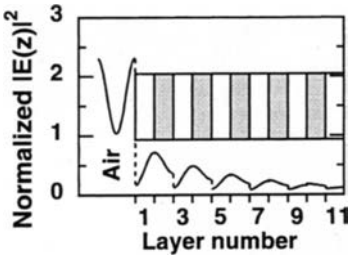


Fig. 5-72 Plotted for *p* polarization.

At a wavelength of λ_0 and incidence angle of 56.1°, normalized square electric field [see Eq. (2-63g)] vs optical thickness of

air $\mathbf{H}\,(\mathbf{L}\,\mathbf{H})^5$ *substrate*

where *substrate*, **H** and **L** represent media of refractive indices 1.49, 1.95 and 1.45, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$, *matched* at 56.1°.

$$|\hat{E}_i|^2 \propto (1 \pm |\hat{\rho}|)^2 \cos^2\theta_0 + (1 \mp |\hat{\rho}|)^2 \sin^2\theta_0 \qquad (5\text{-}12c)$$

and the values at the minima are

$$|\hat{E}_i|^2 \propto (1 \mp |\hat{\rho}|)^2 \cos^2\theta_0 + (1 \pm |\hat{\rho}|)^2 \sin^2\theta_0 \qquad (5\text{-}12d)$$

provided that the phase shift upon reflection is 180°. Equation (5-12d) describes the maxima and Eq. (5-12c) the minima if the phase shift is zero. In either case there is no standing wave pattern in the incident medium and the E field is uniform everywhere, provided the incidence angle is 45°. Wiener observed this nearly a century ago (see §7.4 of Born & Wolf [75]).

The magnitude of the square electric field for the TM polarized wave for the coating in §5.3.10.4.1 is computed from Eq. (2-65e). Equation (5-11d) shows that the diagonal matrix elements are zero and the off diagonal elements are

$$m''_{12} = N_L^5 \; N_H^{-6} = (1.768)^5 \; (2.155)^{-6} = \; 0.172 \qquad (5\text{-}13a)$$

and

$$m''_{21} = \left(m''_{12}\right)^{-1} = 5.8 \qquad (5\text{-}13b)$$

where the effective indices in Table 5-9 are substituted for N. The m''_{12} in Eq. (5-13b) is small and m''_{21} is relatively large. The numerator of Eq. (2-65e), which contains m''_{21}, is multiplied by the fourth power of the refractive index. A consequence is that at the upper interfaces of the odd numbered layers, the discontinuity in the square electric field is proportional to that fourth power. This is illustrated in Fig. 5-72 and is contrasted with Fig. 5-71 which shows the fields for the TE polarization, where there are no discontinuities.

At the even-numbered interfaces, the converse is true: m''_{12} is quite large and m''_{21} is small. It follows from Eq. (2-65e) that the field is quasi-continuous at that interface, since the coefficient of the "index to the fourth power" term is relatively small.

5.3.10.5 Ultimate reflectance

The ultimate reflectance might be required for a variety of reasons. It might be necessary to extract the maximum output from a laser with low gain or to reflect from a mirror many times, as in laser isotope separation. Table 5-8 illustrates that the absorption in the outer layers of the stack determines the minimum loss—i.e., maximum reflectance. Little can be done to reduce the loss if a quarterwave stack is used and the extinction coefficients k_H and k_L of the **H** and **L** layers are nearly the same.

In fact, it is often the case that one of the components—typically the **L** layer —is less absorbing than the other. Several authors [Lissberger 78, Carniglia 80] observe that a 2:1 or 3:1 stack has smaller loss than a quarterwave stack. Those authors also survey other publications on this subject and provide nonlinear equations that may be used in their design. The problem is that the aforementioned equations are useful *if* k_H and k_L are known. They are, in fact, usually not only unknown but also are quite difficult to measure.

Figure 5-73 shows the standing wave ratio vs number of periods of a stack in which the Koppelmann limit of loss is exceeded by 13% in a stack in which k_H is ten times larger than k_L. This is accomplished by thickening the **L** layers and thinning the **H**. In contrast to the standing wave ratio plot in Fig. 5-66, the ordinate is linear in Fig. 5-73. Shown in the same plot is the standing wave ratio vs number of periods where the extinction coefficient is the same for both layers. The coefficient is the arithmetic mean of the k_H and k_L mentioned at the beginning of this paragraph.

Figure 5-74 graphs the layer thicknesses when k_H is ten times larger than k_L. The stack was generated by adding pairs of layers to the existing stack, starting with a single QWOT **H** layer. A computer code adjusted the metric thicknesses of each pair of layers so that the optimum reflectance was obtained. At the bottom of the stack the QWOTs of both the **L** and **H** layers are quite close to that of the tuned wavelength of 361 nm. As more periods are added, the thicknesses of the **L** layers gently increase and the **H** layers (which have a higher k value) shrink in thickness.

5.3.10.6 Discussion of loss

At the time of this writing (in the year 2000), the author is not aware of any published experimental data that show how to apportion the loss between scattering and absorption. And in the case of scattering, there are three possible mechanisms:

1. Scattering at the interfaces between the layers.
2. Scattering in the bulk of the layers.
3. Scattered flux is trapped in high index layers and propagates laterally in them in a waveguide mode of propagation.

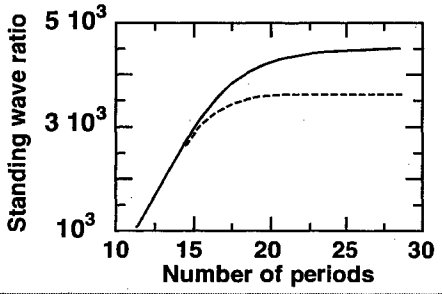


Fig. 5-73 At a wavelength of λ_0 of 351 nm, standing wave ratio vs number of periods q of

air **H**(**L H**)^q *silica*

where the *silica* refractive index is 1.49. (*Dash*) The optical constants of **H** and **L** are $1.95 - j0.00015$ and $1.45 - j\,0.00015$ and their optical thicknesses are $\lambda_0/4$ at $\lambda_0 = 334$ nm and at 368 nm, respectively. (*Solid*) The optical constants of **H** and **L** are $1.95 - j0.0003$ and $1.45 - j0.00003$, respectively. Their QWOTs appear in Fig. 5-74.

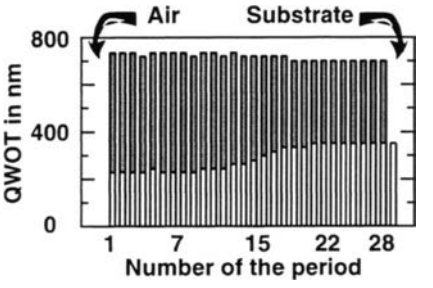


Fig. 5-74 At a wavelength λ_0 of 351 nm, the layer thicknesses are adjusted to maximize the reflectance. Optical thickness (in QWOTs, units of nm) of each pair of layers vs the number of the period. The optical constants appear in Fig. 5-73. **H** layers (*lower, light shaded*) and **L** layers (*upper, dark shaded*). Period 1 is adjacent to air and the last period is adjacent the substrate.

The standing wave analysis of the foregoing sections is applicable if absorption and bulk scattering are the predominant loss mechanisms. §5.3.10.3.1 relates the absorption to the square electric field and Fig. 5-71 illustrates the rapid decrease of the fields in the interior of the stack. Table 5-8 and Fig. 5-68 illustrate that the absorption is concentrated in the outer layers of the stack.

A consequence is that materials with the lowest absorption (i.e., k value) should be used near the top of the stack. An example is a multilayer reflector at $\lambda_0 = 250$ nm. Scandium oxide has a slightly lower k than hafnia or zirconia at that wavelength and, hence, would produce a mirror of higher reflectance. At one time the scandia was quite

expensive and an acceptable compromise was to use hafnia/silica layers at the bottom of the stack and scandia for a few high index layers near the top of the multilayer.

Another example is the production of a multilayer mirror with alternate layers of tantala and silica. Two evaporation sources of the former are used. As eleven or more high index layers are successively deposited from source *A*, the oxygen is depleted from the evaporant and the absorption of those layers slowly creeps upwards. But for the crucial two layers near the top of the stack, the virgin tantala source *B* is used, which results in films with lower absorption.

5.3.10.6.1 Experimental data

Arnon [78a] constructed a precision reflectometer that measured the reflectance of quarterwave stacks consisting of alternate layers of tantala and silica. The reflectance was converted into standing wave ratio and is plotted in Fig. 5-75. The "saturation" effect is quite evident — the standing wave ratio levels out at about 37 layers.

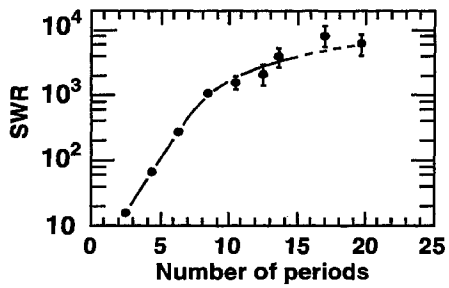


Fig. 5-75 At a wavelength of 520 nm, standing wave ratio (from measured reflectance data) vs number of periods q of
air **H** (**L H**)^q *substrate*
where **H** and **L** represent layers of quarterwave optical thickness of tantala and silica, respectively. The *substrate* is BK-7. The error bars represent the experimental uncertainty.
After Arnon [78a].

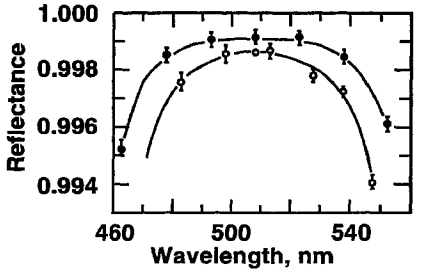


Fig. 5-76 Measured $R(\lambda)$ of (*solid circles, upper*)
air **H** (**L H**)^8 *substrate*
and (*open circles, lower*)
air 5**H** (**L H**)^8 *substrate*
where **H** and **L** represent layers of quarterwave optical thickness of tantala and silica, respectively. The error bars represent the experimental uncertainty. The *substrate* is BK-7. After Arnon [78a].

Another experiment is to determine the absorption of the various materials by adding to the top of the stack a layer whose optical thickness is a halfwave or multiple thereof [Arnon 78a]. Suppose that the objective is to measure the extinction coefficient k_H of the high index component. The reflectances of

air 4**H H** (**L H**)^8 *glass* and *air* **H** (**L H**)^8 *glass*

should be equal at λ_0 if there were no absorption in the layers. The standing wave electric field in the outer "4**H**" layer (of fullwave optical thickness) reduces the reflectance when absorption is present. k_H is determined through computer modeling. Figure 5-76 illustrates that the fullwave overcoat decreases the peak reflectance by at least 0.05%. This coating was baked in air at 400 °C for four hours after it was coated to reduce the absorption.

One precaution must be taken in the deposition of the stacks mentioned in the foregoing paragraph — the coatings should be deposited simultaneously. That is, a 17-

period quarterwave stack is deposited on two substrates. One of the substrates is covered with a shutter and the fullwave "4H" layer deposited on the other substrate.

5.4 Edge filter design

5.4.1 Introduction

An edge filter (see §1.5.2) transmits on either the short-wavelength or long-wavelength side of its rejection band. The rejection is produced by the periodic stacks mentioned previously. Both optimization (see §5.4.2) and the introduction of equivalent layers (see §5.4.3) are used to enhance the passband transmittance. The reader who has an efficient optimization computer code and desires a quick solution might be tempted to read §5.4.2 and ignore the remainder of §5.4.

5.4.2 Optimization

§5.3.9.2 shows a rejection filter produced by optimization. That technique is quite effective in the production of an edge filter, which begins with a starting design such as

$$air\ (\mathbf{H}\,\mathbf{L})^2\ (\mathbf{H}\,\mathbf{L})^5\ \mathbf{H}\,(\mathbf{L}\,\mathbf{H})^2\ substrate$$

which may be generalized to

$$air\ (\mathbf{H}\,\mathbf{L})^{q1}\ (\mathbf{H}\,\mathbf{L})^{q2}\ \mathbf{H}\,(\mathbf{L}\,\mathbf{H})^{q3}\ substrate$$

where the total number of periods q is

$$q \equiv q1 + q2 + q3\ . \tag{5-14a}$$

The stack that is exponentiated $q2$ times—for example $(\mathbf{H}\,\mathbf{L})^5\ \mathbf{H}$—is termed the *core stack*. The multilayers on the exterior (exponentiated by $q1$ and $q3$) are termed *matching stacks*.

The *core stack* provides virtually all of the attenuation. Although the *matching stacks* may attenuate in the rejection band, it is prudent not to count on it. The matching stacks increase the passband transmittance.

5.4.2.1 Periods in the core stack

The number of periods in the core stack is usually determined by either of two considerations: (1) The required attenuation; or (2) the slope. §5.3.4.1.2 lists a coating in which the number of periods is determined by the level of the attenuation. As an example of the slope, suppose that an edge filter is required to transmit less than 2% in the spectral region 600 nm to 700 nm and more than 90% in the region 500 nm to 570 nm. Figure 5-77 graphs the transmittance of stacks with six periods and ten periods. It is evident that, although the six-period stack meets the requirements for the low transmittance in the 600 nm to 700 nm spectral region, its slope is inadequate. The slope of the ten-period stack is adequate, although perhaps marginal. In either case, it is the task of the optimization computer code to remove the transmittance dips in the 500 nm to 570 nm spectral region.

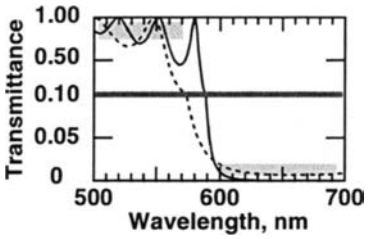


Fig. 5-77 $T(\lambda)$ of the edge filter (*dashed*) *air* **H**(**L H**)^6 *glass* and (*solid*) *air* **H**(**L H**)^{10} *glass* where *glass*, **H** and **L** represent media of refractive indices 1.52, 2.31 and 1.45, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$ at λ_0 = 680 nm (*solid*) and λ_0 = 660 nm (*dashed*). The ordinate changes scale at 0.10.

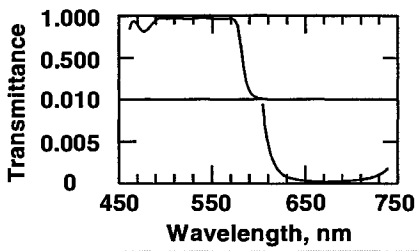


Fig. 5-78 $T(\lambda)$ of the edge filter *air* 1.056**H** 0.371**L** 0.646**H** 1.02**L** 0.151**H** 0.759**L** (0.68**H** 0.68**L**)^5 0.428**H** 0.991**L** 0.525**H** 0.563**L** 0.855**H** 0.704**L** 0.841**H** *glass* where *glass*, **H** and **L** represent media of refractive indices 1.52, 2.31 and 1.45, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$ at λ_0 = 1000 nm. The ordinate changes scale at 0.01.

5.4.2.2 Core stack and matching stacks

The reason that the coating in §5.4.2 is split into three sections—a core stack and two matching stacks—is that it is usually unnecessary for the optimization computer code to modify the layer thicknesses in the center of the stack. A common procedure is to not vary the thicknesses in the core stack and to permit the computer to adjust the layer thicknesses of the matching stacks.

Figure 5-78 shows the transmittance of an edge filter that meets the specifications mentioned in §5.4.2.1. It contains 23 layers, of which 10 are of the same optical thickness and constitute the core stack.

5.4.3 Equivalent layers in edge filters

5.4.3.1 Equivalent layers in the stop band spectral region

The equivalent index in Fig. 2-71 is plotted in the spectral region in which it is pure real. This index is pure imaginary in the stop band region, as given by

$$n_e = j\, m_{21} / \sqrt{m_{11}^2 - 1} \qquad (5\text{-}14b)$$

and its phase thickness is

$$\beta_e = \operatorname{argcosh}(m_{11}) \ . \qquad (5\text{-}14c)$$

In the stop band region the equivalent layer may be modeled as a FTR layer (see §2.4.2.8). The cascading of two or more equivalent layers is the same as a thickening of the FTR layer, which, as depicted in Fig. 2-47, decreases the transmittance monotonically.

5.4.3.2 Design procedure

The design procedure is quite similar to that of §5.4.2. These steps are followed:

(1) Establish refractive indices of the two materials to be deposited.

(2) Choose a core stack such as

air $(\frac{1}{2}$**L H** $\frac{1}{2}$**L**$)^q$ *substrate* or *air* $(\frac{1}{2}$**H L** $\frac{1}{2}$**H**$)^q$ *substrate* .

(3) Adjust the exponent q so that both the required slope and attenuation are attained (see §5.4.2.1). (4) Select an *LHL* or an *HLH* type of equivalent layer that best matches the index of the equivalent layer to both the substrate and incident medium. (5) If the passband transmittance is adequate, then the design is complete. (6) If not, one or more matching stacks (similar to those of §5.4.2) are added to enhance the passband transmittance.

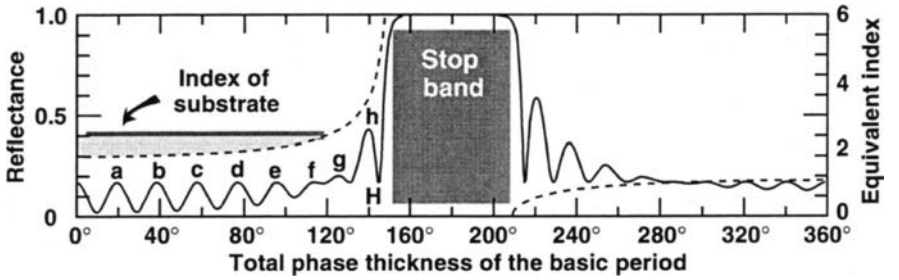


Fig. 5-79 Reflectance (*left ordinate, solid*) and index of the equivalent layer (*right ordinate, dashed*) of the coating captioned in Table 5-10. The abscissa is the total phase thickness of the basic period [see Eq. (2-46g)]. In the lightly shaded spectral region the equivalent layer index is less than the substrate index. The darkly shaded spectral region is the stop band.

5.4.3.3 Use of equivalent index to model passband reflectance

The following example illustrates that at each wavelength the equivalent layer (see step (2) of the foregoing paragraph) can either reduce the reflectance (i.e., function as an antireflection coating) or enhance the reflectance, depending upon the relationship of its index to the surrounding media. The same equivalent layer referenced in §2.6.7.3.1 is used, with two exceptions: (1) It is nine times as thick and (2) the substrate index (n_s = 2.40) is considerably higher than that of BK-7 glass.

Figure 5-79 depicts the reflectance and equivalent layer index vs the total phase thickness of the basic period [see Eq. (2-46g)]. In the β_t range from 0 to 120°, the index of the equivalent layer is less than that of the substrate; hence, the conditions of Eq. (2-23i) are satisfied. In other words, the equivalent layer acts like an antireflection coating (see §2.4.2.6). The reflectance minima are at the quarter, three-quarter, five-quarter, etc., wave points—where β_e of the cascaded nine periods is *nearly* 90°, 270°, 450° and so on. The reflectance equals that of the uncoated substrate at the halfwave, fullwave, etc., points—these are the maxima. The qualifier *nearly* in a foregoing sentence is needed because dispersion of the equivalent index shifts the extremum reflectance slightly (see Fig. 2-43).

Table 5-10 Maximum and minimum reflectances in Fig. 5-79

Extremum	β_t	n_e	β_e	R
Minimum between 0 and *a*	10°	1.78	90°	0.020
Maximum at *a*	20°		180°	U
Minimum between *a* and *b*	30°	1.79	270°	0.022
Maximum at *b*	40°		360°	U
Minimum between *b* and *c*	49°	1.83	450°	0.027
Maximum at *c*	58°		540°	U
Minimum between *c* and *d*	67°	1.88	630°	0.037
Maximum at *d*	76°		720°	U
Minimum between *d* and *e*	85°	1.95	810°	0.061
Maximum at *e*	95°		900°	U
Minimum between *e* and *f*	105°	2.13	990°	0.095
Maximum at *f*	115°		1080°	U
Minimum between *f* and *g*	118°	2.35	1122°	0.163
Maximum at *g*	126°	2.55	1202°	0.201
Minimum between *g* and *h*	130°		1249°	U
Maximum at *g*	140°	3.48	1366°	0.432
Minimum at *H*	144°		1440°	U

The features *a*, *b*, ... *g*, *h* refer to extrema of the reflectance R in Fig. 5-79. *U* is the reflectance of an uncoated substrate. β_t and n_e are, respectively, the total phase thickness [see Eq. (2-46g)] and equivalent index of a basic period $\frac{1}{2}$**L H** $\frac{1}{2}$**L**. The design of the equivalent layer is *air* $(\frac{1}{2}\mathbf{L}\,\mathbf{H}\frac{1}{2}\mathbf{L})^9$ *glass*
where refractive indices of *air*, **L**, **H** and *glass* are 1.0, 1.38, 2.30 and 2.4, respectively.

The equivalent index equals the substrate at $\beta_t = 121°$ and exceeds n_s in the frequency region between 121° and the stop band edge. The quarterwave points are maxima and the halfwave points are minima, for reasons stated in §2.4.2.6. These results are summarized in Table 5-10.

As another example, Fig. 5-80 reproduces the equivalent index vs total phase thickness of Fig. 2-71, with the exception that the shaded areas represent those values of β_t for which the inequality Eq. (2-23i) is satisfied. In other words, the index of the equivalent layer is less than the substrate index of index 1.52 but greater than 1.00. In this instance, the $\frac{1}{2}$**H L** $\frac{1}{2}$**H** combination is preferable if higher transmittance is attained on the low frequency (long wavelength) side of the stop band. The converse is true on its high frequency side — $\frac{1}{2}$**L H** $\frac{1}{2}$**L** performs better here. It is tempting to generalize and assert that the $\frac{1}{2}$**H L** $\frac{1}{2}$**H** type of equivalent layer should *always* be used for low frequency edge filters. The generalization is not valid because the choice depends upon the indices of the substrate and incident medium.

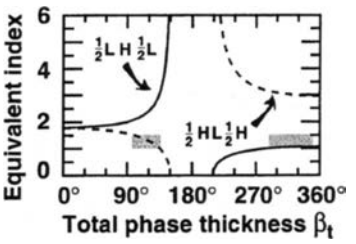


Fig. 5-80 Equivalent index vs total phase thickness [see Eq. (2-46g)] of the basic period
(*solid*) $\frac{1}{2}$**L H** $\frac{1}{2}$**L**
and (*dashed*) $\frac{1}{2}$**H L** $\frac{1}{2}$**H**
where the refractive indices of **H** and **L** are 2.30 and 1.38, respectively. In the shaded area the index of the equivalent layer exceeds that of *air* and is less than 1.52.

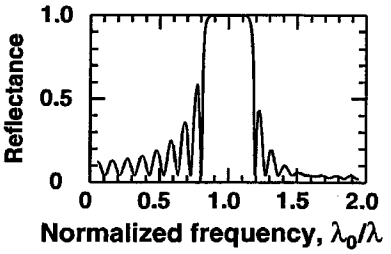


Fig. 5-81 $R(\lambda_0/\lambda)$ of the coating
air $(\frac{1}{2}$**L H** $\frac{1}{2}$**L**$)^9$ *glass*
where *glass*, **H** and **L** represent media of refractive indices 1.52, 2.30 and 1.38, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$.

5.4.3.4 Addition of matching stacks

Following Epstein [52], matching stacks may be added to the incident and emergent sides of an equivalent layer to reduce the reflectance in the passband spectral region. The core stack selected for the first example is the same as that in Fig. 5-79. It is evident in Fig. 5-81 that the reflectance on its low frequency side is considerably higher than in Fig. 5-79, which is attributed to the lower index substrate of the former. The abscissa is in units of normalized frequency g. The design goal is the reduction of the reflectance in the interval between $g = 0.7$ and $g = 0.8$.

5.4.3.4.1 Single-layer matching stacks

The layout of the coating is

air **A** (core stack) **B** *glass*

where the *core stack* is represented by an equivalent layer. Layer **A** reduces the reflectance between the incident medium and upper surface of the equivalent layer. Layer **B** accomplishes the same antireflection function between the lower surface and the substrate.

The design procedure is: (1) Compute the index of the equivalent layer — the core stack — in the passband spectral region. (2) Concoct an antireflection coating — represented by **A** — that matches the equivalent layer to the incident medium. The equivalent layer is treated as a pseudo-substrate in this calculation. (3) Choose the refractive index and thickness of **B** are so that it functions as an antireflection coating between the equivalent layer (which is regarded as a pseudo-incident medium) and substrate.

Table 5-11 Matching layers for a core stack

Parameter	$g = 0.70$	$g = 0.75$	$g = 0.80$
n_e	2.56	3.01	4.25
n_a	1.60	1.73	2.06
n_b	1.97	2.14	2.54

Equivalent indices n_e of the core stack captioned in Fig. 5-81. n_a is the square root of n_c. n_b is computed from Eq. (5-15b). g is the normalized frequency appearing in Eq. (5-5b).

The first step is to find the index of the equivalent layer at $g = 0.7$ and 0.8. Normalized frequency is translated into β_t by

$$\beta_t = 180° \, g \ . \tag{5-15a}$$

β_t is 126° at $g = 0.7$. Figure 5-80 shows that the equivalent index is increasing quite rapidly in the spectral near $\beta_t = 126°$. Its index (see Table 5-11) ranges from 2.56 to 4.25 in the spectral range of the passband.

The next step is to find the index of **B**, which is modeled as a single-layer antireflection coating. Its index should satisfy Eq. (2-23j). At $g = 0.75$, for example,

$$n_b = \sqrt{n_c \, n_s} = \sqrt{(3.01)(1.52)} = 2.14 \ . \tag{5-15b}$$

Table 5-11 shows that its index ranges from 1.97 to 2.54, depending upon the particular frequency that is chosen. A compromise is an index of 2.10. **B** should be thicker than the layers of the core stack, so that its optical thickness is a QWOT at a wave number of 0.75 λ_0^{-1}, which corresponds to a QWOT at a wavelength of $(0.75)^{-1} = 1.33 \, \lambda_0$.

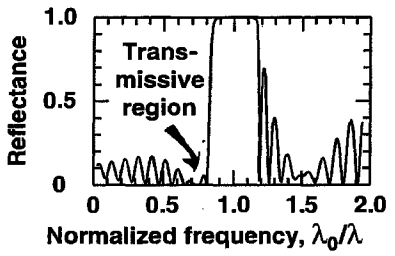


Fig. 5-82 $R(\lambda_0/\lambda)$of the edge filter

air **A** $(\frac{1}{2}\mathbf{L}\,\mathbf{H}\,\frac{1}{2}\mathbf{L})^9$ **B** *glass*

where *glass*, **H**, **A**, **B** and **L** represent media of refractive indices 1.52, 2.30, 1.70, 2.10 and 1.38, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$. The optical thickness of **A** and **B** is $\lambda_0/3$.

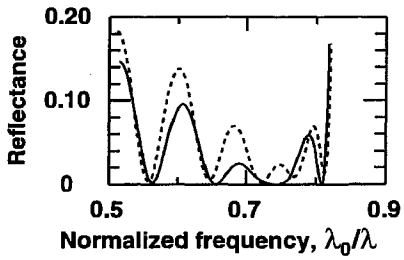


Fig. 5-83 $R(\lambda_0/\lambda)$ of the edge filter captioned in Fig. 5-82 (*solid*) and (*dashed*)

air 0.49**L** 0.42**H** 0.49**L** $(\frac{1}{2}\mathbf{L}\,\mathbf{H}\,\frac{1}{2}\mathbf{L})^9$ 1.33**H** *glass*

where *glass*, **H** and **L** represent media of refractive indices of 1.52, 2.30 and 1.38, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$.

The same logic is used to select the index of **A**, which functions as an antireflection coating between the equivalent layer (core stack) and air. Table 5-11 shows that its optimum index ranges from 1.60 to 2.06. An index of 1.70 is a compromise. Its optical thickness is the same as **B** in the previous paragraph. The final coating is

air **A** $(\frac{1}{2}\mathbf{L}\,\mathbf{H}\,\frac{1}{2}\mathbf{L})^9$ **B** *glass*

and its reflectance is plotted in Fig. 5-82. The equivalent layer is *absentee* (see §2.6.3.1) at frequencies where the phase thickness of the core stack (represented by an equivalent layer) is a multiple of 180°. The remaining layers **A** and **B** still function as an antireflection coating because the indices satisfy Eq. (4-2a).

5.4.3.4.2 Two material coating

The coating in §5.4.3.4.1 is less than desirable for several reasons. First, layers with indices of 1.70 and 2.10 may not be available. It is also possible that the materials are deposited from separate electron guns and that only two guns are available. The solution is to replace **A** and **B** with equivalent layers.

The 2.1 index film is replaced with a 2.30 index film. Although it is possible to synthesize an equivalent index layer with a 2.10 index, the **L** layers in that film would simply be too thin.

The 1.70 index is synthesized with an equivalent layer with a β_e of 90°. Either an *LHL* or an *HLH* type of equivalent layer (see §2.6.7.2) may be used. The dispersion of the equivalent layer is important. The n_e of the *LHL* equivalent layer increases as g gets larger, which means that its dispersion tracks that of the core stack. Although the n_e of an *LHL* equivalent layer does not increase rapidly enough to satisfy the indices in the final column of Table 5-11, any increase of n_e with frequency is welcomed! It is for these reasons that the *LHL* type of equivalent layer is used in preference to the *HLH* equivalent film for the **A** layer. The procedure of §2.11.3 is used to synthesize it. The edge filter coating is

air (0.49**L** 0.42**H** 0.49**L**) $(\frac{1}{2}\mathbf{L}\,\mathbf{H}\,\frac{1}{2}\mathbf{L})^9$ 1.33**H** *glass* .

Figure 5-83 graphs its reflectance.

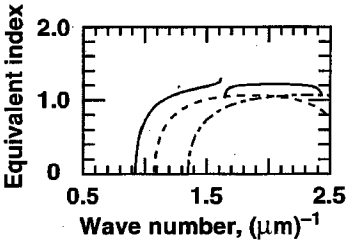


Fig. 5-84 Equivalent index vs wave number of (*solid*) 0.57(**L M H M L**) , (*long-short dash*) 0.5**L H** 0.5**L** and (*short dash*) 1.25(0.5**L H** 0.5**L**). **L**, **M** and **H** represent layers of optical thickness $\lambda_0/4$ and refractive indices of 1.38, 1.90 and 2.30, respectively. λ_0 is 860 nm. The *short dash* is the "shifted period" design.

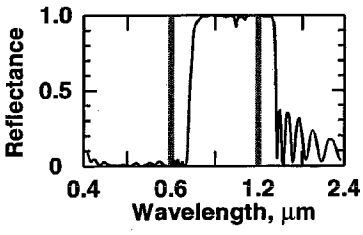


Fig. 5-85 $R(\lambda)$ of the heat reflector

air 1.1(0.5**L H** 0.5**L**) (0.5**L H** 0.5**L**)^5 1.25(0.5**L H** 0.5**L**) 0.57(**L M H M L**)^8 0.5**L** *glass*

where the symbols have the same meaning as in the caption to Fig. 5-84. The *glass* index is 1.52. The abscissa changes scale at 0.6 μm and 1.2 μm.

5.4.3.4.3 Two stack heat reflector

Two stack means that the edge filter contains more than one periodic stack. Following Thelen [63], a heat reflector is designed that transmits in the visible spectral region. This coating combines a three-component stack (see §5.3.5.2.2) such as $(\mathbf{L}\,\mathbf{M}\,\mathbf{H}\,\mathbf{M}\,\mathbf{L})^8$ with a quarterwave stack. The equivalent indices of each stack are plotted in Fig. 5-84.

The first question is, which stack should be laid on top? Figure 5-84 shows that the equivalent index n_e of the quarterwave stack matches the index of air more closely than does the three-component stack. The latter is deposited next to the glass and the quarterwave stack is adjacent to air. Because the n_e of both stacks is close to air, a single layer of MgF_2 (n = 1.38) is deposited on the glass substrate, which then lowers its normalized admittance to a value closer to 1.00.

The antireflection coating is a shifted period—its n_e is plotted in Fig. 5-84 (*short dash*). It is effective as an antireflection coating because its n_e exceeds that of the core stack in the visible spectral region. In other words, its index simulates the graded index antireflection coating similar to the one in §4.9.2. Figure 5-85 shows the spectral reflectance of the final design which could, of course, be improved by computer optimization.

5.4.3.5 Edge filters at nonnormal incidence

This is rather vast subject—a few illustrations serve as examples of edge filter design.

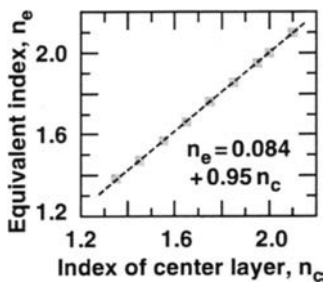


Fig. 5-86 Index of the equivalent layer vs the index of the center layer—when the phase thickness of the equivalent layer is 135°. Its design is

$\frac{2}{3}\mathbf{H}\ \frac{2}{3}\mathbf{L}\ \frac{2}{3}\mathbf{H}$

where the index of **H** is 2.30.

5.4.3.5.1 Immersed dichroic long wavelength pass

The goal is an immersed beamdivider cemented in BK-7 that reflects the blue and transmits in the green and red spectral regions. This is the blue dichroic in Fig. 1-182. The angle of incidence is 30° for an axial ray. The coating that transmits both polarizations quite well (see Fig. 1-194) contains a rare earth fluoride (n_L = 1.5) and zinc sulfide (n_H = 2.30)

$$glass\ (\tfrac{2}{3}\mathbf{H}\,\tfrac{2}{3}\mathbf{L}\,\tfrac{2}{3}\mathbf{H})^{10}\ glass\ .$$

The graphic data of Ufford [74] provides insight related to the effectiveness of the two-to-one stack. Figure 5-86 plots the index n_e of the equivalent layer (when its phase thickness β_e is 135°) vs the index of the center **L** layer. This shows that its n_e nearly equals the **L** layer index. The equivalent index also equals the index n_0 of the incident and emergent media, provided $n_L \approx n_0$. The rare earth fluoride is ideal because its index nearly matches that of the glass. At nonnormal incidence the *effective indices* (see §2.5.3) of the incident medium and equivalent layer index are nearly the same—both p and s polarized light are transmitted equally well.

5.4.3.5.2 Nonpolarizing edge filter

The data in Thelen [89], which summarizes this subject, show that it is easier to design a low frequency pass filter and more difficult to concoct a high frequency pass coating. Thelen [81a] found that the multilayer

air $(\frac{2}{3}\mathbf{L}\,\mathbf{M}\,\frac{2}{3}\mathbf{H}\,\mathbf{M}\,\frac{2}{3}\mathbf{L})^{10}$ *glass*

manifests minimal polarization splitting at the lower frequency edge of its stop band, as depicted in Fig. 5-87. It is, however, a second-order stop band (see §5.2.3.4), which means that it is intrinsically narrower.

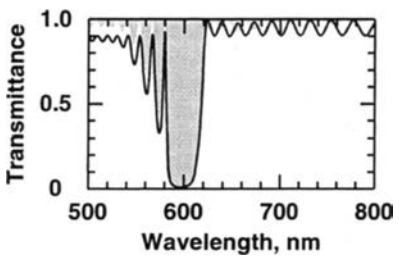


Fig. 5-87 $T(\lambda)$ at 45° incidence angle for *p* polarized flux(*shaded*) and *s* polarized flux (*solid*) of the nonpolarizing edge filter

air $(0.654\mathbf{L}\,0.98\mathbf{M}\,0.654\mathbf{H}\,0.98\mathbf{M}\,0.654\mathbf{L})^4$ $(0.667\mathbf{L}\,\mathbf{M}\,0.667\mathbf{H}\,\mathbf{M}\,0.667\mathbf{L})^{12}$ $(0.654\mathbf{L}\,0.98\mathbf{M}\,0.654\mathbf{H}\,0.98\mathbf{M}\,0.654\mathbf{L})^4$ *glass*

where *glass*, **H**, **M** and **L** represent media of refractive indices of 1.52, 2.35, 1.65 and 1.38, respectively. The optical thickness of **H**, **M** and **L** is $\lambda_0/4$, *matched* at 45°. λ_0 is 600 nm.

5.4.3.5.3 Edge filter in convergent flux

An edge filter is often designed at normal incidence and then inserted into a convergent beam (see Fig. 1-17). Its angle shift (see §1.4.1) should be minimal. A substantial part of this angle shift may be attributed to the decrease of effective thickness of each layer (see §2.5.2) that occurs with increasing angle. Curves in Fig. 2-59 illustrate that the change of effective thickness vs angle is less for high index layers than low index. It follows that the stop band center shifts less with angle when a greater proportion of high index material is used in the basic period.

This is illustrated in Fig. 5-88, which shows the transmittance of the edge filters

air $(0.25\mathbf{H}\,1.5\mathbf{L}\,0.25\mathbf{H})^{12}$ *glass* and *air* $(0.75\mathbf{H}'\,0.5\mathbf{L}'\,0.75\mathbf{H}')^{12}$ *glass* .

The stack centers are adjusted so that the 50% transmittance points at normal incidence are nearly the same. At a 35° incidence angle, Fig. 5-88 shows that the latter of the two coatings manifests a smaller angle shift.

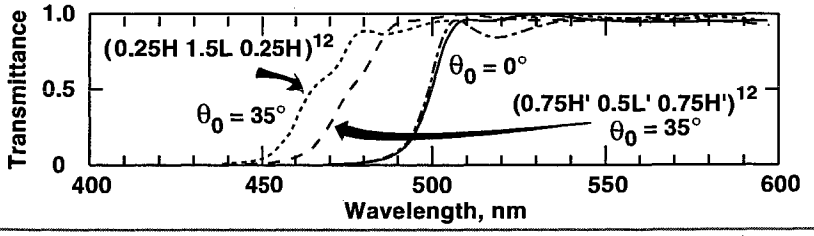


Fig. 5-88 $T(\lambda)$ at $\theta_0 = 0°$ incidence angle of the edge filter (*long-short dash*)

air $(0.25\mathbf{H}\,1.5\mathbf{L}\,0.25\mathbf{H})^{12}$ *glass*

and at $\theta_0 = 35°$ (*short dash*). At normal incidence $T(\lambda)$ of (*solid*)

air $(0.75\mathbf{H}'\,0.5\mathbf{L}'\,0.75\mathbf{H}')^{12}$ *glass*

and (*long dash*) at $\theta_0 = 35°$. The refractive indices of **H**, **H'**, **L**, **L'** and *glass* are 2.35, 2.35, 1.45, 1.45 and 1.52, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$ at λ_0 = 434 nm. **H'** and **L'** are QWOTs at λ_0 of 437 nm. At nonnormal incidence the light is randomly polarized.

5.4.3.5.4 Edge filter polarizer

This linear polarizer is usually deposited on a slab with plane parallel sides. It transmits *p* polarized flux and reflects the *s* polarization over a quite narrow band of wavelengths (see §1.8.3.2). The coating is an edge filter in which the cuton is at a shorter wavelength for the *p* polarization than for the *s* [Buchman 71a, 71b]. The polarization is strong in the gap between the *p* and *s* stop bands, as shown in Fig. 5-89.

A starting design consists of a *core stack* of 21 alternate layers of evaporated silica and titania

air AR $(\frac{1}{2}\mathbf{H}\,\mathbf{L}\,\frac{1}{2}\mathbf{H})^{10}$ AR *glass* .

The antireflection coatings enhance T_p in the passband. The number of periods is a compromise. A consequence of using fewer than 10 is that the *p* transmittance does not rise sharply enough at the edge of the stop band and T_p is not high enough at λ_0. The ripple at the stop band edge decreases T_p at λ_0 if as many as 14 periods are used. Figure 5-89 depicts $T(\lambda)$ — T_s is below 0.1% at λ_0.

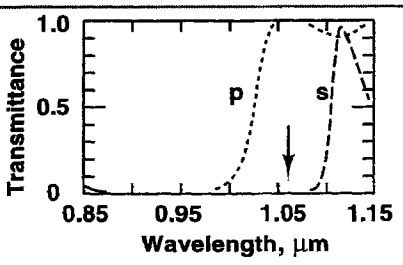


Fig. 5-89 $T(\lambda)$ an incidence angle of 56.4° for *p* polarized flux (*short dash*) and *s* polarized flux (*long dash*) for the linear polarizer

air 1.25**H'** 0.48**L'** 1.25**H'**
(0.465**L** 0.931**H** 0.465**L**)^{10}
1.25**H'** 0.48**L'** 1.25**H'** *glass*

where the refractive indices of **L**, **L'**, **H**, **H'** and *glass* are 1.45, 1.45, 2.16, 2.16 and 1.508, respectively. **L**, **H** and **L'**, **H'** represent layers of optical thickness $\lambda_0 / 4$ at λ_0 = 1.00 μm and 1.06 μm, respectively. All thicknesses are *matched* at 56.4°.

The angle of incidence of 56.4° is Brewster's angle for the substrate. The effective index N_0 = 1.81 of the incident medium equals that of the substrate [see Eq. (2-39e)]. Table 5-12 lists the equivalent index (for *p* polarized flux) at several wavelengths. The effective index N_1 of the antireflection coating is [see Eq. (2-23j)]

$$N_1 = \sqrt{n_e\, 1.81} \qquad (5\text{-}15c)$$

and should have a *matched* phase thickness of 90° or 270°.

The procedure of §2.11.3 is used to synthesize an equivalent layer with a phase thickness of 270° and an index of 2.84 for *p* polarization. Figure 5-89 graphs the transmittance of the final 27 layer filter.

5.4.3.6 Absorption in the passband

The most favorable situation that can occur is that the net absorption *A* of the filter in its passband spectral region is

$$A = \sum_{j=1}^{m} \alpha_j\, h_j \qquad (5\text{-}15d)$$

Table 5-12 Indices of core stack and matching layer

λ	n_e	N_1	n'_e	β'_e
1.04 μm	6.04	3.31	2.85	275.3°
1.06 μm	4.45	2.84	2.84	270.0°
1.08 μm	3.80	2.62	2.83	264.9°
1.10 μm	3.43	2.49	2.83	260.0°

Equivalent index n_e (for p polarized flux incidence angle in air of 56.4°) vs λ of 0.465**L** 0.931**H** 0.465**L** . N_1 is obtained from Eq. (5-15c). Equivalent thickness (in units of phase) β'_e and equivalent index n'_e (for p polarization) of the coating 1.25**H'** 0.48**L'** 1.25**H'** . The symbols are captioned in Fig. 5-89.

where α_j and h_j are the bulk absorption coefficient (see §2.2.4.4) and total metric thicknesses, respectively, of the jth material in the stack. m = 2, for example, for the coating in Fig. 5-89. This is the lowest attainable net absorption and occurs when the magnitude of the standing wave fields in the passband region is comparable to those of a single layer on a substrate. Higher standing wave fields occur close to the stop band edge which, in turn, dramatically increases the absorption in that spectral region.

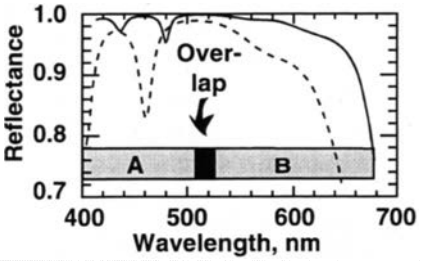


Fig. 5-90 $R(\lambda)$ of the broadband reflectors *air* stack *A* stack *B* *glass*
where stack *A* is (0.5**H** **L** 0.5**H**)^q
and stack *B* is (0.5**H'** **L'** 0.5**H'**)^q
for q = 4 (*dashed*) and q = 6 (*solid*). The refractive indices of **L**, **L'**, **H**, **H'** and *glass* are 1.46, 1.46, 2.30, 2.30 and 1.52, respectively. **L**, **H** and **L'**, **H'** represent layers of optical thickness $\lambda_0/4$ at λ_0 = 450 nm and 580 nm, respectively. Shaded areas are the stop bands of the respective stacks, with the overlap region blackened.

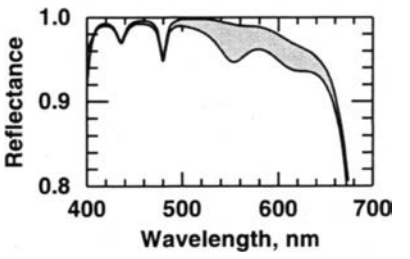


Fig. 5-91 $R(\lambda)$ of the broadband reflector
air (0.5**H** **L** 0.5**H**)^6
(0.5**H'** **L'** 0.5**H'**)^6 *glass*
where the optical thicknesses of **L**, **L'**, **H**, and **H'** are captioned in Fig. 5-90. (*Upper curve*) The refractive indices of the nonabsorbing materials **L**, **L'**, **H**, **H'** and *glass* are 1.46, 1.46, 2.30, 2.30 and 1.52, respectively. (*Lower curve*) The extinction coefficient of **H** and **H'** is 0.001. The shaded area represent the absorption

5.5 Broadband reflectors and rejection filters

This section focuses on coatings in which two or more stacks are deposited on the same substrate, similar to the coating in §5.4.3.4.3. The goal is to broaden the region of high reflectance or, alternatively, to extend the low transmittance spectral region, as in a blocking filter. Both goals require that the stop bands be located spectrally so that they

overlap. The principal design problems are the avoidance of transmittance leaks in blocking filters and the minimization of reflection loss (due to absorption) in mirrors.

5.5.1 Two-component stacks with overlapping stop bands

Figure 5-90 schematically shows the spectral location of the stop bands of coatings *A* and *B*. The overlap of the stop bands is deliberate. The conditions are found such that the reflectance in the overlap region increases monotonically as an additional period is added to both *A* and *B*.

The basic period of each stack is written as an equivalent layer. In the spectral region of the overlap, Turner [66] demonstrated that the magnitude of the elements of the characteristic matrix increases monotonically as more periods are added to both the *A* and *B* stacks, provided either of the following conditions is met:

(*a*) Both the *A* stack and the *B* stack have high index center layers; or,

(*b*) the center layers are low index in both stacks.

As an illustration, these coatings fulfill the foregoing criteria:

$$(\tfrac{1}{2}\mathrm{L}\,\mathrm{H}\tfrac{1}{2}\mathrm{L})^q\,(\tfrac{1}{2}\mathrm{L}'\mathrm{H}'\tfrac{1}{2}\mathrm{L}')^q$$

$$(\tfrac{1}{4}\mathrm{L}\tfrac{3}{2}\mathrm{H}\tfrac{1}{4}\mathrm{L})^q\,(\tfrac{2}{3}\mathrm{L}'\tfrac{2}{3}\mathrm{H}'\tfrac{2}{3}\mathrm{L}')^q$$

$$(\tfrac{3}{5}\mathrm{H}\tfrac{4}{5}\mathrm{L}\tfrac{3}{5}\mathrm{H}')^q\,(\tfrac{1}{3}\mathrm{H}'\tfrac{4}{3}\mathrm{L}'\tfrac{1}{3}\mathrm{H}')^q$$

where the optical thicknesses of the **L**, **H**, **L'** and **H'** components are chosen so that there is overlap of the stop bands. But the reflectance of

$$(\tfrac{1}{2}\mathrm{H}\,\mathrm{L}\tfrac{1}{2}\mathrm{H})^q\,(\tfrac{1}{2}\mathrm{L}'\mathrm{H}'\tfrac{1}{2}\mathrm{L}')^q$$

severely dips in the overlap region.

The dip may be explained by modeling the two stacks as reflectors in a "spacerless" Fabry-Perot interferometer (see §2.6.4.3). The transmittance is a maximum (and reflectance a minimum) when the mean phase shift upon reflection (see (Eq. (2-44h)] of the two reflectors is 180° or an even multiple thereof. The 180° condition occurs when the *A* and *B* stacks have different index center layers. This produces a dip in reflectance and is the basis of the phase conjugate interference filter (see §7.9.1). That dip is eliminated with the insertion of an extra layer **L"** between the stacks, as

$$(\tfrac{1}{2}\mathrm{H}\,\mathrm{L}\tfrac{1}{2}\mathrm{H})^q\,\mathrm{L}''\,(\tfrac{1}{2}\mathrm{L}'\mathrm{H}'\tfrac{1}{2}\mathrm{L}')^q\ .$$

L" is a QWOT at the wavelength of the stop band overlap.

Figure 5-90 depicts the reflectance of two stacks piled on top of each other with overlapping stop bands. The reflectance increases in the spectral region of stop band overlap (cross-hatched) as the number of periods increases from 4 to 6.

5.5.1.1 Broadband reflectors with nonabsorbing layers

It is surprising that it is more difficult to construct a reflector whose *R* fluctuates between 88% and 91% across the visible spectrum, than it is to achieve $R > 99\%$ over the same region. The reason is that notwithstanding the fluctuations of *R* in the stop band spectral regions, these eventually disappear (on a reflectance scale) as more periods are added.

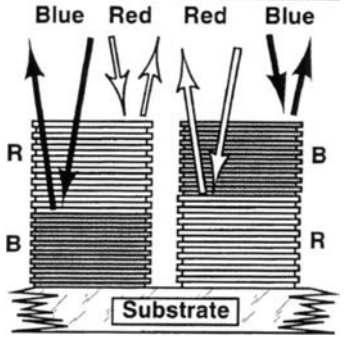


Fig. 5-92 In a broadband reflector, arrangement of a red-reflecting multilayer **R** and blue reflecting stack **B**. (*Left*) Red stack is on top.(*Right*) Blue stack is on top.

The shaded area in Fig. 5-90 shows that the stop band of stack *A* extends from slightly below 400 nm to 523 nm and the stop band of the *B* stack from 508 nm to 677 nm. This is a comfortable overlap. The *A* stack, which reflects in the short wavelength (blue) spectral region, is adjacent to the incident medium of air.

Figure 5-90 depicts its $R(\lambda)$ with four periods in each stack. The minima in *R* at 450 nm may be modeled using the concept of partitioning (see §2.6.4)—the interference between the *A* and *B* stacks. In other words, the reflectance peaks in the "*non*-stop-band" spectral region of stack *B* are interfering with stack *A* in its stop band. The dip in *R* weakens as more periods are added to both stacks. In the four period stack depicted in Fig. 5-90, the minimum *R* at 450 nm is 84%. Figure 5-90 shows that R_{min} progresses to 96% with six periods in each stack.

As another example, consider the reflectance at λ = 660 nm, which is within the stop band of stack *B*. With four periods, *R* = 60%, which is off scale in Fig. 5-90. But when six periods are used in both stacks, its *R* never drops below 94% in the spectral region 420 nm to 660 nm.

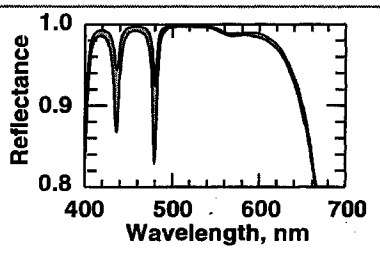


Fig. 5-93 $R(\lambda)$ of a broadband reflector

air (0.5H' L' 0.5H')^6 (0.5H L 0.5H)^6 *glass* where the optical thicknesses of **L**, **L'**, **H** and **H'** are captioned in Fig. 5-90. (*Upper curve*) The refractive indices of the nonabsorbing materials **L**, **L'**, **H**, **H'** and *glass* are 1.46, 1.46, 2.30, 2.30 and 1.52, respectively. (*Lower curve*) The extinction coefficient of **H** and **H'** is 0.001.

5.5.1.2 Broadband reflectors with absorbing layers

The assumption in §5.5.1.1 that the layers are nonabsorbing is unrealistic. The extinction coefficient of each high index layer is modeled as 0.001 in §5.5.1.2. Figure 5-91 compares the reflectance of the nonabsorbing stack of Fig. 5-90 with the absorbing. The quite noticeable dip in reflectance in the red is explained below.

Figure 5-92 shows two configurations of the cascaded reflectors: (1) The blue reflector is on top; and (2) the red reflector is on top. The red and blue rays that impinge upon the multilayers are also shown. In both coatings, either the blue or red light must penetrate an entire outer stack before it can reflect from the reflector on the bottom. The light is absorbed as it does so.

Figure 5-93 depicts the reflectance when the red reflector is adjacent to air. The reflectance dips to as low as 60% in the blue. The addition of more periods to both the red and blue reflectors worsens the reflectance in the blue. The problem is that the flux at 450 nm reflecting from the blue reflector must traverse all 14 layers of the red reflector

before it even impinges upon the top layer of the blue reflector. In essence, high standing wave electric fields are established that induce considerable absorption.

Figure 5-91 shows the reflectance of the alternative configuration — the blue reflector is on top. In contrast to Fig. 5-93, the drop in reflectance is now in the red spectral region. This is because the red flux must penetrate the blue stack before it can reflect from the red reflector.

The absorption in the **H** layers is responsible for degrading the reflectance and is consistent with the standing waves mentioned in §5.3.10.3. The foregoing examples demonstrate that it is usually preferable to locate the blue reflector on top — i.e., the configuration of Fig. 5-91. There are several reasons: (1) The total metric thickness of the layers in the short wavelength reflector is less than that of the long wavelength — hence thinner layers have less absorption. (2) The absorption of most coating materials is greater at shorter wavelengths than at longer.

The broadband reflectors consisting of two stacks and the "one-stack" coatings of §5.3.10 are quite different. Figure 5-66 shows that if too many layers are added to a "one- stack" coating, the worst that can happen is that the reflectance saturates at the Koppelmann limit and money is wasted depositing layers that in no way increase the reflectance. In "two-stack" designs, however, the residual absorption causes the reflectance to actually degrade if too many periods are used.

If, however, the purpose is to reduce the transmittance, then more and more periods may be added with impunity to either the "one-stack" or the "two-stack" coatings.

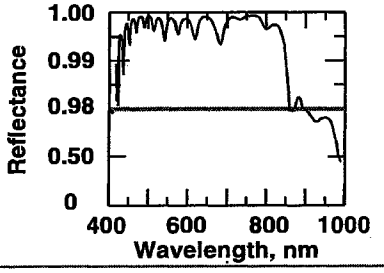


Fig. 5-94 $R(\lambda)$ of a "reflector with progressive layer thicknesses" containing 45 alternate layers of refractive indices 1.45 and 2.35. Its design appears in §5.8.5.2. The ordinate changes scale at 0.98.

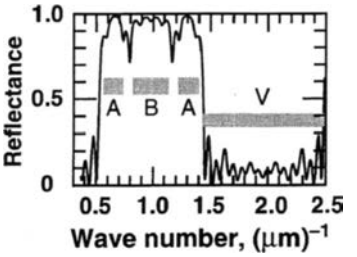


Fig. 5-95 $R(\sigma)$ of the heat reflector

air $(\mathbf{HL})^4\,\mathbf{H}\,(\mathbf{LLH})^5$ *glass*

where the refractive indices of **H**, **L** and *glass* are 2.30, 1.38 and 1.52, respectively. The optical thickness of **H** and **L** is $\lambda_0 / 4$ at $\lambda_0 = 1.02$ μm. The shaded areas *A* and *B* are the stop bands of **LLH** and **HL**, respectively. *V* is the visible spectral region.

Fig. 5-95 $R(\sigma)$ of the heat reflector

air $(\mathbf{HL})^4\,\mathbf{H}\,(\mathbf{LLH})^5$ *glass*

where the refractive indices of **H**, **L** and *glass* are 2.30, 1.38 and 1.52, respectively. The optical thickness of **H** and **L** is $\lambda_0 / 4$ at $\lambda_0 = 1.02$ μm. The shaded areas *A* and *B* are the stop bands of **LLH** and **HL**, respectively. *V* is the visible spectral region.

5.5.1.3 Reflectors with layers of progressively increasing thicknesses

The foregoing broadband reflectors were designed by depositing say, thirteen layers at a relatively long wavelength and then surmounting that stack with one centered at a shorter wavelength. The net effect is that the layer optical thicknesses abruptly change at a point in the stack. Another approach is to shift the layer thicknesses in more or less a continuous manner [Heavens 66].

An example is a stack in which the thicknesses change in a geometric progression:

air $(\mathbf{H}\,\mathbf{L})^q$ 1.04**H L** 1.08**H L** 1.13**H L** 1.17**H L**...

In other words, each period is 4% thicker than the previous one. The exception is the outermost period, which contains three or four periods—enough to build up the reflectance to its "starting value." Figure 5-94 shows that such a broadband reflector can achieve a substantial reflectance over a 2:1 bandwidth. This curve, however, is computed with no absorption in the layers. Similar to the coating captioned in Fig. 5-93, its reflectance would be degraded by absorption.

Heavens [66] also considers coatings in which the layer thicknesses change in an arithmetic, rather than geometric, progression.

5.5.2 Hot mirror without overlapping stop bands

There are heat reflectors in which the stop bands may not overlap. Figure 5-95 shows the reflectance of

air $(\mathbf{H}\,\mathbf{L})^4\,\mathbf{H}\,(\mathbf{L}\,\mathbf{L}\,\mathbf{H})^5$ *glass*

which consists of a quarterwave stack nested between the stop bands of a 2:1 stack. As Turner [66] points out, the stop bands are not quite wide enough to overlap.

The reflectance drops sharply in a gap between. The first-order stop band center of the 2:1 stack occurs at 1.47 μm. Its third-order reflectance peak, which would occur at 1.47/3 = 0.49 μm, is absent (see §5.2.5.2). The fourth-order reflectance peak almost crowds into the visible spectral region. Turner [66] shows that the reflectance dips disappear when the stop bands are widened by the use of higher index coating materials. This coating has also been used with lower refractive index materials [Martin 87].

5.5.3 Broadband IR rejection filter

The mission of this filter is to attenuate over a quite broad spectral region—its rejection bandwidth spans a 6:1 range of wavelengths. An example is a bandpass that transmits at least 90% average in the FLIR band—7.85 μm to 11.25 μm. A shortwave pass is coated on one side of the zinc selenide substrate and the longwave pass on the other. The discussion here focuses on the longwave pass. Its average transmittance below 7.22 μm should be less than 1%. The coating contains zinc sulfide ($n_L = 2.2$) and germanium ($n_H = 4.2$).

The starting design consists of five stacks—each with a different passband center. The layout of the coating is

air **A B C D E** *ZnSe*

where **A, B, C, D** and **E** each represents a stack, rather than a layer. Each stack is an equivalent layer

$(\frac{1}{2}\mathbf{H}\,\mathbf{L}\,\frac{1}{2}\mathbf{H})^q$.

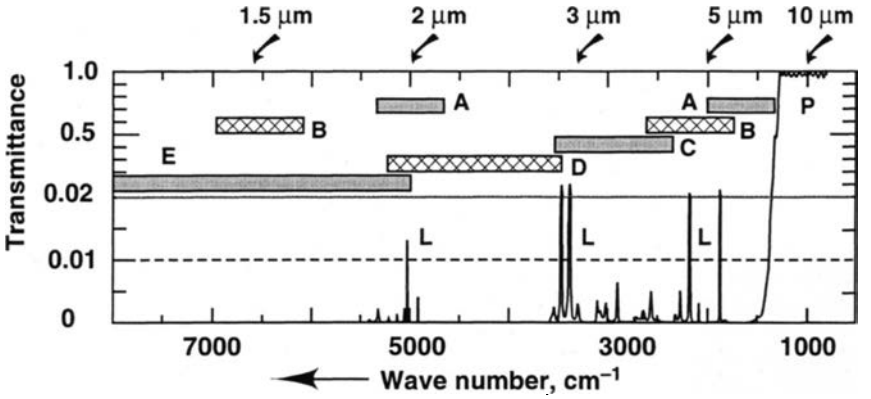


Fig. 5-96 Transmittance vs wave number of the longwave pass edge filter
air 10.3L 12.9H 2.17L 7.12H 5.94L 5.02H 6.84L 5.69H 6.0L 6.0H 6.0L 6.0H 6.15L 5.17H 5.11L 9.58H 1.49L 8.38H 16.0L 4.65L 4.10H 4.3L 5.23H 4.71L 5.0H 4.78L 5.5H 5.04L 3.95H 4.62L 3.62H 3.38L 3.38H 3.38L 2.78H 3.22L 3.22H 3.22L 4.67H 3.66L 3.66H 3.66L 2.37H 3.14L 1.68H 2.35L 2.59H 2.43L 2.32H 2.32L 2.32H 2.32L 2.0H 2.05L 2.4H 2.31L 2.31H 2.31L 3.28H 2.28L 2.28H 2.28L 2.11H 1.52L 1.5H 2.31L 1.13H 1.52L 1.55H 1.50L 1.04H 4.77L 1.32H 10.77L *ZnSe*
where the refractive indices of H, L and *ZnSe* are 4.2, 2.20 and 2.44, respectively. The extinction coefficient of H increases from zero at 5000 cm^{-1} to 0.1 at 8000 cm^{-1}. L and H represent layers of optical thickness $\lambda_0/4$ at $\lambda_0 = 1.0$ μm. The shaded or cross-hatched rectangles represent the stop bands of the stacks listed in Table 5-12. The ordinate changes scale at 0.02. *P* designates the passband and *L* leaks in the transmittance in the rejection region. The wavelength equivalents to wave number are shown. The dashed line is the 0.1% transmittance level.

Figure 5-96 depicts the spectral regions covered by the stop bands of each stack. The A stack determines the bandpass edge near 7.55 μm. It contains 19 layers and its stop band center wavelength λ_c is 6.0 μm. This coating was optimized separately so that the cuton wavelength was correct and the passband transmittance was high—using air as both the substrate and incident medium.

The stop band of the B stack ($\lambda_c = 4.6$ μm) is at a shorter wavelength and overlaps the stop band of the A stack. Figure 5-96 shows that the C stack ($\lambda_c = 3.4$ μm) rejects in the spectral region near 3 μm and the D stack ($\lambda_c = 2.3$ μm) attenuates at wavelengths as short as 2 μm. The rejection of the E stack ($\lambda_c = 1.6$ μm), which has the thinnest layers, extends to wavelengths as short as 1.25 μm. The absorption in the germanium layers enhances the attenuation at wavelengths shorter than 1.5 μm.

Figure 5-96 illustrates that the third-order stop bands of the A and B stacks overlap with the stop bands of the D and E stacks in the short wavelength part of the spectrum. At worst, such overlap produces a single-cavity bandpass filter so narrow in bandwidth that little energy is transmitted though it. This narrowness is attributed to the thick "spacer," which is the entire thickness of the C stack.

It is possible that the order of the layers in the foregoing multilayer could be reversed and the thinnest—the E stack—deposited next to air. The problem is that its equivalent index in the passband region (*P* in Fig. 5-96) is approximately 3.0. Its index mismatch to the index of air is quite severe.

The **B**, **C**, **D** and **E** stacks each contained six periods in the starting design, which totaled 83 layers. An optimization computer code adjusted the thicknesses of the outer layers and the layers near the boundaries of the different stacks. The thicknesses of the outer six layers of the **A** and **B** stacks would be adjusted, for example, to optimize the passband transmittance and also to suppress any transmission leaks in the stop band overlap region near 2000 cm^{-1}.

Figure 5-96 depicts the narrow transmittance leaks L that occur in the rejection region. The dashed line is at 1% transmittance level. These leaks are acceptable, as long as the specification for the filter is an average transmittance. It is evident that a narrowband laser could penetrate the coating in the "leaks."

5.5.4 Design techniques

An optimization computer program is an invaluable design tool, as witness the results of §5.5.3 and the high transmittance shown in Fig. 1-53. The fact that the loss is linear with the extinction coefficient k is useful in reflector design. This is valid, at least, when $|k|<<1$. The k values may be made a decade larger than they really are to avoid rounding errors in the computation of R on the computer. The absorption in a broadband laser mirror may be minimized by optimizing R and T simultaneously.

Examples of broadband reflectors abound in the literature. Berthold [72] patented a broadband reflector in which the thicknesses of **H L** pairs grew progressively thicker or thinner.

5.6 Phase shift upon reflection

5.6.1 Introduction

The analysis of the periodic structures in previous sections focuses upon radiant reflectance and transmittance. The phase shift upon reflection is important when multilayers are used in bandpass filters, interferometers, liquid crystal displays and laser cavities.

5.6.2 Quarterwave stacks and similar periodic media

Figure 5-97 plots the magnitude of the reflection coefficient and the phase shift upon reflection of a quarterwave stack. The phase has a constant slope in the stop band spectral region and oscillates rapidly outside of that region. An appendix (see §10.6.3.1) uses dispersion theory to show that the slope of the phase shift at the stop band center is

$$\frac{d\delta}{d\sigma} = \frac{constant}{2\Delta g} \tag{5-16a}$$

where $2\Delta g$ is the width of the stop band [see Eq. (5-5c)]. The foregoing is true in the limit as the number of periods becomes infinite. It applies to any periodic structure — not necessarily a quarterwave stack.

Equation (5-16a) shows that the slope is inversely proportional to the stop band spectral width. This is illustrated by the stack with the broad stop band in Fig. 5-19. It manifests a relatively small slope, whereas the three-quarterwave stack (*shaded* in Fig. 5-23) has a greater slope. This is important in the design of a multiple-cavity bandpass, which consists of halfwave spacers surrounded by reflectors. §7.2.3 shows that the bandwidth of the bandpass is influenced by the slope of the phase shift of the aforementioned reflectors.

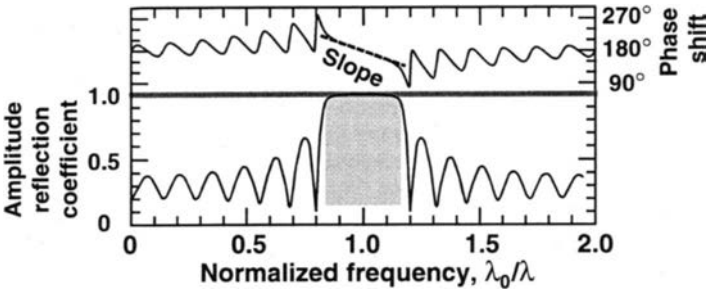


Fig. 5-97 Phase shift upon reflection (*upper panel*) and magnitude of the amplitude reflection coefficient (*lower panel*) of the quarterwave stack

air **H** (**L H**)^6 *glass*

where the refractive indices of *glass*, **L** and **H** are 1.52, 1.45 and 2.32, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$. The slope of the phase shift near the center frequency is shown, as well as the stop band (*shaded*).

The slope equation of Abelès [58] is derived from the characteristic matrix (see §2.4.4). The phase thickness of each layer is 90° or an odd multiple thereof. The films alternate between a high index n_H and a low index n_L. When the outer layer (contiguous to the incident medium of index n_0) is high index, the derivative (i.e., slope) of phase shift upon reflection is

$$\frac{d\delta}{d\sigma} \approx \frac{-\pi\, m\, \lambda_0\, n_0}{n_H - n_L}\left(1 + \frac{1}{n_H}\frac{dn_H}{\lambda_0\, d\sigma}\right) \tag{5-16b}$$

where m (which is odd) is the order number of the stop band (see §5.2.3.4). This slope is depicted in Fig. 5-97.

The foregoing applies in the limit as the number of periods becomes infinite. The second term in the brackets represents the dispersion of the high index material. It is assumed that the dispersion of the low index component may be neglected. The dispersion can be significant—increasing $\frac{d\delta}{d\sigma}$ by as much as 10% above that of a nondispersive coating. When the low index component is at the top of the stack and the dispersion of both components is neglected,

$$\frac{d\delta}{d\sigma} \approx \frac{-\pi\, m\, \lambda_0\, n_0^{-1}}{n_L^{-1} - n_H^{-1}} \,. \tag{5-16c}$$

Suppose that a quite large number of periods (rather than six) is used in the three-quarterwave stack captioned in Fig. 5-40. Its slope is

$$\frac{d\delta}{d\sigma} = \frac{-\pi \times 3 \times (0.633\ \mu\text{m}) \times 1.0}{1.63 - 1.45} = -33.1 \text{ radian } \mu\text{m} = -1900 \text{ degree } \mu\text{m} \tag{5-16d}$$

when the outer layer is **H** and

$$\frac{d\delta}{d\sigma} = \frac{-\pi \times 3 \times (0.633\ \mu\text{m}) \times (1.0)^{-1}}{(1.45)^{-1} - (1.63)^{-1}} = -78.3 \text{ radian } \mu\text{m} = -4488 \text{ degree } \mu\text{m} \tag{5-16e}$$

when **L** is on top. The latter is about 2.3 times larger in magnitude than the former. French [86] reviews the influence of the phase dispersion of the stack in distorting the temporal shape of dye laser pulses of femtosecond duration.

5.6.3 Broadband reflectors

A characteristic of the phase shift curves for these reflectors is a quite rapid change in certain spectral regions, as illustrated in Fig. 5-98. Two of the curves are for the coatings

air (**blue**) (**red**) *glass* and *air* (**red**) (**blue**) *glass*

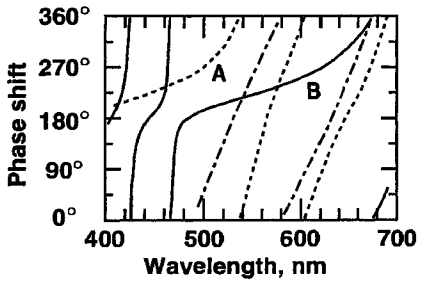


Fig. 5-98 Phase shift upon reflection vs wavelength of the broadband reflector (*solid*)

air $(0.5\mathbf{H}\,\mathbf{L}\,0.5\mathbf{H})^6$ $(0.5\mathbf{H}'\,\mathbf{L}'\,0.5\mathbf{H}')^6$ *glass*, (*short dash*)

air $(0.5\mathbf{H}'\,\mathbf{L}'\,0.5\mathbf{H}')^6$ $(0.5\mathbf{H}\,\mathbf{L}\,0.5\mathbf{H})^6$ *glass* and (*long-short dash*)

air $12\mathbf{F}'$ $(0.5\mathbf{H}\,\mathbf{L}\,0.5\mathbf{H})^6$ *glass*

where the indices of **H**, **H'**, **L**, **L'** and **F'** are 2.30, 2.30, 1.46, 1.46 and 1.00, respectively. The optical thickness of each layer is $\lambda_0/4$, with $\lambda_0 = 563$ nm for the **H** and **L** layers. λ_0 is 450 nm for the **F'**, **H'** and **L'** layers.

where (**blue**) and (**red**) represent symbolically the components of the reflectors in the captions to Figs. 5-90 and 5-92. That is, **red** and **blue** are quarterwave stacks "tuned" to reflect in those spectral regions.

At wavelengths near 400 nm, the phase (*short-dashed* curve) of the "blue on top" coating is close to that of the quarterwave stack in Fig. 5-97—remembering that the sign of the slope is reversed because of the use of wavelength (rather than wave number) on the abscissa. At wavelengths near 520 nm, however, the blue reflector is relatively transparent and the light is reflecting from the red reflector underneath the blue. The blue reflector functions as a "spacer" layer of considerable thickness.

The *long-dashed* curve in Fig. 5-98 shows the effect of this spacer layer. Depicted is the phase shift of

air $12\mathbf{F}'$ (**red**) *glass*

where the fictitious **F** layer is composed of air and has the same optical thickness as that of the (**blue**) stack, namely 12 QWOTs. At a particular wavelength λ, the phase shift δ_f of this stack is given by a modification of Eq. (2-22i)

$$\delta_f = \delta_0 - 4\times\pi\times 3\times 450/\lambda = \delta_0 - 4\times\pi\times 3\times 450\,\sigma \qquad (5\text{-}16f)$$

where δ_0 is the phase shift at the outer surface of the (**red**) stack. There is a sign reversal of the second term in Eqs. (2-21i) and (5-16f) because the phase reference surface is moved *away*, not *towards*, the substrate. The 3×450 in Eq. (5-16f) represents the optical thickness (12 QWOTs = 3 waves) of the **F** "spacer layer."

The next step is to differentiate Eq. (5-16f) with respect to σ. After this is done, the second term in that equation makes a much larger contribution to slope than the contribution from Eq. (5-16b) for a quarterwave stack.

The converse of the foregoing arguments hold for the "red on top" coating. The behavior of the phase in region *B* in Fig. 5-98 is quite similar to that of a quarterwave stack. Due to the fact that the red reflector reflects most of the light in the red, the (**blue**) stack on the bottom has little influence on the reflectance in that spectral region. In the

blue part of the spectrum, however, the (**red**) stack is relatively transparent and functions like a thick spacer layer. The phase varies rapidly with wavelength in this spectral region.

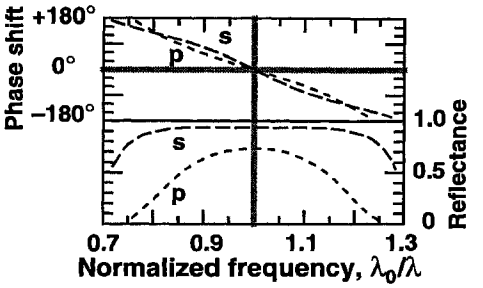


Fig. 5-99 The quarterwave stack design is *air* $(\mathbf{L}\,\mathbf{H})^4$ *glass*

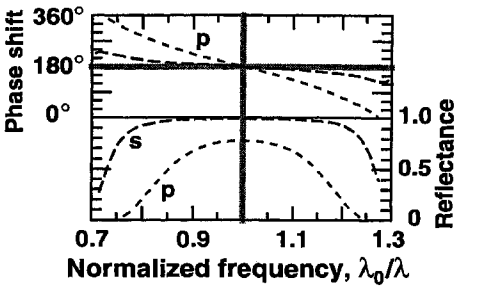


Fig. 5-100 The quarterwave stack design is

air $\mathbf{H}(\mathbf{L}\,\mathbf{H})^3$ *glass* .

At 60° angle of incidence, reflectance (*lower*) and phase shift upon reflection (*upper*) vs λ_0/λ for *p* polarized (*solid*) and *s* polarized light (*dashed*). The refractive indices of *glass*, **L** and **H** are 1.52, 1.45 and 2.35, respectively. **H** and **L** represent layers of optical thickness $\lambda_0/4$, *matched* at 60°.

§2.2.5.3 examines the analogy between a multilayer and an electrical transmission line. Suppose that a two-terminal network of inductors and capacitors is encased in a "black box." A length of transmission line, such as coaxial cable, connects the "black box" to an instrument that measures impedance at rf frequencies. At a given frequency, there is a change of phase in the signal between the connector and the instrument. In other words, the measured impedance must be corrected for the phase shift of the transmission line. If, for example, the electrical length of the cable is many wavelengths, then the phase of the return signal varies quite rapidly with frequency—similar to the *long-dash* curve in Fig. 5-98. The analogy is that the "air" film on top of the stack is analogous to the lengthy transmission line connecting the "black box" impedance to the instrument.

5.6.4 Phase shift due to thickness changes

As mentioned in §2.3.2, the reflection of a plane wave from a coated surface is from a pseudo-surface that is determined by the node of the standing wave. This is discussed in §10.7.3.

5.6.5 Phase shift upon reflection—nonnormal incidence

Figures 5-99 and 5-100 depict the reflectance and phase shift upon reflection at 60° incidence angle for both polarizations of quarterwave stacks. In Fig. 5-99, there is a **L** layer adjacent to air, where in Fig. 5-100 there an **H** layer on both ends. Inspection of these Figs. shows that the *p* reflectance of both stacks is nearly the same. This is attributed to the fact that the 60° incidence angle is close to 55°, which is Brewster's angle for the low index material of index 1.45. The consequence is that the Fresnel reflection coefficient at the air-film interface is quite small (see §2.5.3.1).

Inspection of Fig. 5-99 shows that the slope of the phase shift upon reflection for *p* polarization is slightly less than the slope for *s* polarization. Figure 5-100 shows that when an **H** layer is next to air, the opposite is true. The slope of the phase shift for the *p* polarization exceeds that of the *s*. Because Eqs. (5-16b) and (5-16c) are derived from the elements of the characteristic matrix (see §2.4.4), effective indices (see §2.5.3) may be substituted into them at nonnormal incidence. As an illustration, the effective indices of *air*, **L** and **H** are 2.00, 1.81 and 2.53, respectively, for the TM polarization. Assuming a stack center at 1.00 μm and first-order stop band, the magnitude of the slope is

$$\frac{d\delta}{d\sigma} = \frac{\pi \times (0.633\,\mu\text{m}) \times 2.0}{2.53 - 1.81} = 8.7 \text{ radian } \mu\text{m} \ . \qquad (5\text{-}17a)$$

The slope for *s* is 1.54 radian μm — about 5.5 times less than the *p* slope.

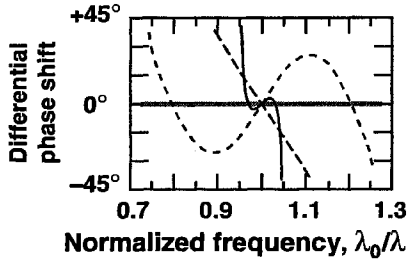


Fig. 5-101 At 60° angle of incidence, differential phase shift upon reflection [see Eq. (2-20f)] vs λ_0/λ of the quarterwave stack (*short dash*)

air $(\mathbf{L}\,\mathbf{H})^4$ *glass,*

(*long dash*)

air $\mathbf{H}(\mathbf{L}\,\mathbf{H})^3$ *glass*

and (*solid*)

air $(\mathbf{L}\,\mathbf{M})^{32}$ *glass*

where the refractive indices of *glass*, **L**, **M** and **H** are 1.52, 1.45, 1.80 and 2.35, respectively. The optical thickness of **H**, **M** and **L** is $\lambda_0/4$, *matched* at 60°.

5.6.5.1 Differential phase shift upon reflection

There are optical systems in which the differential phase shift of the reflected beam must be controlled, as, for example, a laser cavity or interferometer. Figure 5-100 shows the phase splitting of the *p* and *s* components and Fig. 5-101 plots the differential phase shift Δ [see Eq. (2-20f)].

The Δ of a quarterwave stack is precisely zero at the passband center wavelength λ_0. That is, the coatings

air $(\mathbf{H}\,\mathbf{L})^q$ *substrate*

and *air* $\mathbf{L}(\mathbf{H}\,\mathbf{L})^q$ *substrate*

both achieve a zero differential phase shift at λ_0. In practice, however, a small centering error in the deposition of the stack introduces a nonzero differential phase. The slope, as measured from Fig. 5-101, is 3.4° per percent centering for the stack with an **L** layer adjacent to air. A 2% shift of the stack center (due to thickness errors during deposition) introduces a differential phase shift of 6.8°.

One solution — which is admittedly an act of desperation — is to use a lower refractive index for the **H** layer. As illustrated by the solid curve in Fig. 5-101, the use of a 1.80 index produces a spectral region of a few percent in which the Δ is relatively flat.

The index of the "zero slope Δ" coating is found by substituting effective indices into Eq. (5-16c) and equating the *p* and *s* slopes. Given the angle of incidence θ_0 and the angles of refraction θ_H and θ_L in the **H** and **L** layers, this leads to

$$\frac{\sec\theta_0}{n_L^{-1}\sec\theta_L - n_H^{-1}\sec\theta_H} = \frac{\cos\theta_0}{n_L^{-1}\cos\theta_L - n_H^{-1}\cos\theta_H} \ . \qquad (5\text{-}17b)$$

Figure 5-102 shows the index of the **H** layer that satisfies the foregoing equation vs angle of incidence. There are practical solutions at angles less than 30°.

5.6.5.1.1 Differential phase shift coatings

There are coatings that produce a prescribed differential phase shift, rather than the zero shift of §5.6.5.1. Figure 5-103 shows several simple coatings that produce nonzero differential phase shifts. A thickening of the top **L** layer of the stack by 15% produces a Δ of 34° and a 30% increase drives Δ to 50°.

5.6.5.1.2 Ninety degree differential phase shift coating

A widely used multilayer mirror has a 90° differential phase shift and equal reflectance for *p* and *s* polarized flux. Such coatings are often developed with optimization computer codes [Southwall 79, 80]. The goal in this section is to gain insight.

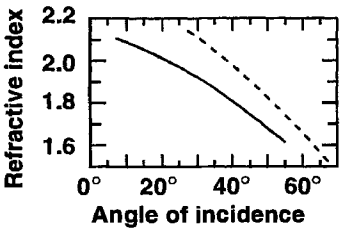


Fig. 5-102 Versus angle of incidence in air, refractive index of the **H** layer in a quarterwave stack that satisfies Eq. (5-17b) for an **L** layer of indices 1.45 (*solid*) and 1.38 (*dashed*). There is an **L** layer adjacent to air. The optical thickness of each layer is *matched* at each angle of incidence.

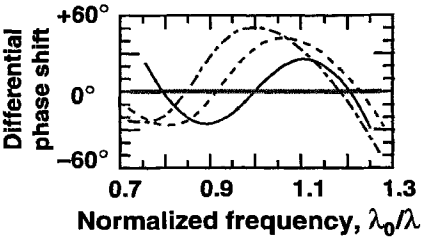


Fig. 5-103 At 60° angle of incidence, differential phase shift upon reflection [see Eq. (2-20f)] vs λ_0/λ of the coating (*long-short dash*)

air 1.3**L H** (**L H**)^2 *glass,*

(*short dash*)

air 1.15**L H** (**L H**)^2 *glass*

and (*solid*)

air (**L H**)^3 *glass* .

The symbols are captioned in Fig. 5-101.

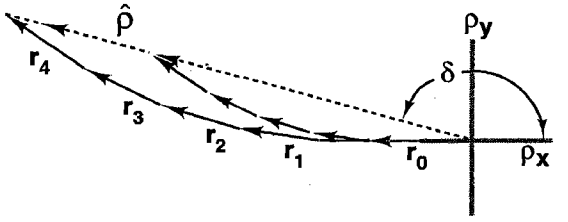


Fig. 5-104 At $\lambda = 0.95\ \lambda_0$ and at 45° incidence angle in air, vector addition of amplitudes computation of the amplitude reflection coefficient $\hat{\rho}$ (*short dash*, *s* polarization, *short-long dash*, *p* polarization) for the coating captioned in Fig. 5-4 . ρ_x and ρ_y are the Cartesian coordinates of $\hat{\rho}$. The optical thicknesses are matched at 45°. The phase shift upon reflection δ is shown.

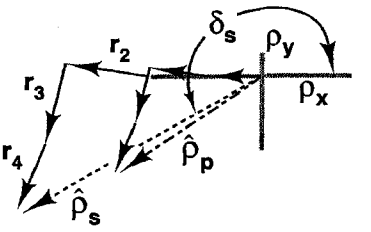


Fig. 5-105 The caption to Fig. 5-104 obtains, with the exceptions that the incidence angle is 60° and the coating is

air **H** 0.5**L H** *substrate*

where the indices of **H**, **L** and *substrate* are 4.20, 1.35 and 1.50, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 10.6 μm, *matched* at 60°.

Fig. 5-104 shows a vector diagram at 45° incidence angle of the coating in Fig. 5-4 . The optical thicknesses of the layers are matched at that angle (see §2.7.2.2). The

phase retardance β of each layer is precisely the same for *p* and *s*—only the lengths of the vectors are different. The differential phase shift is virtually nil. The vectors, however, point the way to a design technique, which is to obtain the maximum *p-s* phase splitting by adding a layer of eighth-wave optical thickness so that the vectors add in quadrature. An illustration is in Fig. 5-105, which depicts the vector diagram of

air **H** 0.5**L H** *substrate*

where the **H** and **L** layers represent germanium and a rare earth fluoride. The differential phase shift is produced by the vectors that add in quadrature.

The design first step is to deposit enough quarterwave layers so that the required reflectance is attained. Fractional quarterwave layers are now added that produce the differential phase shift. At this point, the coating is

high **L'** $(\mathbf{HL})^5$ *substrate*

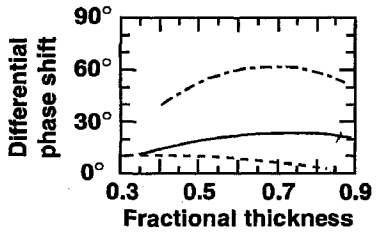


Fig. 5-106 At a wavelength λ_0 and angle of incidence of 45° in air, differential phase shift upon reflection Δ vs thickness of outer layer (in fractions of a QWOT) of (*short dashed*)

high **L'** $(\mathbf{HL})^5$ *substrate* , (*solid*)

low **H** 0.4**L** $(\mathbf{HL})^5$ *substrate* and (*long-short dash*)

low **H** 0.8**L** 0.8**H** 0.4**L** $(\mathbf{HL})^5$ *substrate* where the refractive indices of *high*, **H**, *low*, **L**, **L'** and *substrate* are 2.22, 2.22, 1.40, 1.40, 1.40 and 2.40, respectively.

where the index of *high* is a massive medium of the same index as an **H** layer. This is merely a way of finding the *p-s* phase splitting of the *internal* amplitude reflection coefficient (see §10.1.2.1) vs the optical thickness (as fractions of a QWOT) of the **L'** layer. This is the *dashed* curve in Fig. 5-106. A thickness of 0.4 QWOT—slightly less than an eighth wave—produces a Δ of 10°. The next layer **H'** is now added, resulting in the coating

low **H'** 0.4**L** $(\mathbf{HL})^5$ *substrate*

where the pseudo-incident medium *low* has the same index as an **L** layer. Figure 5-106 (*solid* curve *A*) shows that the differential phase now exceeds 20° at an optical thickness of 0.8 QWOT. The foregoing process is continued for an **L'** layer. Figure 5-106 shows that the *p-s* phase splitting of

low **H'** 0.8**L** 0.8**H** 0.4**L** $(\mathbf{HL})^5$ *substrate*

attains nearly 60°.

Additional layers are added until the differential phase is close to 90°. It may be necessary to add more periods to enhance the reflectance. The final multilayer is

air 0.615**H** 0.7**L** 0.8**H** 0.8**L** 0.8**H** 0.4**L** $(\mathbf{HL})^9$ *substrate* .

Several aspects of this coating are that: (1) It is not unique—there are many other similar ways of producing it; (2) the 90° phase retardance is quite sensitive to thickness errors in the outer layers of the stack; and (3) there are quite large standing wave electric fields in the outer layers of the stack—the stack is quite absorbing if there is any absorption in either the **L** or **H** layers.

5.7 Miscellaneous topics

5.7.1 Overcoated metallic reflector

§1.7.5 discusses several uses of these coatings, which consist of dielectric layers on an opaque metal. They are seldom used as laser mirrors for several reasons: (1) deposition temperatures; and (2) contamination.

Metals such as aluminum and silver must be evaporated upon substrates at relatively low temperatures. This is incompatible with dielectric oxide coating materials such as silica, titania and tantala that should be deposited at temperatures in excess of 200 °C to minimize their absorption.

There is always a concern that during the deposition process flakes of metal will contaminate the sources of the dielectric coating materials and thus cause absorption.

5.7.1.1 Reflectors with minimal number of layers

This coating usually contains only two or four layers and the intent is to reflect over a relatively narrow range of wavelengths. Two options are available: (1) the **HL** periods of QWs are added directly on top of the metal; or (2) the layer adjacent to the metal is adjusted in thickness so that the combination of **L**' metal produces a pure real admittance. In the latter case, a slightly higher reflectance is obtained, although the gain may be in the third decimal place.

Table 5-13 Reflectance of overcoated metals

Design	Optimum?	Reflectance	R uncoated	$\hat{Y}$
air **H L** *nickel*	No	0.8023	0.5555	
0.796**L** *nickel*			0.5555	$0.265 + j0$
air **H** 0.796**L** *nickel*	Yes	0.8183	0.5555	
0.662**H** *nickel*			0.5555	$0.621 + j0$
1.662**H** *nickel*			0.5555	$8.52 + j0$
air **H L** 1.662**H** *nickel*	Yes	0.8445	0.5555	
air **H L** *aluminum*	No	0.9679	0.9151	
air **H** 0.866**L** *aluminum*	Yes	0.9692	0.9151	
air **H L** 1.782**H** *aluminum*	Yes	0.9714	0.9151	

Reflectance at $\lambda = \lambda_0$ of uncoated metals *nickel* ($\hat{n} = 2 - j3$) and *aluminum* ($\hat{n} = 0.93 - j6.33$), as well as dielectric layers on those metals. The normalized admittance $\hat{Y}$ of several partial designs is listed. The optical thickness of **H** and **L** is $\lambda_0/4$ and their refractive indices are 2.30 and 1.38, respectively. *Optimum* means that the thicknesses are adjusted to obtain the maximum reflectance at λ_0.

The first example uses nickel as the metal. Consider first the use of an **L** layer next to the metal. *Ab intuito*, this produces a larger Fresnel reflection coefficient at the dielectric-metal interface than when an **H** layer is used. Equation (2-46a) is used to find the optical thickness of the **L** layer, which is 0.797 QWOT. Table 5-13 shows that the normalized admittance of

0.796**L** *nickel*

is pure real and a fraction, namely 0.265. The foregoing value is obtained from Eq. (2-46b). The stack 0.796**L** *nickel* may be regarded as a pseudo-substrate. It follows that the addition of an **H** layer to such a low index pseudo-substrate produces a high reflectance. Table 5-13 shows that the coating

air **H** 0.796**L** *nickel*

increases the peak reflectance from 55% to 82%.

The foregoing procedure is repeated for the case where the **H** layer is contiguous to the metal. Table 5-13 lists its normalized admittance as 0.62. The addition of an **HL** pair to the single **H** layer results in

HL0.662**H** *nickel*

that at λ_0 has a (pure real) normalized admittance of

$$Y = \frac{(2.30)^2\ 0.62}{(1.38)^2} = 1.72 + j\,0 \qquad (5\text{-}18a)$$

and thus acts like an antireflection coating, rather than a reflectance enhancement coating.

The solution is to add the three-layer stack **HLH**. From Eq. (2-48c) the normalized admittance is

$$Y = \frac{(2.30)^4}{(1.38)^2\ 0.62} = 23.7 + j\,0\ . \qquad (5\text{-}18b)$$

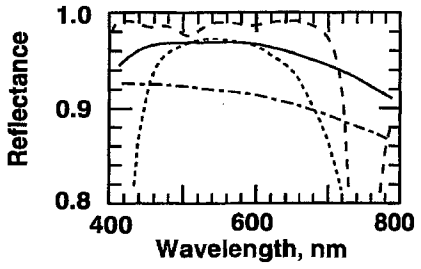


Fig. 5-107 $R(\lambda)$ of uncoated aluminum (*short-long dash*), the overcoated metal (*solid*)
air **H**' 0.866**L**' *metal* ,
(*short dash*)
air **H**' **L**' 1.782**H**' *metal*
and (*long dash*)
air 0.381**L** 1.969**H** 1.135**L** 0.225**H** 0.662**L** 0.9677**H** 0.5186**L** 0.5129**H** 0.5111**L** 0.511**H** 0.433**L** *metal*
where the indices of **H**, **H**', **L** and **L**' are 2.30,.2.30,.1.38 and 1.38, respectively. The optical thickness of each layer is $\lambda_0/4$, with λ_0= 1.00 for the **H** and **L** layers. λ_0 is 546 nm for the **H**' and **L**' layers. The optical constants of the *metal* (aluminum) are from Hass [72].

Table 5-13 shows that the reflectance is enhanced from 55% (for the uncoated substrate) to nearly 85%. The three-layer **HLH**' stack has a peak reflectance only 3% higher than the two layer **HL**' coating.

Figure 5-107 compares the reflectance of overcoated metals with that of the uncoated aluminum substrate, the n and k of which are dispersive. The reflectance of **HLH**' (i.e., **H** layer next to the metal) drops below that of the opaque aluminum at wavelengths 20% longer than the stack center. This Fig. also shows that the two-layer **HL**' overcoat has superior bandwidth as compared to **HLH**'.

5.7.1.2 Broadband reflector

Figure 5-107 also depicts the reflectance of an eleven-layer overcoat that was spewed out by the optimization process (see §10.2). A Coursen type of starting design (see §10.2.4.5) was used. I have no idea if it is possible to enhance the reflectance still further or to extend the bandwidth beyond the 400 nm to 700 nm *desideratum* given to the optimization computer code.

5.7.1.3 Cold mirror on metallic substrate

A cold mirror is often deposited upon a metallic substrate, such as electroformed nickel or pressed aluminum. Its installation is similar to that shown in Fig. 1-60, except that the IR is absorbed, rather than transmitted by the mirror.

The first step is to design a *dark mirror* (see §8.5) that covers both the visible spectrum and near IR to at least 2.5 μm. The substrate is overcoated with titanium metal until it is opaque. It is then overcoated with layers of titanium interdispersed between

dielectric films. Beginning with a Coursen starting design (see §10.2.4.5), it is quite surprising that the optimization computer code utilized four dielectric and two metallic layers to reduce the reflectance over a 5:1 bandwidth, as shown in Fig. 5-108.

The design of the cold mirror, which contains alternating layers of silica and titania, follows the procedure of §5.5.1. The starting design consists of a "blue stack on top" (QWs at 470 nm—13 layers) and a "red stack on bottom" (QWs at 600 nm —15 layers) deposited on top of the dark mirror. The stacks have overlapping stop bands. An optimization computer code was unleashed to reduce the reflectance in the IR and to optimize the visible. The computer strove to reduce the thicknesses of the metallic layers to 7 nm. Since such ultrathin layers would not manifest bulk optical constants, the thicknesses were constrained to exceed 11 nm. Figure 5-108 depicts the final result.

Table 5-14 Reflector that includes an antireflection coating

		Simulation		
ID	Index	Thickness **H**	Thickness **L**	Thickness **H**
High index	2.30	None	None	None
1	2.16	0.454**H**	0.084**L**	0.454**H**
2	1.87	0.343**H**	0.287**L**	0.343**H**
3	1.62	0.215**H**	0.539**L**	0.215**H**

A reflector with a spectrally flat spectral response, in which layers 1, 2 and 3 of a Chebyshev antireflection coating are interposed between the high index layer and substrate (index 1.52). Each of those three layers is simulated with an equivalent layer.
The indices of **H** and **L** are 2.30 and 1.45, respectively.
The optical thicknesses are $\lambda_0/4$ at λ_0 of 533 nm.

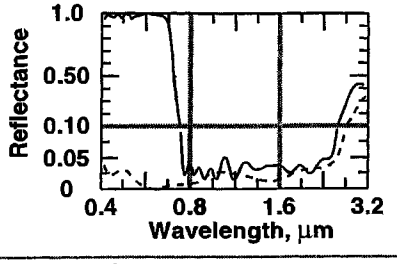


Fig. 5-108 $R(\lambda)$ of the six layer dark mirror (*dashed*) and a 32 layer coating (*solid*) consisting of a dark mirror overcoated with a cold mirror. The designs appear in §5.8.5.3. The ordinate changes scale at 0.1 and the abscissa at 0.8 μm and 1.6 μm.

5.7.2 Broadband reflector with relatively low reflectance

The objective is to produce a reflectance of close to 10% to 15% in the visible part of the spectrum. A reflectance of 30% to 35% is possible in the IR, where higher refractive index coating materials are available. §6.1.1.3 describes a method of beamdivider design in which a three-layer Chebyshev antireflection coating is "inverted." The "inversion" of the antireflection coating produces a normalized admittance in the range of 4.00 to 5.00, which produces a reflectance of close to 40%.

There is a different approach, one similar to the edge filter design in §5.4.3.4.1 [Gordon 83]. The passband transmittance of the edge filter is enhanced by producing an antireflection coating for each *interface* of the *core stack.* The patent of Gordon [83]

references earlier publications in which this coating produces a reflector with neutral color.

The design procedure is: (1) Choose a layer whose refractive index is in the range 2.00 to 2.30. This is termed the *high index layer* and is the outer layer of the stack. (2) Considering the *high index layer* as a massive medium, design an antireflection coating that matches it to the substrate. In other words, step (2) matches the *bottom interface* of the high index layer to the substrate. The Fresnel reflection coefficient of the upper interface of the *high index layer* produces a reflectance that is spectrally "flat."

An example is a coating on a BK-7 substrate containing materials of index 2.30 and 1.45. The *high index layer* (n = 2.30) is the outer layer. The procedure of §4.7.2 is used to design a three-layer Chebyshev antireflection coating that matches the 2.30 to the substrate of index 1.52. A 2:1 bandwidth—400 nm to 800 nm—is spanned. Table 5-14 shows that the indices of sections of the antireflection coating are 2.16, 1.87 and 1.62. Table 5-14 also lists the simulations of each of those indices with an equivalent layer of 90° phase thickness at the band center wavelength of 533 nm, as obtained from Eq. (4-10). Figure 5-109 shows that the reflectance is relatively flat over the aforementioned bandwidth, even when the optical thickness of the *high index layer* is changed from a halfwave to a quarterwave.

5.7.3 IR reflector with visible transmittance

An example of such a coating is a stack of layers of a rare earth fluoride **L** and zinc sulfide **H** that reflects in the IR at 8 μm, yet transmits 633 nm. The design method is quite similar to that of §4.9.1. The coating is

air VisAR1 (**H'** VisAR2 **L'** VisAR2)^3 **H'** VisAR2 **L'** VisAR3 *substrate*

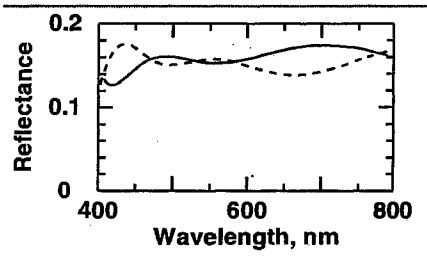


Fig. 5-109 $R(\lambda)$ of the broadband reflector (*dashed*)
air 2.454**H**0.084**L**0.797**H**0.287**L** 0.558**H**0.538**L**0.215**H** *glass*
and (*solid*)
air 1.454**H**0.084**L**0.797**H**0.287**L** 0.558**H**0.538**L**0.215**H** *glass*
where *glass*, **H** and **L** represent media of refractive index 1.52, 2.30 and 1.45, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$ at λ_0 of 533 nm.

where the optical thicknesses of the **H'** and **L'** layers are QWs near 8 μm. **H'** and **L'** are termed the IR-reflecting layers. The coatings VisAR1 and VisAR3 perform the same functions as they do in Fig. 4-46. They function as an antireflection at 633 nm that matches the outer surface of the outermost **H'** layer to air and the bottom **L'** layer to the bulk zinc selenide substrate, respectively. VisAR2 performs the same function at the interface between each **H'** and **L'** layer. The argument is that if *all* interfaces of the IR-reflecting layers are surrounded with antireflection coatings at 633 nm, the transmittance of the entire assembly must be close to 100% at that wavelength, *independent* of the thicknesses of the IR-reflecting layers.

The simplest type of antireflection coating is the two-layer unequal thickness (see §4.3.4). Following the procedure of §4.3.4.2.2, for example, a two-layer stack

high index 0.32**L**0.33**H** *low index*

has zero reflectance at 633 nm, provided **H** and **L** are QWs at 633 nm. The massive media *high index* and *low index* have the same refractive index as a **H** and **L** layer, respectively. VisAR1 consists of a single **L** layer.

Figure 5-110 shows the reflectance in the IR and the visible. The optical thicknesses of the IR-reflecting layers must be reduced slightly to compensate for the optical thickness increase of the VisAR2 coating. Rancourt [80] lists a coating that has a broader bandwidth because the antireflection coating contains four layers, rather than the two layers used above.

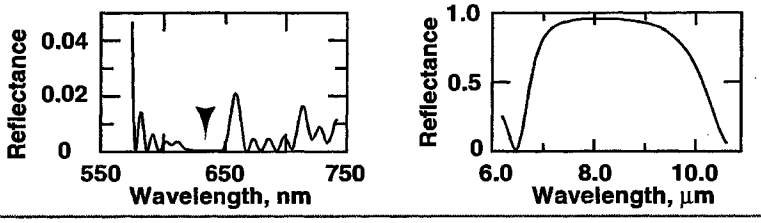


Fig. 5-110 $R(\lambda)$ of an IR reflector that transmits 633 nm

air **L** (**H'** 0.32**L** 0.33**H L'** 0.33**H** 0.32**L**)^3 **H'** 0.32**L** 0.33**H L'** 0.42**H** 0.28**L** *ZnSe*

where the optical thickness of **L'** and **H'** is $\lambda_0/4$ at λ_0 of 7.7 μm. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 633 nm. The refractive indices of **H'**, **L'**, **H** and **L** are 2.22, 1.40, 2.30 and 1.50, respectively. The index of *ZnSe* is 2.56 in the visible and 2.40 in the IR. The arrow designates 633 nm.

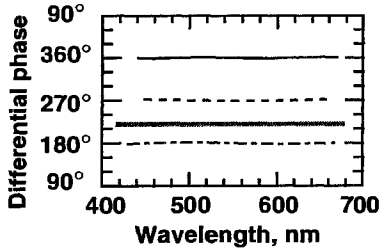


Fig. 5-111 At an angle of 45°, differential phase shift upon reflection (modulo 360° using the Abelès sign convention) of (*short dash*)

glass 0.132**H** 0.305**L** 0.672**H** *air* ,

(*long-short dash*)

glass 0.166**H** 0.315**L** 0.473**H** 0.722**L** *air*

(*solid*)

glass 0.15**H** 0.344**L** 0.347**H** 0.318**L** 0.338**H** 0.221**L** 0.939**H** 0.617**L** 0.402**H** *air*

and (*shaded* horizontal line)

glass air

where the refractive indices of glass, **H** and **L** are 1.517, 2.30 and 1.45, respectively. The optical thickness (which is *not* matched) of **L** and **H** is $\lambda_0/4$ at λ_0 of 1000 nm.

5.7.4 TIR differential phase shift coatings

The coatings of this section are not periodic structures and need not enhance the reflectance, since it is already 100% because of the total internal reflection TIR. The goal is to produce a "flat" differential phase shift upon reflection (see §2.3.4) across the visible portion of the spectrum. Figures 5-111 and 2-29 show that the differential phase shift of the uncoated prism is 220°. It is not surprising that since 270° and 180° are relatively close to 220°, only three or four layers are needed to produce the "flat" differential phase shifts shown in Fig. 5-111. Nine layers are used to produce a zero differential phase.

A Coursen type of starting design (see §10.2.4.5) was used in each instance to produce the aforementioned multilayers. For example, for the 0° differential phase it is

glass (0.1**H** 0.1**L**)^{14} 0.1**H** *air*

where **H** and **L** are captioned in Fig. 5-111. The fact that the optimization computer code leaves thin layers in the design is a source of

frustration. I have attempted to drop these thin layers and also to thicken them — to no avail. I was unable to design a coating with a differential phase shift of 90° ± 3° over the visible spectral region.

5.7.5 Protective coatings

A topic hitherto neglected in Chap. 5 is the addition of layers to the outer surface of the mirror to protect it. An example are the mirrors in gas lasers sealed onto the ends of the plasma tube. These mirrors are bombarded with ions from the discharge and also the intense UV from the plasma. The term "plasma stable" refers to mirrors that do not degrade under such conditions and reduce laser power.

Holmes [91] has shown, for example, that tantala layers are *not* plasma stable. He conjectures that the UV flux in the plasma produces electron traps that eventually increase the radiant absorption at 633 nm. A standard prescription in the industry is to overcoat a mirror for a He-Ne gas laser with a halfwave of fused silica. This increases its plasma stability.

5.8 Appendices

These appendices contain mathematical details that would clutter the main text and, perhaps, even obscure the physical models.

5.8.1 Analysis of a periodic structure using matrices

The first step is to develop a matrix for the basic period— either the Kard matrix [see Eq. (2-29h)] or the characteristic matrix [see Eq. (2-31c)]. This matrix is represented by $\mathbf{M}$. The goal is to compute the exponentiated matrix $\mathbf{M}^q$.

5.8.1.1 Cayley-Hamilton theorem

This theorem is derived in standard texts (see Frazer [63]). The first step is to find the eigenvalues G of the 2X2 characteristic matrix, which are the roots of a quadratic

$$G^2 - xG + D = 0 \ . \tag{5-19a}$$

D is the determinant [see Eq. (2-40g)] of the matrix and x is the trace [see Eq. (2-41e)].

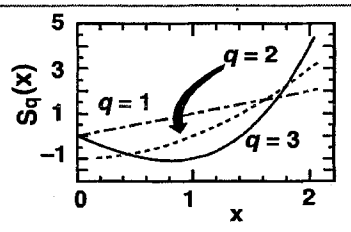


Fig. 5-112 Chebyshev polynomials $S_q(x)$ of the second kind and order q for q = 1, 2 and 3.

D is unity for both the Kard and the characteristic matrices. The Cayley-Hamilton theorem states that $\mathbf{M}$ may be substituted for G in Eq. (5-19a)

$$\mathbf{M}^2 - x\mathbf{M} + \mathbf{I} = 0 \tag{5-19b}$$

where $\mathbf{I}$ is the unit matrix (see §2.6.2.2). In other words, a matrix satisfies its own eigenvalue equation.

5.8.1.2 Recursion relationships

There are several ways to find $\mathbf{M}^q$. The first is to use Eq. (5-19b) recursively. This is written

$$\mathbf{M}^2 = x\mathbf{M} - \mathbf{I} \tag{5-19c}$$

or, in other words, an expression for $\mathbf{M}^2$ in terms of $\mathbf{M}$ and $\mathbf{I}$ has been found. The foregoing is multiplied by $\mathbf{M}$ and $\mathbf{M}^2$ is eliminated. The results for $\mathbf{M}^3$ and $\mathbf{M}^4$ are

$$\mathbf{M}^3 = (x^2 - 1)\,\mathbf{M} - x\,\mathbf{I} \tag{5-19d}$$

$$\mathbf{M}^4 = (x^3 - 2x)\,\mathbf{M} - (x^2 - 1)\,\mathbf{I} \tag{5-19e}$$

or in general

$$\mathbf{M}^q = S_{q-1}(x)\mathbf{M} - S_{q-2}(x)\,\mathbf{I} \tag{5-19f}$$

Table 5-15 Chebyshev polynomials of the second kind of order q

q	$S_q(x)$	q	$S_q(x)$
0	1	5	$x^5 - 4x^3 + 3x$
1	x	6	$x^6 - 5x^4 + 6x^2 - 1$
2	$x^2 - 1$	7	$x^7 - 6x^5 + 10x^3 - 4x$
3	$x^3 - 2x$	8	$x^8 - 7x^6 + 15x^4 - 10x^2 + 1$
4	$x^4 - 3x^2 + 1$	9	$x^9 - 8x^7 + 21x^5 - 20x^3 + 5x$

where $S_q(x)$ is a Chebyshev polynomial of the second kind, not to be confused with the Chebyshev polynomial of the first kind used in the design of antireflection coatings (see Table 4-5) and bandpass filters (see §7.10.1). Listed below are several ways of generating these polynomials. Regardless of which method is used, Eq. (5-19f) is the basic equation used to iterate a matrix. Table 5-15 lists these polynomials.

5.8.1.3 Recursive generation of the polynomials

It may be shown by induction that the Chebyshev polynomials follow the recursion relationship

$$S_q(x) = x\,S_{q-1}(x) - S_{q-2}(x) \tag{5-20a}$$

which may be used to generate $S_q(x)$, given the two initial values

$$S_0(x) = \text{a constant} = 1 \tag{5-20b}$$

and

$$S_1(x) = x\ . \tag{5-20c}$$

$S_q(x)$ is graphed in Fig. 5-112 for $q = 1$, 2 and 3.

5.8.1.4 Absorbing layers

If the basic period represented by the matrix $\mathbf{M}$ contains one or more absorbing layers, then each element of $\mathbf{M}$ is complex [see Eq. (2-31c)], as is its trace $\hat{x}$. The recursion equation [see Eq. (5-19f)] is still valid, however.

5.8.1.5 Polynomial generation by circular and hyperbolic functions

Another form of the Chebyshev polynomial is to define

$$\phi \equiv \arccos(\tfrac{1}{2}x) \tag{5-21a}$$

for $|x| < 2$

$$\phi \equiv \operatorname{argcosh}(\tfrac{1}{2}x) \tag{5-21b}$$

for $|x| > 2$, where x appears in Eq. (2-41e). Then

$$S_q(x) = \frac{\sin[(q+1)\phi]}{\sin\phi} \tag{5-21c}$$

or

$$S_q(x) = \frac{\sinh[(q+1)\phi]}{\sinh\phi} \tag{5-21d}$$

depending upon whether the circular or hyperbolic expression for ϕ is used, respectively.

5.8.1.6 Exponentiation in terms of the eigenvalues

An equivalent form of the Cayley-Hamilton theorem is to express $\mathbf{M}^q$ in terms of the eigenvalues. The first step is to find the roots G_1 and G_2 of Eq. (5-19a)

$$G_1 = \tfrac{1}{2}\left(x + \sqrt{x^2 - 4}\right) \qquad \text{and} \qquad G_2 = \tfrac{1}{2}\left(x - \sqrt{x^2 - 4}\right) . \tag{5-22a}$$

Equation (5-19f) becomes

$$\mathbf{M}^q = \left[\left(G_1^q - G_2^q\right)\mathbf{M} - G_1 G_2\left(G_1^{q-1} - G_2^{q-1}\right)\mathbf{I}\right](G_1 - G_2)^{-1} . \tag{5-22b}$$

In the case where the matrices represent nonabsorbing layers, x is pure real and the eigenvalues are shown in Fig. 5-113. When $|x| < 2$, the eigenvalues lie on the unit circle and are complex conjugates. When $|x| > 2$, the eigenvalues are pure real and are reciprocals of each other. That is, given G_1 from Eq. (5-22a), $G_2 = 1/G_1$.

5.8.1.7 Special form for symmetrical matrices

Another formulation uses the Chebyshev polynomials of the first kind, as they appear in Table 4-5.

Starting with the identity

$$2\,\mathcal{T}_q(x/2) = S_q(x) - S_{q-2}(x) , \tag{5-23a}$$

Eq. (5-19f) becomes

$$\mathbf{M}^q = \begin{bmatrix} \mathcal{T}_q(x/2) + \frac{1}{2}(m_{11} - m_{22})\,S_{q-1}(x) & m_{12}\,S_{q-1}(x) \\ m_{21}\,S_{q-1}(x) & \mathcal{T}_q(x/2) - \frac{1}{2}(m_{11} - m_{22})S_{q-1}(x) \end{bmatrix} \tag{5-23b}$$

which is particularly useful when the matrix is symmetrical. In such instances, $m_{11} = m_{22}$ and the diagonal elements of Eq. (5-23b) simplify.

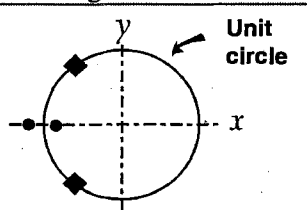


Fig. 5-113 On an unit circle plotted on an Argand diagram, nondegenerate eigenvalues of the characteristic matrix of the bilayer 1.3H0.7L at a phase thickness β of 90° (where a stop band exists—*circles*) and at $\beta = 60°$ (where a passband exists—*squares*). n_L is 1.38 and n_H is 2.30.

Any of the foregoing methods may be applied to matrices that represent either nonabsorbing or absorbing layers. In the latter case, x in Eq. (5-19f) is complex and the complex arithmetic becomes more cumbersome.

5.8.1.8 Tour-de-force multiplication of matrices

The distributive rule of matrix multiplication may be used, as, for example,

$$\mathbf{M}^{16} = (((\mathbf{M}^2)^2)^2)^2 \tag{5-24a}$$

and only four multiplications are needed to exponentiate the matrix, rather than 15.

The foregoing is most efficient when the exponent q is close to 2^m, where m is an integer. It may still be used in other instances, although not as efficiently. An example is a matrix iterated 15 times. Starting with the definitions

$$\mathbf{M}_2 \equiv \mathbf{M}^2 , \qquad \mathbf{M}_4 \equiv \mathbf{M}_2^2 \qquad \text{and} \qquad \mathbf{M}_8 \equiv \mathbf{M}_4^2 \tag{5-24b}$$

then

$$\mathbf{M}^{15} \equiv \mathbf{M}_8\,\mathbf{M}_4\,\mathbf{M}_2\,\mathbf{M} \tag{5-25a}$$

and six matrix multiplications are needed.

5.8.1.9 A comparison of methods

The particular method used to compute $\mathbf{M}^q$ depends upon the environment. For a desktop computer in which the co-processor may be accessed in assembler, there may exist hardware methods of finding the arccos—in which case the equations of §5.8.1.5 are attractive. In the computer code that the author used to prepare this book, the exponent q is reduced to two parts

$$q \equiv qa + qb \tag{5-26a}$$

where

$$qa \equiv q \bmod 16\ , \tag{5-26b}$$

$$qb = (q - qa)/16 \tag{5-26c}$$

and

$$\mathbf{M}^q = \mathbf{M}_a\,\mathbf{M}_b = \mathbf{M}^{qa}\left[\mathbf{M}^{16}\right]^{qb}\ . \tag{5-26d}$$

The recursion relationship [see Eq. (5-20a)] is used to compute the matrix $\mathbf{M}_a$ generated by the exponent qa. The procedure of §5.8.1.8 generates the matrix $\mathbf{M}_b$ created by the exponent qb. The final exponentiated matrix is the product of $\mathbf{M}_a$ and $\mathbf{M}_b$.

5.8.1.10 Stop band and passband

Equations (5-21a) and (5-21b) apply to the passband and stop band spectral regions, respectively. The hyperbolic functions in Eq. (5-21d) ensure that the matrix elements of $\mathbf{M}^q$ monotonically increase in magnitude with increasing q. The transmittance of a stack [see Eq. (2-35e)] depends upon the inverse square of the matrix elements. Hence T becomes vanishingly small in the stop band spectral region as more periods are added.

5.8.1.11 Nomenclature and history

If the results in §5.8.1.7 are compared with Born & Wolf [75], note that x [see Eq. (2-41e)] is defined differently. The book of Brillouin [56] reviews this subject in depth. Floquet's theorem and the Cayley-Hamilton theorem were developed in the nineteenth century. Rayleigh [17] found the reflectance of a periodic structure equivalent to a quarterwave stack. He developed a recursion formula utilizing hyperbolic functions that is equivalent to a Chebyshev polynomial.

5.8.2 Amplitude reflection coefficient of a periodic structure

The approach in §5.8.1 is to find the elements of the characteristic matrix and thence the transmittance from Eq. (2-34b). Jacobsson [75] developed an alternative approach that combines amplitude reflection coefficients and matrices.

A basic period consisting of two or more homogeneous layers—or, alternatively, combinations of homogeneous and inhomogeneous layers—is sandwiched between an incident medium (index n_0) and a substrate (index n_s). The elements of its characteristic matrix have been computed, so that x [in Eq. (2-41e)] is known. Its amplitude reflection coefficient $\hat{\rho}_1$ and transmission coefficient $\hat{\tau}_1$ are computed. When the basic period is repeated q times, $\hat{\rho}_q$ and $\hat{\tau}_q$ are

$$\frac{\hat{\rho}_q}{\hat{\tau}_q} = \frac{\hat{\rho}_1}{\hat{\tau}_1}\,S_{q-1}(x) - \frac{n_0 - n_s}{2\,n_0}\,S_{q-2}(x) \tag{5-27a}$$

S_q and x have the same meaning as in Eqs. (5-19f) and (2-41e).
In the case where $n_0 = n_s$, Eq. (5-27a) simplifies to

$$F_q \equiv \frac{R_q}{T_q} = \frac{R_1}{T_1} \left[S_{q-1}(x)\right]^2 = F_1 \left[S_{q-1}(x)\right]^2 \tag{5-27b}$$

where R_1, T_1, R_q and T_q are the radiant reflectances and transmittances of the single period and multiple periods, respectively.

5.8.2.1 Relationships valid at the stop band edge

At the stop band edge, by definition

$$|x| = 2 \tag{5-28a}$$

in which case $S_q(x)$ simplifies to

$$S_q(\pm 2) = (q+1)\,(\pm 1)^q. \tag{5-28b}$$

Equation (5-22b) is no longer usable because the eigenvalues are degenerate—that is, $G_1 = G_2$. Equation (5-28b) is used with the simple result

$$F_q = F_1\, q^2\ , \tag{5-28c}$$

which is the basis of Eq. (5-9a).

5.8.3 Reflectance and phase envelopes

For a given basic period, incident medium index and substrate index, it is possible to establish upper and lower limits for the radiant reflectance and phase shift upon reflection. These are often useful because the designer knows that as more layers are added to a stack, the reflectance must remain bounded between these limits. Although a graphical method has been used to construct the reflectance envelope [Dufour 54], the mathematical treatment here is more general [Arndt 66, Mouchart 81].

It is assumed that a nonabsorbing iterated stack is deposited on a substrate whose optical constant is $\hat{n} = n_s - j\,k_s$. Starting with the matrix elements of the basic period [using Eq. (2-32g)], the following auxiliary parameters are defined:

$$\xi \equiv (m_{11} - m_{22})/(2\sin\phi) \tag{5-29a}$$

$$d_{12} \equiv m_{12}/\sin\phi \tag{5-29b}$$

$$d_{21} \equiv m_{21}/\sin\phi \tag{5-29c}$$

where ϕ is defined in Eq. (5-21a). The envelope is obtained by

$$\frac{\partial R}{\partial q} = 0 \tag{5-29d}$$

and solving for q. This q is then used to compute $S_q(x)$ in Eq. (5-21c) and thence the reflectance. Two values of q are found: q_1 and q_2. One provides the upper limit of the envelope and the other the lower limit. These are computed using the arctangent of

$$\tan(2\,q_1\,\phi) = \tan(2\,q_2\,\phi + \pi) = A/B \tag{5-29e}$$

where

$$A \equiv 2\left[\xi\,(n_s^2 - n_0^2) - |k_s|\,(n_0^2\,d_{12} - d_{21}) + k_s^2\,\xi\right], \tag{5-30a}$$

$$B \equiv (n_s^2 + n_0^2)\,(\xi^2 - 1) + n_0^2\,n_s^2\,d_{12}^2 + d_{21}^2 + 2\,\xi\,|k_s|\,(n_0^2\,d_{12} + d_{21}) + k_s^2\,(n_0^2\,d_{12}^2 + \xi^2 - 1)\ . \tag{5-30b}$$

Reflectance envelopes appear in Figs. 5-17 and 5-18.

5.8.4 Missing stop bands — solution for the refractive indices

The design **A x y B y x A** in Table 5-5 is chosen as an example. Its basic period contains seven layers. According to §5.3.5.1.3, there are four "index to be found layers." Ergo, this coating can suppress the reflectance of as many as *four* contiguous higher order stop bands.

The objective is to produce zero reflectance for the basic period, which is composed of nonabsorbing layers. The constraints are that it is symmetrical and that the incident medium and substrate have the same refractive index. According to Eq. (5-27a), the reflectance of q iterated periods is zero if the reflectance of the basic period is zero. This zero reflectance is *independent* of the refractive index of the incident medium and substrate.

5.8.4.1 Development of the linear equations

The starting point is the Kard R/T polynomial for seven layers of equal optical thickness (see §2.7.4.3). Taking the square root of Eq. (2-51d) results in

$$\sqrt{R/T} = a_1\, \mathcal{S}\, C^6 - a_3\, \mathcal{S}^3\, C^4 + a_5\, \mathcal{S}^5\, C^2 - a_7\, \mathcal{S}^7 \qquad (5\text{-}31a)$$

where $\mathcal{S}$ and C—the sine and cosine of the phase of retardation β — appear in Eq. (2-50a). The stack is symmetrical and hence $u_1 = u_7$, $u_2 = u_6$ and $u_3 = u_5$. The symmetry — plus the fact that $n_0 = n_s$— causes the B_1 term in Eq.(2-51d) to vanish. Each of the a coefficients in Eq. (5-31a) is a function of the log of the refractive indices. Table 2-15 cites approximate values for them. For a given phase thickness β_3, Eq. (5-31a) contains four variables — u_1, u_2, u_3 and u_4.

$$\underbrace{-60u_1 + 20\mathbf{u_2} + 4\mathbf{u_3} - 6u_4}_{a_1 \cos^6(\beta_3)} - \underbrace{a_3 \sin^2(\beta_3) \cos^4(\beta_3)}_{-4u_1 + 4\mathbf{u_2} - 4\mathbf{u_3} + 2u_4} + \underbrace{a_5 \sin^4(\beta_3) \cos^2(\beta_3)}_{-60u_1 - 20\mathbf{u_2} + 4\mathbf{u_3} + 6u_4} - \underbrace{a_7 \sin^6(\beta_3)}_{-4u_1 - 4\mathbf{u_2} - 4\mathbf{u_3} - 2u_4} = 0$$

Fig. 5-114 Equation (5-31a), with a factor of sinβ ($\equiv \mathcal{S}$) removed, is equated to zero and evaluated for a symmetrical stack in which the phase thickness of each layer is β_3. The symbols $\boldsymbol{u_2}$ and $\boldsymbol{u_3}$ (in boldface) are variables. u_1 and u_4 are constants.

The first step is to constrain the refractive indices of the inner layer and the outer layers of the seven layer stack. The result is that u_1 (= u_7) and u_4 are constants in Eq. (5-31a). Fig. 5-114 shows approximate values for each of the coefficients in Eq. (5-31a) in terms of the log index. The two unknowns are in boldface type and the constrained u values are in ordinary type.

The procedure is to write Eq. (5-31a) for two phase thicknesses β_3 and β_2, which are the values at which the reflectance is to be zero. The next step is to collect the constant terms shown in Fig. 5-114 on the right of each equation. The result is two linear equations, with u_2 and u_3 as variables. They are solved and the resulting u values are converted into refractive index using Eq. (2-53b). In rare instances, the linear equations are singular and there is no solution for the indices. The numerical values for β_3 and β_2 are discussed in §5.8.4.2.

An alternative approach to the procedure in the foregoing paragraph is to impose the condition of *degeneracy* and require that $u_2 = u_3$. The equation in Fig. 5-114 is then solved for the remaining index at one phase angle β_3, as discussed in [Baumeister 92b].

5.8.4.2 Selection of the phase thickness values

This section discusses the selection of the phase thickness values β at which the reflectance of the basic period is to be zero—as, for example, the β_3 and β_2 in the previous section. If m is the order number of the stop band at which the reflectance is to be suppressed, β_m is

$$\beta_m = 180\, m\, /\ell \qquad (5\text{-}31b)$$

where ℓ is the number of *quarterwave* layers in the basic period. ℓ is *odd* in this example and ℓ is *even* when the center layer is a halfwave—as discussed in §5.3.5.1.1.

An example is the seven layer stack cited in §5.8.4.1. It suppress the reflectances at stop bands of orders 2, 3, 4 and 5. Due to the symmetry, however, the reflectance is always suppressed at *pairs* of stop bands symmetrical about order number 3.5. That is, both $m = 3$ and $m = 4$ will be suppressed. By the same line of reasoning, the reflectance at order numbers $m = 2$ and $m = 5$ is suppressed.

For the suppression of the third-order stop band ($m = 3$), Eq. (5-31b) provides $\beta_3 = 3\times180°/7 = 77.143°$. By the same line of reasoning, β_2 is $\beta_2 = 2\times180°/7 = 51.43°$.

5.8.4.3 Five layers

The square root of the R/T polynomial for five layers is

$$\sqrt{R/T} = a_1\, S\, C^4 - a_3\, S^3\, C^2 + a_5\, S^5 \ . \qquad (5\text{-}31c)$$

a_1, a_3 and a_5 are obtained from Table 2-15. The suppression of the reflectance at the second and third-order stop bands requires that the foregoing is zero at $\beta = 72°$. Collecting terms provides

$$u_1(-4\,C^4 + 24\,C^2S^2 - 4S^4) + u_1(-2\,C^4 - 4\,C^2S^2 - 2S^4) \approx u_2\, 4\,(C^4 - S^4) \qquad (5\text{-}31d)$$

which results in

$$-1.236\, u_1 - 2\, u_3 \approx -3.2361\, u_2 \ . \qquad (5\text{-}31e)$$

Because the refractive indices of the outer and inner layers are constrained, u_1 and u_3 are constants and u_2 is the unknown. Equation (5-9a) results when refractive indices [see Eq. (2-29c)] replace the u variables. Solution of Eq. (5-31e) provides the constant 0.382—the first entry of the "a" column of Table 5-5.

5.8.4.4 Refractive indices of a quasi-rugate reflector

Figure 5-48 plots the reflectance vs frequency of a stack that suppress the reflectances of stop bands with order numbers $2 \le m \le 22$. The design of the basic period is **A p q r t u v w x y z BB z y x w v u t r q p A** and the refractive indices are

$$n_p \approx \exp[0.966\ln[n_a) + 0.034\ln[n_b)\,] \qquad [5\text{-}32a)$$
$$n_q \approx \exp[0.900\ln(n_a) + 0.10\ln(n_b)\,] \qquad (5\text{-}32b)$$
$$n_r \approx \exp[0.807\ln(n_a) + 0.193\ln(n_b)\,] \qquad (5\text{-}32c)$$
$$n_t \approx \exp[0.693\ln(n_a) + 0.307\ln(n_b)\,] \qquad (5\text{-}32d)$$
$$n_u \approx \exp[0.566\ln(n_a) + 0.434\ln(n_b)\,] \qquad (5\text{-}32e)$$
$$n_z = n_a\ n_b\ /\ n_p \qquad (5\text{-}32f)$$
$$n_y = n_a\ n_b\ /\ n_q \qquad (5\text{-}32g)$$
$$n_x = n_a\ n_b\ /\ n_r \qquad (5\text{-}32h)$$
$$n_w = n_a\ n_b\ /\ n_t \qquad (5\text{-}32i)$$

$$n_v = n_a \; n_b / n_u \tag{5-32j}$$

where n_a and n_b are the refractive indices of the A and B layers, respectively. The indices of the layers in Fig. 5-48 are 1.450, 1.473, 1.518, 1.585, 1.671, 1.772, 1.883, 1.996, 2.104, 2.196, 2.264 and 2.300 for n_a, n_p, n_q, n_r, n_t, n_u, n_v, n_w, n_x, n_y, n_z and n_b, respectively.

5.8.5 Coating designs

5.8.5.1 Design of a rejection filter

The coating design in Fig. 5-64 is

air 0.621L 0.895H 0.546L 0.385H 0.323L 0.419H 0.844L 0.369H 0.344L 0.486H 0.555L 0.486H
0.462L 0.486H 0.498L (486H 0.486L)^9 0.486H 0.493L 0.486H 0.473L 0.486H 0.481L
0.486H 0.566L 0.486H 0.335L 0.486H 0.813L 0.366H 0.358L 0.211H 0.824L 0.449H *glass*

where L and H represent layers of optical thickness of $\lambda_0/4$ at λ_0 of 1 μm. The indices of L and *glass* are 1.45 and 1.52, respectively. The index of H is given by Eq. (2-2c) with $A = 3.4$ and $B = 0.0327$.

5.8.5.2 Reflector with layers of progressive thicknesses

The design is:

air 0.500H L H L H L 1.020H 1.040L 1.065H 1.090L 1.110H 1.130L 1.150H 1.170L 1.195H
1.220L 1.245H 1.270L 1.295H 1.320L 1.345H 1.370L 1.395H 1.420L 1.450H 1.480L
1.510H 1.540L 1.570H 1.600L 1.635H 1.670L 1.700H 1.730L 1.765H 1.800L 1.835H
1.870L 1.910H 1.950L 1.990H 2.030L 2.070H 2.110L 1.055H *substrate* .

The refractive indices of L, *substrate* and H are 1.45, 1.52 and 2.35, respectively. L and H represent layers of optical thickness $\lambda_0/4$ at λ_0 of 430 nm. Its $R(\lambda)$ appears in Fig. 5-94.

5.8.5.3 Dark mirror and overcoated dark mirror

The dark mirror design is (*dashed* in Fig. 5-108)

air 0.672L 0.376H 0.0110M .571L 0.0117M 0.36L *metal*

and cold mirror (*solid* in Fig. 5-108) is

air 1.115L 0.306H 0.462L 0.486H 0.432L 0.43H 0.477L 0.514H 0.408L 0.416H 0.454L
0.567H 0.352L 0.861H 0.408L 0.702H 0.505L 0.582H 0.763L 0.474H 0.724L 0.479H
0.865L 0.368H 0.996L 0.2822H 0.757L 0.843H 0.011M 0.506L 0.0117M 0.663L *metal*

where L and H represent layers of refractive indices 1.45 and 2.30, respectively. The optical thickness of L and H is $\lambda_0/4$ at λ_0 of 1.00 μm. The thickness of M is metric (in μm). Optical constants of M and *metal* are in titanium table (p. 35) of Weaver [81].

5.8.6 Koppelmann's equation

The derivation follows Arnon [78a]. The total absorption A in a quarterwave stack (containing ℓ layers) is found by summing Eq. (2-63d) for each layer

$$A \approx (1-R) \sum_{i=1}^{\ell} A_i = (1-R) \sum_{i=1}^{\ell} \alpha_i \, G_i \, h_i \; . \tag{5-33a}$$

The stack contains alternating quarterwave layers of high index n_H and low index n_L. Because the thickness of each layer is a QWOT (see §2.7.2.1), the metric thickness of each high index and low index layer is $\frac{1}{4}\lambda_0 \, n_H^{-1}$ and $\frac{1}{4}\lambda_0 \, n_L^{-1}$, respectively. Conversion of the α_i in Eq. (5-33a) into an extinction coefficient [see Eq. (2-14b)] leads to

$$A \approx (1-R)\,\pi\, k_H \, n_H^{-1} \sum_{i=1\,(odd)}^{\ell-1} G_i \; + (1-R)\,\pi\, k_L \, n_L^{-1} \sum_{i=2\,(even)}^{\ell} G_i \; . \tag{5-33b}$$

The approximations are used that $G_i \approx \frac{1}{2} V_i$ and from Eq. (2-19g) that

$$1 - R \approx 4\, V_m^{-1} \tag{5-33c}$$

where V_m is the maximum standing wave ratio of the entire stack. Consider the first *odd* term of Eq. (5-33b), which contains the V of the entire stack reduced by the *ratio* of the indices of the media on either side of the top interface of the outermost layer

$$A \propto 4\pi\, V_m^{-1}\, k_H\, n_H^{-1}\, \frac{V_m}{2}\, \frac{n_0}{n_H} + \ldots \tag{5-33d}$$

and the next *odd* term is

$$A \propto \ldots + \ldots 4\pi\, V_m^{-1}\, k_H\, n_H^{-1}\, \frac{V_m}{2}\, \frac{n_0}{n_H} \left(\frac{n_L}{n_H}\right)^2 + \ldots \tag{5-33e}$$

The first *even* term and the next even term are, respectively

$$A \propto \ldots + \ldots 4\pi\, V_m^{-1}\, k_L\, n_L^{-1}\, \frac{V_m}{2}\, \frac{n_L n_0}{n_H n_H} + \ldots, \tag{5-33f}$$

$$A \propto \ldots + \ldots 4\pi\, V_m^{-1}\, k_L\, n_L^{-1}\, \frac{V_m}{2}\, \frac{n_L n_0}{n_H n_H} \left(\frac{n_L}{n_H}\right)^2 + \ldots \tag{5-33g}$$

Collecting terms results in

$$A \approx 2\pi\, n_0\, n_H^{-2}\, (k_H + k_L) \sum_{i=0}^{\ell} \left(\frac{n_L}{n_H}\right)^{2i} . \tag{5-33h}$$

In the limit of a large layer count, the foregoing geometric series is summed to infinity

$$A \approx 2\pi\, n_0\, n_H^{-2}\, (k_H + k_L) \left(1 - \frac{n_L^2}{n_H^2}\right)^{-1} \tag{5-33i}$$

which is essentially Eq. (5-10d).

Chapter 6
Beamdividers and polarizers

6.1 Introduction

This chapter is more difficult to organize than others because of the great variety of requirements that a beamdivider (BD) must meet. For example, if a beamdivider design is developed with a division of 50% reflectance R and 50% transmittance T, what probability is there that the reader will find this of little interest because a beamdivider with a 30%-70% division is needed? §1.8.6.3 cites parameters that specify the performance of a beamdivider as:

(1) Whether the coating is nonimmersed or immersed. In §1.8.1.1.1 it is noted that immersed coatings usually manifest no small amount of polarization splitting.

(2) The range of angle of incidence and spectral region over which it must function.

(3) The amount of polarization, by which is meant that the p transmittance is much higher than the s. The amount of polarization of the output beam is often irrelevant, as when the input light is randomly polarized and the receiver (such as the human eye) of the reflected light from the beamdivider is insensitive to polarization. As compared to the coatings in the category (4) below, these are easier to produce.

(4) In another class of beamdividers, both the degree of polarization [see Eq. (1-5c)] and the differential phase shift (see §2.3.4) are to be controlled.

It is natural that the subject of polarizers appears in this chapter, because both beamdividers and polarizers function at nonnormal incidence.

The concepts in previous sections related to periodic structures — especially §2.7.2.2 and §2.5 — are used extensively in this chapter.

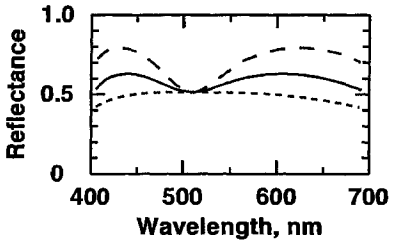


Fig. 6-1 $R_a(\lambda)$ [see Eq. (1-5a)] at 45° incidence of (*short dash*) *air* **L H L H L** *glass*, (*long dash*) *air* **HH L H L H L HH** *glass* and (*solid*) *air* **MM L H L H L MM** *glass* where the refractive indices of **L**, *glass*, **M** and **H** are 1.38, 1.52, 1.60 and 2.30, respectively. **L**, **M** and **H** represent layers of $\lambda_0/4$ optical thickness (*matched* at 45°) at λ_0 of 510 nm.

6.1.1 Examples of beamdividers

The temptation is resisted to delve immediately into the theory of beamdivider design; rather, the design of a few commonly used beamdividers is reviewed. In each case the coating meets the requirements of case (3) in §6.1 — it is designed to split randomly polarized light (or, for the IR, the radiant flux).

6.1.1.1 Reflector coated on a slab with 50% *R* and 50% *T*

This all-dielectric coating functions at 45° incidence and is intended to reflect 50% and transmit 50% of randomly polarized light in the visible spectral region — 400 nm to 700 nm. The multilayer is coated on one surface of a slab of BK-7 with plane parallel surfaces. An antireflection coating is deposited on the reverse side of the slab.

A trial starting design is suggested by Fig. 5-16, which shows that three-quarterwave thickness layers manifest a reflectance of close to 50% — at normal incidence, at least. The "core" of the starting design, however, is not three but rather five alternate layers of magnesium fluoride **L** (n = 1.38) and zinc sulfide **H** (n = 2.30). Two additional layers are used because the p reflectance at nonnormal incidence is less than at normal incidence. Additional layers are added to the "core," for reasons mentioned below. The optical thickness of each layer is a quarterwave at 510 nm (*matched* at 45° — see §6.1.2.2 and §2.7.2.2), which is the center of the visible on a frequency scale [see Eq. (4-10)].

The *short-dash* curve in Fig. 6-1 shows that the coating

air **L HL HL** *glass*

reflects adequately in the green spectral region. The reflectance, however, sags in the blue and red portions of the spectrum. This sag is eliminated by adding a layer of halfwave optical thickness to each end of the stack (see §5.3.6.1.2). The *long-dash* curve in Fig. 6-1 shows the reflectance of

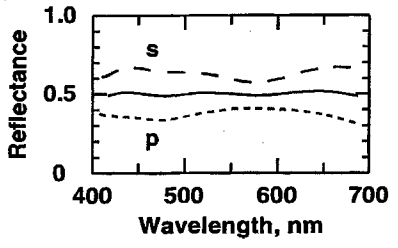


Fig. 6-2 At 45° incidence, $R_a(\lambda)$ (*solid*), $R_p(\lambda)$ (*short dash*) and $R_s(\lambda)$ (*long dash*) of

air 0.173**H** 1.094**L** 0.4205**H** 0.783**L**
0.2545**H** 1.092**L** 0.9047**H** 1.1173**L**
1.1015**H** 2.515**L** 0.350**H** 0.4678**L** *glass*

where the refractive indices of **L**, *glass* and **H** are 1.38, 1.52 and 2.30, respectively. **L** and **H** represent layers of $\lambda_0/4$ optical thickness at λ_0 of 510 nm.

air **HH L H L H L HH** *glass* .

The $R(\lambda)$ is, indeed, increased in the red and blue spectral regions — but increased too much.

One method of lowering that blue-red reflectance is to use a lower refractive index for the halfwave layers. The design is thus modified to

air **MM L H L H L MM** *glass* .

Computer calculations reveal that an index in the neighborhood of 1.6 is preferable for the **M** layers. The $R(\lambda)$ of the resulting beamdivider is the solid curve in Fig. 6-1.

The next step is to replace each **M** layer with an equivalent layer (see §2.6.7). An *lhl* simulation is used on one end of the stack and an *hlh* simulation on the opposite end. As discussed in §5.3.5.3, this helps to compensate for the dispersion of the index of the equivalent layer. The final design is

air 0.4**L** 0.18**H** 0.8**L** 0.18**H** 1.4**L H L H L** 0.235**H** 0.49**L** 0.47**H** 0.49**L** 0.235**H** *glass* .

An optimization computer code (see §10.2) then adjusted the layer thicknesses with the intent of producing a mean reflectance of 50%. The layer count was reduced by two, and the resulting $R(\lambda)$ appears in Fig. 6-2. Its reflectance of the p polarized light is close to 35%, and for the s it is 65%. The higher R_s, as compared to R_p, is a manifestation of the *polarization splitting* mentioned in §6.1.2.5. There is little point in optimizing the design further to smooth out the $R(\lambda)$ curve, inasmuch as the dispersion (see §3.5.2.2) of the zinc sulfide must be taken into account in the final design.

6.1.1.2 Immersed reflector with 50% *R* and 50% *T*

The next example is a beamdivider whose performance is similar to that of the previous beamdivider, with the exception that this one is *immersed*. Examples of immersed systems (see §1.8.1.1) are the coating cemented between Porro prisms in panel (b) of Fig.

1-146 and the cemented cube in the "a" *panel* of Fig. 1-155. Using the same coating materials as in §6.1.1.1, the trial design consists of a quarterwave stack

glass **H L H L** *glass*

where the layers are *matched* QWOTs at $\lambda_0 = 510$ nm. The *short-dash* curve in Fig. 6-3 indicates that, although the stack reflects well at λ_0, the reflectance sags to low values in the blue and the red spectral regions. This is remedied by adding a three-layer "Fabry-Perot" containing the sequence **HH L HH** to one end of the stack (see §5.3.6.1). The trial design is

glass **HL HL HH L HH** *glass* .

Its $R_a(\lambda)$ is the *long-dash* curve in Fig. 6-3. The foregoing design is similar to that of Catalán [63] when the halfwave layers are removed.

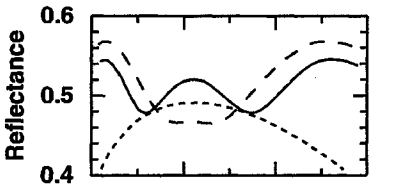


Fig. 6-3 $R_a(\lambda)$ [see Eq. (1-5a)] at 45° incidence in *glass* of the beamdivider (*short dash*)
glass **H L H L** *glass*,
(*long dash*)
glass **HL HL HH L HH** *glass*
and (*solid*)
glass **HL HL HH M HH** *glass*
where the refractive indices of **L**, *glass*, **M** and **H** are 1.38, 1.52, 1.73 and 2.30, respectively. **L**, **M** and **H** represent layers of $\lambda_0/4$ optical thickness (*matched* at 45°) at λ_0 of 510 nm.

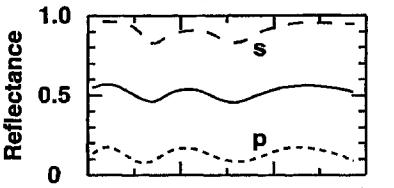


Fig. 6-4 At 45° incidence in *glass*, $R_a(\lambda)$ (*solid*), $R_p(\lambda)$ (*short dash*) and $R_s(\lambda)$ (*long dash*) of
glass **HL HL HH**
(0.28**H** 0.343**L** 0.28**H**) **HH** *glass*
where the refractive indices of **L**, *glass* and **H** are 1.38, 1.52 and 2.30, respectively. **L** and **H** represent layers of $\lambda_0/4$ optical thickness (*matched* at 45°) at λ_0 of 510 nm.

The trial design would be improved if its reflectance at λ_0 were increased. This is accomplished by substituting an intermediate index layer ($n = 1.73$) for one of the **L** layers, as in the coating

glass **HL HL HH M HH** *glass* .

Its $R_a(\lambda)$ is the solid curve in Fig. 6-3. An equivalent layer (see §2.6.7) is substituted for that **M** layer. Figure 6-4 shows that its reflectance is 51% ± 5% across the visible spectrum. As illustrated in §6.1.1.1 and by Clapham [71], an optimization computer code could flatten this curve still further by adjusting layer thicknesses. The *p* reflectance in Fig. 6-4 is an anemic 10%, whereas the *s* reflectance approaches 90%. The polarization splitting (see §6.1.2.5) of this "glass-to-glass" design is even more pronounced than that of the "air-to-glass" design in §6.1.1.1.

6.1.1.3 IR beamdivider coated on a slab

The designs in the previous sections of §6.1 are derived from Chap. 5. The design philosophy is: "The construction of a broadband reflector is the first step in the design of a beamdivider." The beamdivider in §6.1.1.3, however, stems from an all-dielectric

antireflection coating. This coating was originally designed for use in a Fourier transform spectrometer [Heinrich 67].

The beamdivider is required to reflect at least 40% in a broad spectral region centered at $\lambda_0 = 4.5$ μm. The substrate is a slab of barium fluoride—its nominal index is 1.44 at λ_0. The starting point of the design procedure is to assume that a slab of material exists with an index of 5.50. According to Eq. (2-12f), the Fresnel reflectance R at its surface is

$$R = (1 - 5.50)^2 (1 + 5.50)^{-2} = 0.48 \ , \qquad (6\text{-}1a)$$

which is reasonably close to the target of 50%. The problem is that a nonabsorbing bulk material with that index does not exist—in the IR, at least. It is possible, however, to design a three-layer stack that has a surface admittance (mentioned in §2.2.5.4) of 5.50. This surface admittance simulates the aforementioned bulk material.

Figure 6-5 presents the design procedure. The first step is the design of a three-layer Chebyshev antireflection coating using the methodology of §4.7.2. The substrate index is 5.50 and the incident medium index is 1.44—the barium fluoride mentioned above. Each layer has an optical thickness of $\lambda_0/4$ (*matched* at 45°).

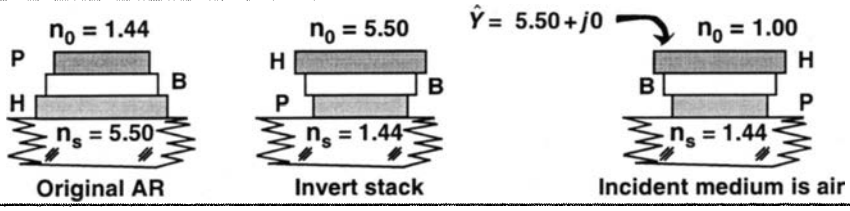


Fig. 6-5 Steps in IR beamdivider design: (*Left panel*) A three-layer antireflection coating matches an incident medium of refractive index 1.44 to a pseudo-substrate of index 5.50;
(*center panel*) the system is inverted; and (*right panel*) *air* replaces the medium of index 5.50. $\hat{Y}$ is the stack admittance at λ_0.

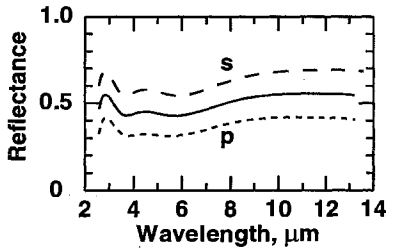


Fig. 6-6 At 45° incidence in air, $R_a(\lambda)$ (*solid*), $R_p(\lambda)$ (*short dash*) and $R_s(\lambda)$ (*long dash*) of *air* **H B P** *substrate* .

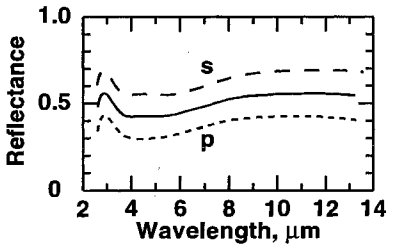


Fig. 6-7 The caption to Fig. 6-6 obtains, with the exception that the design is *air* **H** 0.109**L** 0.101**H** 0.218**L** 0.101**H** 0.218**L** 0.101**H** 0.109**L P** *substrate* .

For both Figs., the refractive indices of the **H**, **B**, **L** and **P** layers are 4.20, 2.75, 2.22 and 1.73, respectively. Each layer is $\lambda_0/4$ optical thickness (*matched* at 45°) at λ_0 of 4.5 μm. The dispersive *substrate* is barium fluoride [Malitson 64].

According to Eq. (4-13a), the refractive index of the center layer (whose symbol is **B**) should be $2.81 = \sqrt{(5.50)(1.44)}$. A slightly lower value of 2.75 is used so that the coating will perform better at longer wavelengths, where the index of the barium fluoride

substrate is lower. According to §4.2.3, the three-layer antireflection coating (see *left panel* of Fig. 6-5) steps the admittance $\hat{Y}_s$ of a high index pseudo-substrate ($n_s = 5.50$) downward to the incident medium of index 1.44.

The antireflection coating is inverted in the *center panel* of Fig. 6-5. After this is accomplished, the substrate is a material of index 1.44 and the pseudo-incident medium has an index of 5.50. The antireflection coating transforms the admittance of the 1.44 index substrate *upwards* to 5.50. The incident medium is air in the *right panel.* At $\lambda = \lambda_0$, the admittance $\hat{Y}_s$ at the outer surface of the inverted antireflection coating stack is 5.50. It is the admittance mismatch between air and this stack of admittance of 5.50 (i.e., the Fresnel reflectance) that provides the required reflectance of 48%. The spectral reflectance of the three-layer beamdivider

air **H B P** *substrate*

appears in Fig. 6-6. The **H** layer is evaporated germanium and the **P** layer is lead fluoride. The temptation to use silicon monoxide for the **P** layer is resisted because of its absorption at longer wavelengths.

The remaining problem is to find a coating material with the proper refractive index for the center **B** layer. It is tempting to replace it with an equivalent layer of 90° phase thickness. This would work well over a relatively narrow spectral region. The dispersion of index of the equivalent layer would, however, limit the bandwidth of this coating. The effects of dispersion are reduced by cascading three equivalent layers — each with a 30° phase thickness. As shown in the caption to Fig. 6-7, those equivalent layers are fabricated from germanium ($n = 4.2$) and zinc sulfide ($n = 2.22$).

A final comment relates to the spectral bandwidth of this coating, which is limited by the index of the outer **H** layer. The design procedure is to decrease the bandwidth w_q [see Eq. (4-17b)] until the **H** layer index is 4.2 — which is germanium and, therefore, is attainable. For example, a beamdivider with a broader bandwidth could be fabricated if a higher index material, such as lead telluride ($n = 5.48$), were used for the outer **H** layer. A four-section or five-section Chebyshev antireflection coating would reduce the ripple in the reflectance.

6.1.2 Terminology associated with polarizer design and usage

The purpose of §6.1.2 is to review terminology used in subsequent parts of this chapter.

6.1.2.1 Nonpolarizing

Suppose that a beamdivider is designated as *nonpolarizing* at a particular wavelength λ_0 and incidence angle θ_0. *Nonpolarizing* has two meanings: (1) The polarization state of the incident light is unaltered after it reflects from the coating; or (2) the radiant reflectance for the *p* polarized and *s* polarized light is the same.

The first of the foregoing definitions describes a *polarization state preserving coating.* It would be used, for example, in an optical system that transfers polarized light to an ellipsometer. The coating has the property that not only is its radiant reflectance the same for *p* polarized light and *s* polarized light, but, in addition, the phase shift upon reflection is equal for these two components, as well. The second definition relaxes the requirements for the phase. Unless mention is made to the contrary, *nonpolarizing* requires only that R_p and R_s be equal.

6.1.2.2 Matched at angle

This term is defined in §2.7.2.2. For example, the optical thicknesses of the **H**, **B** and **P** layers captioned in Fig. 6-6 are adjusted so that the phase thickness of each is 90° at the incidence angle of 45°. Those thicknesses would be readjusted if the incidence angle were changed.

6.1.2.3 Matched at all angles

This phrase refers to a stack in which the optical thickness of each layer has a prescribed phase thickness, regardless of the angle of incidence. This is accomplished by adjusting all layer thicknesses each time the incidence angle is altered. In other words, suppose the properties of a particular coating were actually to be determined by depositing the coating and then measuring its R or T on a spectrophotometer. A different coating would be fabricated for each angle of incidence.

6.1.2.4 Polarizing angle

The polarizing angle (symbol θ_p) at which a nonabsorbing stack has zero reflectance at a particular wavelength. This may occur for light polarized either p or s. It is the analog of Brewster's angle (which applies strictly to a dielectric interface) for a multilayer. Several examples appear in §6.2.2.

6.1.2.5 Polarization splitting

As indicated in §1.8.1.1.1, polarization splitting is a measure of the ratio of the reflectance (or transmittance) of s polarized and p polarized light. Although there are exceptions to every general rule, it is true that most multilayer dielectric coatings reflect the s more strongly than the p. Polarization splitting is evident in Fig. 6-4, which shows that R_p hovers around 17%, whereas R_s is close to 85%. Thus the ratio R_s:R_p exceeds 4:1.

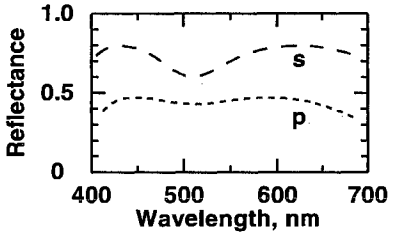


Fig. 6-8 At 45° incidence angle in air, $R_p(\lambda)$ (*short dash*) and $R_s(\lambda)$ (*long dash*) of *air* **MM L H L H L MM** *glass* where the refractive indices of **L**, *glass*, **M** and **H** are 1.38, 1.52, 1.60 and 2.30, respectively. **L**, **M** and **H** represent layers of $\lambda_0/4$ optical thickness (*matched* at 45°) at λ_0 of 510 nm.

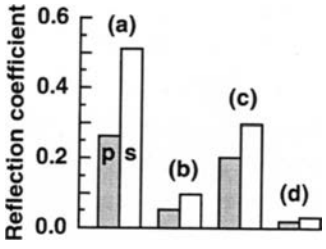


Fig. 6-9 At 45° incidence angle in air, magnitude of the Fresnel amplitude reflection coefficients for p polarization (*shaded*) and s polarization (*non-shaded*) at the interface between media with the following refractive indices: (a) 1.00, 1.60; (b) 1.60, 1.38; (c) 1.38, 2.30; and (d) 1.60, 1.52.

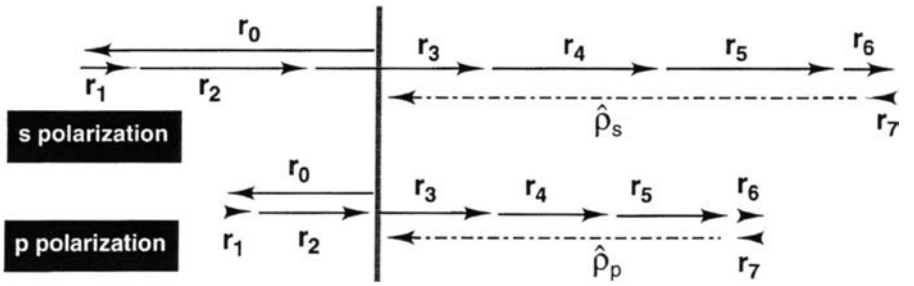


Fig. 6-10 For the coating captioned in Fig. 6-8, at 45° incidence angle in air and at $\lambda = 510$ nm, the vector addition of amplitudes approximates the net amplitude reflection coefficients (*short dash*) $\hat{\rho}_s$ and $\hat{\rho}_p$. The vertical *long-short dashed* line is the imaginary axis that intersects the origin.

The vector addition of amplitudes provides insight to the origin of polarization splitting. An example is the coating cited in §6.1.1.1

air **MM L H L H L MM** *glass* .

Its R_p and R_s are graphed in Fig. 6-8.

The first step is to compute the Fresnel amplitude reflection coefficient at each interface, using Eq. (2-8c) for *s* polarization and Eq. (2-9c) for *p* polarization. Figure 6-9 graphs the magnitude of the amplitude reflection coefficients at the interfaces — at the *glass*-**M**, **M**-**L**, **L**-**H** and **M**-*glass* interfaces. In each instance, the magnitude of the Fresnel coefficient is larger for *s* polarized light than for *p*.

It is the aforementioned effect that produces the polarization splitting, since the phase retardation β of each layer is precisely the same for both *p* and *s*. The cumulative effect of the difference between the Fresnel coefficients for *p* and *s* is evident in Fig. 6-10, which depicts a vector diagram at $\lambda = \lambda_0$. The vectors are collinear because each of the layers is *matched* in phase thickness (see §2.7.2.2 and §6.1.2.2) so that its β is 90° or a multiple thereof. A consequence is that, at this wavelength, the net amplitude reflection $\hat{\rho}$ lies on the real axis in the complex $\hat{\rho}$ plane. The difference in length between the vectors $\hat{\rho}_p$ and $\hat{\rho}_s$ is accentuated even more when the radiant reflectances $R_s \equiv |\hat{\rho}_s|^2$ and $R_p \equiv |\hat{\rho}_p|^2$ are computed. The square of a number less than unity makes it even smaller.

6.2 Nonimmersed linear polarizers containing quarterwave layers

6.2.1 Introduction

The *quarterwave* in the section title means that the optical thickness of each dielectric layer is $\lambda_0/4$—*matched at all angles.* Multilayers are analyzed that manifest a *polarizing angle* θ_p (see §6.1.2.4). Randomly polarized light incident upon such a coating at θ_p is reflected as linearly polarized. The objective of §6.2 is to offer insight into how such coatings function. Although it is tempting to believe that many of these coatings are impractical because θ_p is often close to 90°, there are applications related to an optical waveguide (see §1.7.8).

6.2.1.1 The use of reduced admittance to compute the reflectance

The computation of the reflectance of the coatings in §6.2.2.1.1, §6.2.2.2.1 and §6.2.3.1 by the vector addition of amplitudes is only an approximation, although it offers insight. It is inaccurate if the stack contains many layers. An exact value for the reflectance can be computed using reduced admittance. It is applicable to both single layers and multilayers.

6.2.1.1.1 Recapitulation of relationships

§2.7.3.3 presents equations for the reduced admittance y of a nonabsorbing stack in which the phase thickness of each layer is 90° (*matched at all angles).* At nonnormal incidence, y can be expressed as either the product or quotient of two functions, U [Eq. (2-48d)] and W [Eq. (2-48e)]. Summarized below are the relationships between U and W, from which the polarizing properties of the stack may be computed. If

$$R_p = 0 \quad \text{at } \theta_p, \quad \text{then} \quad U = W, \tag{6-2a}$$

or if

$$R_s = 0 \quad \text{at } \theta_p, \quad \text{then} \quad U = W^{-1}. \tag{6-2b}$$

In other words, once U is computed and W is tabulated at various angles, many of the properties of the multilayer are evident, such as:

(a) The angle at which θ_p occurs.

(b) Whether R_p or R_s is zero at θ_p.

(c) The angle θ_n at which R_p equals R_s (i.e., the stack is nonpolarizing).

It follows that if $U = 1$, the stack is nonpolarizing at all angles (see §6.5.2). Whereas, if $W = 1$ at a particular angle θ_n, the multilayer is nonpolarizing only at that angle.

6.2.1.1.2 Reflectance in the limit of the angle of incidence → 90°

Consider a coating containing dielectric layers—the phase thickness of each layer is 90° *matched at all angles.* Its reflectance can be found by examining the change of W with angle of incidence.

(a) Provided that the incident medium refractive index is less than that of all layers of the stack and also is less than the substrate index, the reflectance (regardless of the polarization) always approaches 100% as grazing incidence occurs, i.e., as $\theta_0 \rightarrow 90°$.

Proof: It can be shown that, as long as the critical angle (see §2.2.2.2) does not occur in any of the layers, all but one of the cosines in Eq. (2-48e) tend to a nonzero limit as $\theta_0 \rightarrow 90°$. There is a unique cosine—the cosine of the incidence angle $\cos\theta_0$. It approaches zero as $\theta_0 \rightarrow 90°$. Because $\cos\theta_0$ appears in the denominator of W [Eq. (2-48e)], W diverges as $\theta_0 \rightarrow 90°$. Hence $|\hat{\rho}_p| \rightarrow 1$ and $|\hat{\rho}_s| \rightarrow 1$, where $\hat{\rho}_s$ and $\hat{\rho}_p$ appear in Eqs. (2-48f) and (2-48g), respectively.

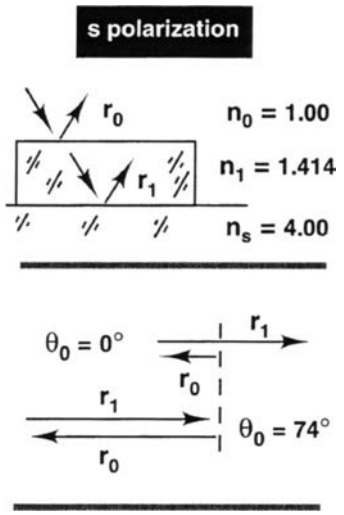


Fig. 6-11 (*Upper panel*) Showing the Fresnel amplitude reflection coefficients r_0 and r_1 at the two interfaces of a single layer. (*Lower panel*) Vector addition of amplitudes is used to compute the net amplitude reflection coefficient at normal incidence and at an incidence angle θ_0 of 74° for *s* polarized light. The phase thickness of the layer is 90°, *matched at all angles.*

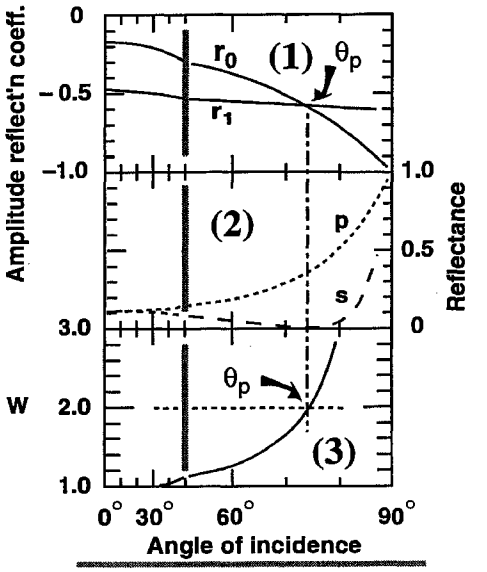


Fig. 6-12 Versus angle of incidence θ_0 in air for the layer captioned in Fig. 6-11:
[*Panel* (1)] Fresnel amplitude reflection coefficients at the two interfaces for *s* polarization.
[*Panel* (2)] $R_p(\lambda)$ (*short dash*) and $R_s(\lambda)$ (*long dash*).
[*Panel* (3)] *W* [see Eq. (6-3a)].
The vertical *long-short dashed* line is the polarizing angle. The scale of the abscissa changes at 50°.

(b) A necessary and sufficient condition for a polarizing angle θ_p to exist is that the stack contain an odd number of layers.

Proof: The $\cos\theta_s$ and $\cos\theta_0$ in Eq. (2-48e) cannot cancel if the stack contains an odd number of layers, even though n_0 equals n_s. Using the same arguments of §6.2.1.1.2(a), $W \rightarrow \infty$ as $\theta_0 \rightarrow 90°$. Due to the fact that *W* is identically 1.00 at normal incidence, there must exist an angle θ_p at which either Eq. (6-2a) or Eq. (6-2b) is satisfied.

(c) For a stack with an even number of layers, a necessary and sufficient condition for a polarizing angle θ_p to exist is that $n_0 \neq n_s$.

Proof: The "$\cos\theta_s$" and "$\cos\theta_0$" in Eq. (2-48e) cannot cancel if $n_0 \neq n_s$. Hence, there is a unique "$\cos\theta_0$" in the denominator. Using the same arguments of §6.2.1.1.2(a), *W* diverges as $\theta_0 \rightarrow 90°$. Due the fact that *W* is 1.00 at normal incidence, there must exist an angle θ_p at which either Eq. (6-2a) or Eq. (6-2b) is satisfied.

6.2.2 Single layers on a germanium substrate

Figure 10-52 graphs the reflectance vs phase thickness of a layer of *lower* refractive index (n_1 = 1.414) deposited on a germanium substrate. A layer of *higher* index (n_1 = 2.828)

manifests an identical reflectance curve at normal incidence. At nonnormal incidence, however, the reflectances of the *lower* and *higher* index layers are quite different. In particular, the *lower* manifests a polarizing angle (see §6.1.2.4) for *s* polarization, whereas a polarizing angle occurs for the *higher* for *p* polarization. In the analyses in §6.2.2.1 and §6.2.2.2, it is assumed that the phase thickness of each layer is 90°, *matched at all angles.*

6.2.2.1 Performance of the *lower* refractive index layer

The coating is analyzed by two methods: (1) vector addition of amplitudes; and (2) reduced admittance vs angle.

6.2.2.1.1 Analysis—vector addition of amplitudes

Figure 6-11 depicts the Fresnel reflections r_0 at the outer interface of the layer and r_1 at its inner surface. Both r_0 and r_1 are "rare-to-dense" reflections—hence, they are negative. Shown is a vector addition of amplitudes plot—a graphical representation of Eq. (2-54a). At $\theta = 0$, the vectors are in opposite directions and the length of r_1 exceeds that of r_0.

Figure 6-12 graphs r_0 and r_1 vs angle of incidence θ_0 for *s* polarization. The lengths of both r_0 and r_1 steadily increase as θ_0 increases. The length of r_0 increases at a faster rate, however. The consequence is that for a polarizing angle θ_p of 74°, r_0 and r_1 are of equal length—the net amplitude reflection coefficient is zero. R_s is zero, as shown in *panel* (2).

The question arises: Is it possible for R_p to be zero? The answer is no. If a vector diagram were drawn for *p* polarization, it would show that the length of r_0 shrinks until it is zero at 55°, which corresponds to Brewster's angle (see §2.2.3.3.2) for an index of 1.414. Because at angles in excess of 55° r_0 and r_1 are collinear and in the *same* direction, there is no possibility that they cancel.

6.2.2.1.2 Analysis—reduced admittance

The variable W and parameter U are written for this layer. Starting with Eq. (2-48e) and substituting $n_0 = 1.00$, W is

$$W = \frac{(1 - n_1^{-2}\sin^2\theta_0)}{\sqrt{1 - n_s^{-2}\sin^2\theta_0}\ \sqrt{1 - \sin^2\theta_0}} \tag{6-3a}$$

and from Eq. (2-48c) the reciprocal of U is

$$U^{-1} = \left[\frac{(1.414)^2}{(1.00)\,(4.00)}\right]^{-1} = (0.50)^{-1} = 2.00\ . \tag{6-3b}$$

Panel (3) of Fig. 6-12 plots $W(\theta_0)$. This function is 1.00 at normal incidence and monotonically rises with increasing θ_0. According to Eq. (6-2b), the polarizing angle θ_p occurs when W equals the right of Eq. (6-3b).

6.2.2.2 Performance of the *higher* refractive index layer

The analysis of §6.2.2.1 is repeated for the *higher* index ($n_1 = 2.828$) layer.

6.2.2.2.1 Analysis—vector addition of amplitude

Fig. 6-13 depicts the vector diagram at normal incidence. As in §6.2.2.1.1, both r_0 and r_1 are negative because they are "rare-to-dense" reflections. The vectors are in opposite

directions, however, because of the "*double phase thickness* = $-2\beta = -180°$" of the film. In contrast to §6.2.2.1, however, the external reflection r_0 is relatively large and the internal reflection r_1 is smaller. It is evident that r_0 must *shrink* in order for the vectors r_0 and r_1 to be of equal length. This shrinkage occurs only for p polarization, as shown in Fig. 6-13.

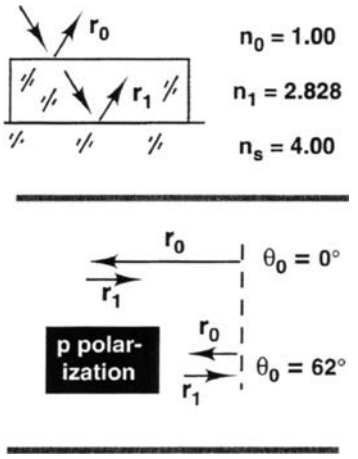


Fig. 6-13 (*Upper panel*) Showing the Fresnel amplitude reflection coefficients r_0 and r_1 at the two interfaces of a single layer. (*Lower panel*) Vector addition of amplitudes is used to compute the net amplitude reflection coefficient at normal incidence and at θ_0 of 62° for p polarized light. The phase thickness of the layer is 90°, *matched at all angles*.

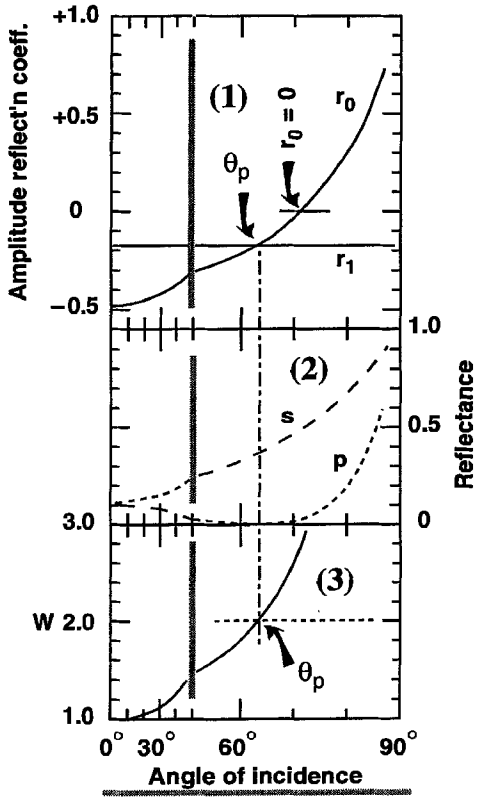


Fig. 6-14 Versus angle of incidence θ_0 in air for the layer in Fig. 6-13:
[*Panel* (1)] Fresnel amplitude reflection coefficients at the two interfaces for p polarization.
[*Panel* (2)] $R_p(\lambda)$ (*short dash*) and $R_s(\lambda)$ (*long dash*).
[*Panel* (3)] W [see Eq. (6-3a)].
The scale of the abscissa changes at 50°. The vertical *long-short dashed* line is the polarizing angle.

Figure 6-14 graphs the Fresnel amplitude reflection coefficients for p polarization vs angle of incidence. *Panel* (1) shows that the magnitude of r_0 shrinks steadily at a faster rate than r_1 until the magnitude of r_0 and r_1 are equal at the polarizing angle θ_p of 62 °. Figure 6-13 shows that the vectors are considerably shorter at nonnormal incidence

than at normal. The vectors oppose (and cancel) each other at θ_p. *Panel* (2) shows that R_p is zero at θ_p.

Inspection of *panel* (1) reveals that r_0 is zero at an angle close to 70°. This is the Brewster angle θ_B for the outer interface. Although there is no reflection (for the *p* polarization) at the air-film interface at 70°, the lower interface is still reflecting. Using the terminology of §2.6.3.2, this is an *equi-admittance layer* at that angle. This illustrates that θ_B and θ_p are distinct and, except coincidentally, occur at different angles.

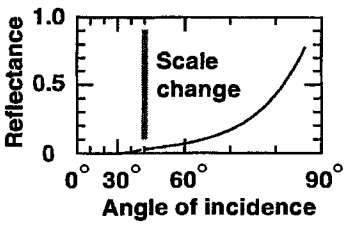


Fig. 6-15 Versus angle of incidence in air, reflectance (for either *p* or *s* linear polarization) of a nonabsorbing layer of index 2.00 deposited upon a substrate of index 4.00. The phase thickness of the layer is 90°, *matched at all angles*. The scale of the abscissa changes at 50°.

6.2.2.2.2 Analysis—reduced admittance

The polarizing angle is computed using $W(\theta_0)$, which appears in Eq. (6-3a). It is graphed in *panel* (3) of Fig. 6-14.

The parameter U [see Eq. (2-48d)] for this layer is

$$U = (2.828)^2\,(1.00)^{-1}\,(4.00)^{-1} = 2.00 \quad .(6\text{-}3c)$$

Figure 6-14 shows that $W(\theta_0)$ is 1.00 at normal incidence and that it increases with angle. At an angle of 62°, $W = U = 2.00$. This is the polarizing angle θ_p.

6.2.2.3 A single layer as a nonpolarizing reflector

The question remains: Is it possible that either of the single layers cited previously in §6.2.2 are nonpolarizing reflectors at any angle of incidence other than zero? The answer is no. The discussion in §2.7.3.3 makes it clear that R_p and R_s are equal when either U [see Eqs. (6-3b) and (6-3c)] is 1.00 or W [see Eq. (6-3a)] is 1.00. Because W in *panel* (3) of Figs. 6-12 and 6-14 is 1.00 at normal incidence and increases monotonically, neither of the single layers in §6.2.2.1 or §6.2.2.2 can function as a nonpolarizing reflector at any nonzero incidence angle.

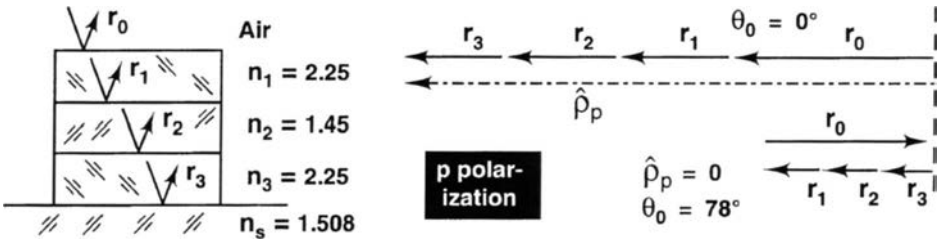


Fig. 6-16 (*Left panel*) Showing the Fresnel amplitude reflection coefficients at the interfaces of a three-layer stack. The phase thickness of each layer is 90°, *matched at all angles*. (*Right panel*) At angles of incidence of normal and 78°, the net amplitude reflection coefficient $\hat{\rho}_p$ is found using the vector addition of amplitudes.

Consider a single layer whose index of 2.00 satisfies Eq. (2-23j)—it is a "perfect" antireflection coating at normal incidence. It is, indeed, nonpolarizing *at all angles*, because its U is 1.00. The reflectance vs angle in Fig. 6-15 shows, however, that its R barely exceeds 20% at angles of incidence as high as 75°.

6.2.3 Quarterwave stack linear polarizers — odd layer count

This section analyzes the performance of designs such as

air **H L H** *glass* .

H and **L** represent layers of high index and low index materials, respectively. Each layer has a phase thickness of $\beta = 90°$, *matched at all angles* (see §6.1.2.3). Similar to the single layers in §6.2.2, these multilayers manifest a polarizing angle—but at a somewhat larger angle of incidence.

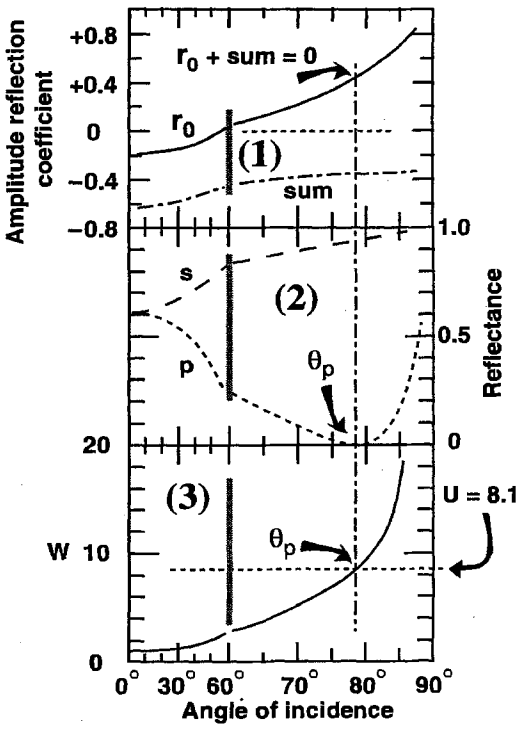


Fig. 6-17 Versus angle of incidence θ_0 in air for the design captioned in Fig. 6-16:
[*Panel* (1)] For *p* polarization, Fresnel amplitude reflection coefficient r_0 at the outer interface and the algebraic sum (*long-short dash*) of r_1, r_2 and r_3.
[*Panel* (2)] $R_p(\lambda)$ (*short dash*) and $R_s(\lambda)$ (*long dash*).
[*Panel* (3)] W [see Eq. (2-48c)].
The scale of the abscissa changes at 60°. The vertical *long-short dashed* line is the polarizing angle.

6.2.3.1 Analysis — vector addition of amplitude

Figure 6-16 shows that there are four vectors that represent the Fresnel amplitude reflection at each interface. r_0 is the reflection at the air-film interface. There are also reflections at the three internal interfaces: r_1, r_2 and r_3. Similar to the quarterwave stack cited in Figs. 5-2 and 5-4, all vectors add constructively to produce the maximum reflectance at normal incidence.

The vector r_0 is negative at normal incidence, as is evident in Fig. 6-16. r_0 for *p* polarization reverses sign and is positive at angles in excess of the Brewster angle, which is 66°. This sign reversal means that r_0 opposes the remaining three vectors r_1, r_2 and r_3. At θ_p of 78 °, the length of r_0 equals the sum of the lengths of the other vectors. The resulting amplitude reflection coefficient $\hat{\rho}_p$ (lower right of Fig. 6-16) is zero at θ_p. The *p* reflectance is zero at that angle, as plotted in *panel* (2) of Fig. 6-17.

Panel (1) of Fig. 6-17 graphs (vs angle of incidence) r_0 and the sum of r_1, r_2 and r_3. The sign reversal of r_0 is evident.

6.2.3.2 Analysis using reduced admittance

The function $W(\theta_0)$ [see Eq. (2-48e)] is graphed in *panel* (3) of Fig. 6-17. W is 1.00 at normal incidence and rises monotonically with increasing angle. For this three-layer coating U [see Eq. (2-48d)] is

$$U = (2.25)^4(1.00)^{-1}\,(1.508)^{-1}\,(1.45)^{-2} = 8.1 \qquad (6\text{-}3d)$$

and is drawn as a horizontal *dashed* line in *panel* (3). The intersection of that line and the $W(\theta_0)$ curve defines the polarizing angle.

6.2.3.3 Stacks with at least five layers

Another example is the five-layer stack

air **H L H L H** *glass*

where the refractive indices are the same as those of the previous section. Figure 6-18 and Table 6-1 indicate that the polarizing angle now exceeds 80°. The foregoing stack is a component of the Ward polarizer (see Figs. 1-176 and 1-177, §4.9.5.5 and §7.9.7). Table 6-1 shows that the polarizing angle increases as more layers are added to the stack. This type of stack was studied by Mahlein [74a].

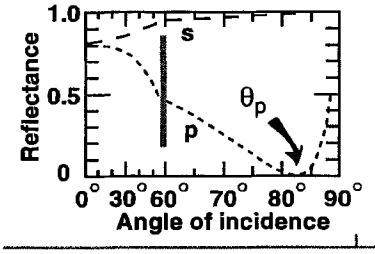


Fig. 6-18 $R_p(\theta_0)$ (*short dash*) and $R_s(\theta_0)$ (*long dash*) of *air* **H L H L H** *glass* where the refractive indices of **L**, *glass* and **H** are 1.45, 1.508 and 2.25, respectively. The phase thickness of each layer is 90°, *matched at all angles*. The scale of the abscissa changes at 60°.

6.2.4 Stack with low refractive index outer layer

The following four-layer coating is studied

air **L H L H** *glass* .

Each layer has a phase thickness of $\beta = 90°$, *matched at all angles* (see §6.1.2.3). An analogy is made with the above coating and the *lower* layer of index 1.414 in §6.2.2.1. In both cases the admittance at normal incidence is less than 1.00. According to Eq. (6-2a), the *s* reflectance is zero at the polarizing angle. Because the reduced admittance at normal incidence is less than 1.00, Eq. (6-2b) predicts that the *s* reflectance plunges to zero at the polarizing angle.

Figure 6-19 graphs both $W(\theta_0)$ and the reflectance of this design. At relatively low angles of incidence, $W(\theta_0)$ dips below 1.00. At $\theta_0 = \theta_n$ of 63°, *W* is 1.00 and, according to the teachings of §6.2.1.1.1, the coating is nonpolarizing. This is not a 50%-50% beamdivider, however, since the reflectance is 35% at θ_n.

Table 6-1 Polarizing angles of quarterwave stacks

Design	*U*	θ_p	R_p at θ_p	R_s at θ_p
air **H L H** *glass*	8.1	78.7°	zero	0.94
air **H L H L H** *glass*	19.5	82.8°	zero	0.989
air **H L H L H L H** *glass*	47	85.5°	zero	0.998
air **L H L H** *glass*	0.26	85°	0.75	zero

For the designs cited above, *U* [using Eq. (2-48b) or Eq. (2-48c)], polarizing angle θ_p, as well as R_p and R_s at that angle. The refractive indices of **L**, *glass* and **H** are 1.45, 1.508 and 2.25, respectively. The phase thickness of each layer is 90°, *matched at all angles*.

At a larger angle of 85°, *W* equals the reciprocal of *U*. Figure 6-19 shows that the *s* reflectance plunges to zero. This coating would not function well as a linear polarizer, however, since R_p at the polarizing angle is a mere 75%.

An analysis of a stack of six layers

air **L H L H L H** *glass*

shows that the nonpolarizing angle θ_n occurs at an angle in the 70° range and that the polarizing angle is even closer to grazing incidence.

6.2.5 Applications

The projected area of the substrate is an obvious consideration. At 84° incidence, the area of the coating is ten times the light beam area. Another consideration is that the angular bandwidth over which R_p or R_s is zero becomes quite narrow as the layer count increases. There are, however, at least two uses for the coatings in §6.2: (1) The Ward polarizer (see §7.9.7 and Fig. 1-176); and (2) a reflecting coating for an optical waveguide.

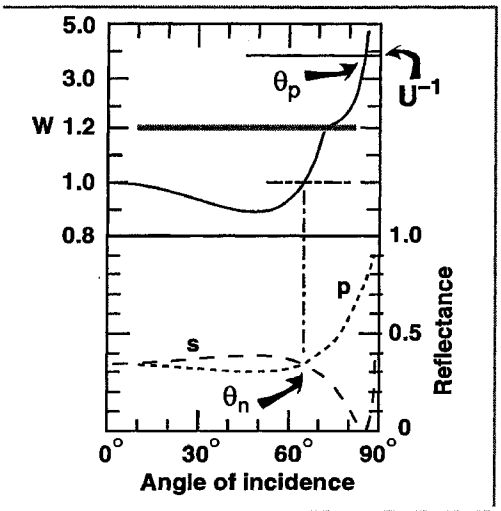


Fig. 6-19 Versus angle of incidence θ_0: [*Upper panel*] W [see Eq. (2-48c)]. The scale of the ordinate changes 1.2. [*Lower panel*] $R_p(\lambda)$ (*short dash*) and $R_s(\lambda)$ (*long dash*). The vertical *long-short dashed* line is the nonpolarizing angle θ_n. The design is *air* **L H L H** *glass*. The refractive indices of **L**, *glass* and **H** are 1.45, 1.508 and 2.25, respectively. The phase thickness of each layer is 90°, *matched at all angles*.

A waveguide, such as the one shown in cross section in Fig. 1-144, is often used to "pipe" infrared radiation. The angles of incidence of the rays are quite close to grazing. Because many reflections occur, a reflectance of 99% for randomly polarized radiation is necessary. A typical approach is to overcoat the metallic wall with a dielectric stack (see §5.7.1). The problem is that if the outer layer of the stack is **H** (see §6.2.3), the p reflectance is low. This is illustrated in §6.2.2.2. The analysis of §6.2.4 shows that the s reflectance is low if the stack terminates with an **L** layer. The conclusion is that neither the "ending with an **H**" or the "ending with an **L**" coating reflects randomly polarized flux effectively.

Figure 1-145 shows that a compromise is to terminate the coating with an eighth-wave optical thickness layer. In this way, the randomly polarized radiation suffers the lowest loss.

6.3 Immersed linear polarizers

6.3.1 Introduction

Unlike the air-to-glass slab polarizers cited in §6.2, there is a host of immersed designs that are excellent linear polarizers. With few exceptions, these coatings — often called MacNeille polarizers — transmit p polarized light and reflect the s. Furthermore, the angle of incidence is quite attainable, usually in the range of 40° to 70°. The performance of this class of coatings is based upon the attainment of the internal Brewster angle (see §2.5.3.6).

6.3.1.1 Overview of this section

§6.3.1.2 discusses the parameters that influence a coating design. §6.3.2 introduces the concept of a MacNeille pair of refractive indices. Other sections address the problems of

broadening the spectral bandwidth and the enhancement of the angular range over which the coating functions.

6.3.1.2 Issues that affect the coating design

Important issues in the design and production of these multilayer dielectric polarizers are:

(1) Joining of the coating to the prism. All immersed coatings must be cemented (see §1.9.2 and §6.6.1)—or possibly optically contacted—to a prism. Problems arise when a high index prism (such as $n = 1.75$) is cemented, since cements that match the refractive index of the prism are not commonly available. If a prism of index 1.75 is joined with a cement of index 1.517, at θ_0 of 45° a Fresnel reflection of 3% (for R_s) occurs at the glass-cement interface. An antireflection coating on the glass could reduce this reflectance.

(2) Refractive index control. This control is stringent for a MacNeille polarizer. §4.1.3.1 illustrates that refractive indices in the range from 2.25 to 1.95 may be used in a narrowband antireflection coating. This freedom to choose different indices is usually not available in the construction of a MacNeille polarizer.

(3) Range of wavelengths. The design of a MacNeille polarizer over a wide range of wavelengths is more difficult, as compared to a narrow bandwidth.

(4) Range of angles. Figure 1-196 depicts a coating in a cemented cube inserted into a convergent beam. The cone angle is ±5° in the glass. The extension of that cone angle to ±10° is by no means a trivial task. §1.8.6.2.1 mentions the cost reduction benefits that result when the manufacturer of the coating is allowed the freedom to select the angle of incidence, as in §6.3.2.1.

(5) Selection of the prism material and choice of incidence angle. §1.8.6.2.1 mentions the possibility that the prism material and angle of incidence are mandated by the customer. There is little point in reading the remainder of this section if such is the case.

Construction of the prism from BK-7 or plastic is often required because of the greater cost of other materials, especially higher index glass. If, however, the angle and/or glass index may be varied, then the following combinations of angles and prism index all produce an immersion constant of 1.13:

$$L = 1.72 \sin 40° = 1.60 \sin 45° = 1.517 \sin 48° = 1.13 \ . \qquad (6\text{-}4a)$$

This illustrates the trade-off between the angle of incidence and the prism refractive index.

6.3.2 Selection of the refractive indices

6.3.2.1 MacNeille pairs of refractive indices

For a given immersion constant L, there are two refractive indices that manifest the same effective index [see Eq. (2-38b)] for p polarization. Those indices are termed a *MacNeille pair* and satisfy Eq. (2-39d). The graphs of effective index N_p vs actual index in Figs. 2-62, 2-63 and 6-20 enable the designer to visualize the aforementioned concept. In Fig. 2-62 there are MacNeille pairs of indices A, B and A', B' that manifest the same effective index.

The question of the availability of indices is now addressed. The quasi-parabolic curves in Fig. 6-20 represent N_p vs the layer index. A MacNeille pair occurs when a

horizontal line intersects either of the two branches of the quasi-parabolic curves. But how many of those index pairs are attainable?

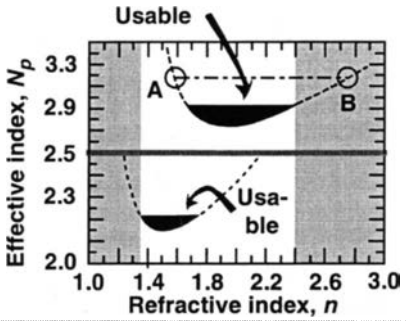


Fig. 6-20 For an incident medium of index 1.517 and incidence angle of 45° (*lower curve*) and 65° (*upper curve*), effective index (*p* polarization) of a layer material [see Eq. (2-38b)] vs its refractive index *n*. The nonshaded areas represent attainable (i.e., *usable*) indices: $1.35 \leq n \leq 2.40$. In the solid black areas, the MacNeille pairs contain *usable* indices. The ordinate changes scale at 2.5.

Assume that usable indices *n* lie in the range $1.35 \leq n \leq 2.40$, as represented by the non-shaded region of Fig. 6-20. The lower curve obtains for an angle of incidence of 45°. The usable indices (*solid black* region) range from 1.35 to 1.77. Consider the indices higher than 1.77. The problem is that the lower MacNeille pair index is below 1.35 — the *dashed* portion of the curve in Fig. 6-20.

Consider the upper curve in that Fig., which is graphed for an incidence angle of 65°. The minimum of the quasi-parabola has shifted to the right. A consequence is that the usable indices now lie in the range $1.67 \leq n \leq 2.40$ — this is the solid black area. Consider the MacNeille pair represented by *A* and *B*. Although the *A* index is attainable, the *B* index exceeds 2.40 — this MacNeille pair is not usable.

As emphasized in §2.5.3.6, one index of the MacNeille pair must be smaller than the index in Eq. (2-39b) and the other index must exceed it. An illustration is that the minimum of the upper curve in Fig. 6-20 is at $n = \sqrt{2}\, n_0 \sin 65° = 1.94$. A MacNeille pair consisting of an index of 1.98 and an index in excess of 1.98 cannot occur, since one of the indices must be less than 1.94.

Several design examples below illustrate that the selection of the indices of a MacNeille pair depends not only upon their availability, but in addition upon: (1) the angle shift (see Fig. 6-22); and (2) the attenuation of the *s* polarization.

6.3.2.2 Effective indices for *s* polarization

The selection of a MacNeille pair of indices (in §6.3.2.1) only guarantees that an internal Brewster angle (see §2.5.3.6) exists — hence, the *p* transmittance is quite large. If this coating is to reflect as much *s* polarization as possible, then layers of quarterwave optical thickness (*matched at angle*) should be used and the Fresnel amplitude reflection coefficient at each interface should be as large as possible (see §5.1.1). It follows from Eq. (2-37g) that the ratio of the effective indices (for *s* polarization) should be maximized. Several examples illustrate this point.

6.3.2.3 Design example

The incidence angle is 45° and the prisms are composed of crown glass ($n_0 = n_s = 1.517$). The goal is to reduce T_s below 1% over as much of the visible spectrum as possible. Two MacNeille pairs of indices are considered in the design.

Table 6-2 Refractive indices for MacNeille polarizers

n_H	n_L	$N_{H,s}$	$N_{L,s}$	Ratio	$N_{H,p}$ $N_{L,p}$	Layers
1.705	1.38	1.3253	0.8682	1.53	2.194	16
1.594	1.45	1.1791	0.9756	1.21	2.155	36

For the indices n_H and n_L in columns one and two, the corresponding effective indices $N_{H,s}$ and $N_{L,s}$ for s polarization [see Eq. (2-38a)], and their ratio. The effective indices $N_{H,p}$ and $N_{L,p}$ are listed for p polarization [see Eq. (2-38b)]. *Layers* is the layer count needed to produce a quarterwave stack with T_s less than 1%. L is 1.073.

6.3.2.3.1 Design procedure

The design procedure is:

(1) A trial design is the stack

glass $(\mathbf{HL})^q$ *glass*

where **H** and **L** represent layers of quarterwave optical thicknesses (*matched at angle)* of high index and low index, respectively. The integer q is the number of times the **HL** pair is repeated. This quarterwave stack is discussed in §5.2.4.3.1 and §5.3.8.

(2) Use Eq. (2-39d) to select the indices of the **L** and **H** layers—the *MacNeille pair.* In this example, one *pair* consists of indices 1.38 and 1.705. Indices of 1.45 and 1.594 are used for another *pair.*

(3) Compute the effective indices for s polarization, as listed in Table 6-2. The effective indices for p polarization should be identical, if the refractive indices for the *MacNeille pair* have been chosen correctly.

(4) Given the desired T_s of 0.01, i.e., 1%, find the standing wave ratio V_s for the s polarization from Eq. (5-6a)

$$V_s \approx 4\, T_s^{-1} = 4 / 0.01 = 400. \qquad \text{(6-4b)}$$

(5) The foregoing standing wave ratio is found at its extremum—the wavelength at which the phase thickness β of each layer is 90°. Given the indices 1.38 and 1.705 for the MacNeille pair, effective indices (i.e., N_H and N_L) are substituted for s polarization into Eq. (5-6b) to find V_s

$$V_s = N_H^{2q}\, N_L^{-2q}$$
$$= (1.325)^{2q}\, (0.868)^{-2q}\,. \qquad \text{(6-4c)}$$

An exponent $q = 7$ provides a V_s of 373, which falls short of the *desideratum* of 400. The exponent of $q = 8$ provides a margin of safety.

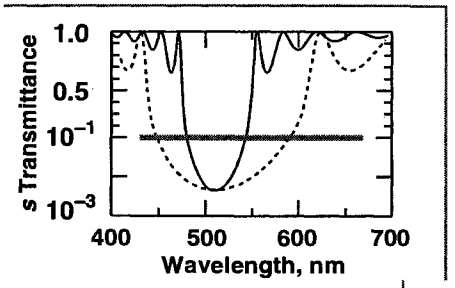


Fig. 6-21 At an angle of 45° in *glass* of index 1.517, $T_s(\lambda)$ of the design

glass $(\mathbf{HL})^q$ *glass* .

For the *dashed* curve q is 8 and the indices of **L** and **H** are 1.38 and 1.705, respectively. The indices of **L** and **H** are 1.45 and 1.594, respectively, for the *solid* curve — q is 18. Each layer is $\lambda_0/4$ in optical thickness (*matched* at 45°) at λ_0 of 510 nm. At 0.1 the ordinate scale changes from linear to log.)

6.3.2.3.2 Performance of MacNeille designs — a comparison

The foregoing calculation is repeated for the MacNeille pair 1.45/1.594. Table 6-2 lists the effective indices and that q = 18 periods are needed to produce a V_s in excess of 900 for this lower index pair.

Figure 6-21 graphs $T_s(\lambda)$ of both polarizers. The T_p of each coating exceeds 99.5% and is not plotted. For the MacNeille pair of 1.38/1.705, the spectral band over which its T_s does not exceed 1% extends from 480 nm to 560 nm. That bandwidth is smaller for the 1.45/1.594 index pair. This result follows from the teachings of §5.2.3.4.1 and §5.3.8. The bandwidth of the stop band in Eq. (5-5c) decreases as the ratio between the high index and low index becomes smaller.

The ratio of the effective indices for *s* polarization listed in Table 6-2 is smaller for the 1.45/1.594 MacNeille pair. The conclusion is that if a large bandwidth in which the *s* polarization is to be rejected is required, the two indices of the coating should be widely separated.

The indices of the substrate and incident medium do not appear in Eq. (6-4c) for V_s because they appear in both the numerator and denominator of Eq. (2-48b). They cancel because the incident medium and substrate have the same refractive index.

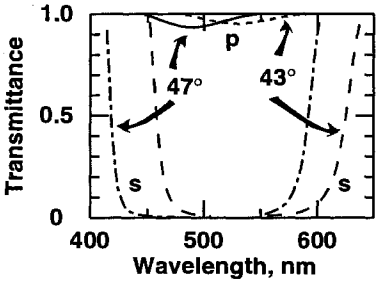


Fig. 6-22 At an incidence angle of 43° in *glass* of index 1.517, $T_s(\lambda)$ (*long dash*) and $T_p(\lambda)$ (*short dash*) of

glass $(\mathbf{HL})^8$ *glass* .

The indices of **L** and **H** are 1.38 and 1.705, respectively. Each layer is $\lambda_0/4$ in optical thickness (*matched* at 45°) at λ_0 of 510 nm. At 47° the *long-short dash* curve is $T_s(\lambda)$ and $T_p(\lambda)$ is *solid*.

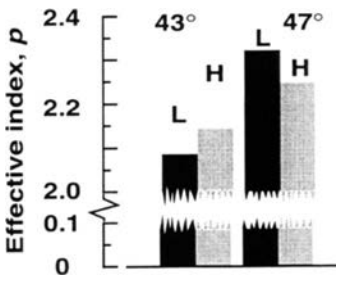


Fig. 6-23 In *glass* of index 1.517 and for *p* polarization, effective index [see Eq. (2-38b)] of **L** (index is 1.38) and **H** (index is 1.705) at 43° incidence angle and at 47°.

6.3.2.4 Range of incidence angles over which the coating performs

The aforementioned coating is optimized (i.e., "tuned") to operate at a design angle of 45°. What happens when a cone of convergent light impinges upon this stack and the angle of incidence is no longer 45°? An example illustrates such effects.

The transmittances in Fig. 6-21 are recalculated when the external angles of the meridional rays impinging upon the prism face vary by $\theta_a = \pm 3°$. The internal angle (inside the glass) θ_g is smaller, due to refraction.

Using small angles (when it is valid to replace the sine by its radian argument), Eq. (2-4a) becomes

$$\theta_g \approx \theta_a / n_s \approx 3 / 1.517 \approx 2° \qquad (6\text{-}4d)$$

where n_s is the prism index. The curves in Fig. 6-22 are computed for

$$\theta_0 = 45° \pm \theta_g = 45° \pm 2°, \qquad (6\text{-}4e)$$

namely 43° and 47°.

Fig. 6-22 shows that the shift of the angle from 43° to 47° produces a 40 nm shift to shorter wavelengths. In addition, the *p* transmittance dips to 93% at 47°. The reasons for this decrease in the *p* transmittance are discussed below.

6.3.2.4.1 Decrease of *p* transmittance as the angle of incidence changes

The *p* transmittance in Fig. 6-22 decreases when the angle of incidence exceeds the "tuned" value of 45°. The dips in transmittance, which approach 93%, are explained in terms of effective index. Figure 6-23 plots the effective indices N_p (for *p* polarization) for the **H** and **L** layers. Due to the fact that N_p for **H** and N_p for **L** are no longer equal at 43° and 47°, the index mismatch creates a narrow spectral region in which there is a stop band (see §5.2.2). It is narrow because the N_p values of the **H** and **L** are close together (see §5.3.3.1). Each period added to the stack reduces the transmittance in that stop band, as illustrated by the dips in T_p in Fig. 6-22.

6.3.2.4.2 Control of the angle shift

Mouchart [89] showed that the designer has at least some control of the angle shift, which is the shift of the transmittance with angle of incidence. It is assumed that the polarizer consists of a two-component basic period (see §6.3.2.3.1) that is repeated many times.

(1) The inclusion of as much high refractive index material as possible in the basic period minimizes the angle shift.

(2) One method of implementing (1) is to use a design (such as a 2:1 or 3:1 stack) in which there is more **H** than **L** in the basic period. Examples are the basic periods: "0.667**L** 1.333**H**" or "0.5**L** 1.5**H**" (see §5.2.4.3.2).

(3) The converse of (1) is true. The use of materials of lower refractive index in the basic period increases the angle shift.

(4) One method of implementing (3) is to use a design (such as a 2:1 or 3:1 stack) in which there is a greater thickness of **L** than **H** in the basic period.

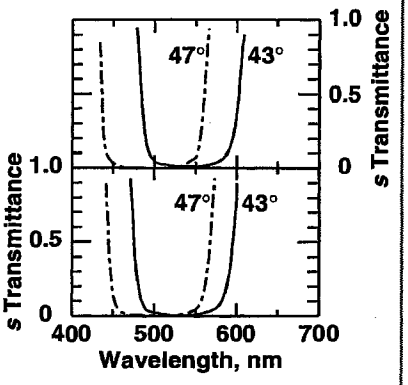


Fig. 6-24 $T_s(\lambda)$ at an incidence angle (in *glass*) of 43° (*solid*) and 47° (*long-short dash*) of the design (*lower panel*)

glass $(1.5\mathbf{H}\,0.5\mathbf{L})^{11}$ *glass*

and (*upper panel*)

glass $(0.5\mathbf{H}\,1.5\mathbf{L})^{11}$ *glass*

The indices of *glass*, **L** and **H** are 1.517, 1.38 and 1.705, respectively. Each layer is $\lambda_0/4$ in optical thickness (*matched* at 45°) at λ_0 of 510 nm.

Figure 6-24 depicts the *s* transmittance of 3:1 stacks. In the *upper panel,* the basic period is 0.5**H** 1.5**L** and in the *lower* it is 1.5**H** 0.5**L**. The coating with a thicker **H** layer in the basic period (*lower panel*) manifests a smaller angle shift. As mentioned in §5.1.3 and §5.2.5.3, however, the first-order reflection peak of a 3:1 stack has a lower reflectance, as compared to a quarterwave stack. It is for this reason that more periods are needed for a 3:1 stack to achieve the same reflectance as a quarterwave stack. The angle shift of an edge filter is the focus of §5.4.3.5.3.

The topic of angle shift is further pursued in §6.3.3.1.

6.3.3 Extension of the angular bandwidth

The angular bandwidth relates to the spread of angles at which rays impinge upon the coating. Computer optimization (see §10.2) is quite useful in the development of coatings with a wide angular bandwidth.

6.3.3.1 Use of multiple refractive indices

The design example is a polarizer at 633 nm that manifests a T_p that approaches 100% and a T_s below 0.001 (i.e., <0.1%) as the angle of incidence in the glass varies by ±10°. Using a BK-7 glass prism, this translates into a ±15° range of angles in air. It is presumed that the designer may choose the angle of incidence. This luxury was, evidently, not available to Mouchart [89], who achieved the aforementioned level attenuation with a range of angles of ±10° in air. In the example below, the lowest angle in the glass is 51° and the highest is 71°. The justification for the choice of these angles is given below.

It is advantageous to include several MacNeille pairs in the design. Each pair contains different indices and is designed to function over a particular set of incidence angles. If we designate the layers for these MacNeille pairs as $\mathbf{A}_1\,\mathbf{A}_2$ for the "A" stack, $\mathbf{B}_1\,\mathbf{B}_2$ for the "B" stack and so on, the design is

glass $(\mathbf{A}_1\,\mathbf{A}_2)^{q1}\ (\mathbf{B}_1\,\mathbf{B}_2)^{q2}\,(\mathbf{C}_1\,\mathbf{C}_2)^{q3}$ *glass*

where the exponents $q1$, $q2$ and $q3$ are chosen so that the *s* polarization is attenuated sufficiently [see step (4) in §6.3.2.3.1]. The indices of the MacNeille pair in the "A" stack are selected such that they satisfy Eq. (2-39d) at the smallest range of angles. The "B" and "C" stacks contain MacNeille pairs "tuned" at a higher and highest range of angles, respectively.

The advantage of this multiple component approach becomes evident when we consider the computer optimization that enhances the *p* transmittance. At the angle at which the "A" stack MacNeille pairs function, the *p* transmittance is relatively insensitive to thickness changes in the "A" stack. The consequence is that those layer thicknesses may be adjusted to improve the *p* transmittance at other angles of incidence and yet not

affect T_p at this lowest angle. Similar arguments may be advanced for the "B" and "C" stacks.

Table 6-3 MacNeille pairs of refractive indices

θ_0	n_H	n_L	$N_{H,s}$	$N_{L,s}$	Ratio	Layers	Shift	Symbols
51°	2.03	1.45	1.65	0.844	1.96	14	1.21	none
51°	1.74	1.60	1.28	1.08	1.19	38	1.11	**A, B**
57°	2.10	1.60	1.67	0.97	1.72	16	-	**A, C**
67°	2.34	1.74	1.88	1.04	1.81	10	-	**B, D**

At an angle θ_0 (measured in an incident medium of index $n_0 = 1.517$) and for the indices n_L and n_H in columns two and three, the corresponding effective indices $N_{H,s}$ and $N_{L,s}$ for s polarization [see Eq. (2-38a)], and their *ratio*. *Layers* is the layer count of a quarterwave stack that produces an s transmittance less than 0.1%. For n_L, *shift* is the ratio of its s effective index at 51° to its effective index at 57°. *Symbols* identifies the n_L and n_H layers, respectively, when used in a design.

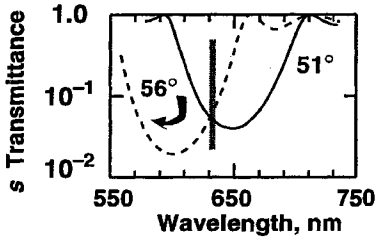


Fig. 6-25 At incidence angles of 51° (*solid*) and 56° (*dashed*) in *glass* of index 1.517, $T_s(\lambda)$ (on an absorbance scale) of

glass 1.51**A** (1.39**B** 1.51**A**)^{13} 1.39**B** 0.75**A** *glass*

where the indices of **A** and **B** are 1.60 and 1.74, respectively. Each layer is $\lambda_0/4$ in optical thickness at λ_0 of 633 nm. The vertical line is at λ_0.

6.3.3.2 Stack that functions at the lowest range of angles

This "A" stack rejects the s polarized light at the lowest angles—from 51° to 57°. Table 6-3 presents two choices for a MacNeille pair at 51°. Although it is tempting to use the pair with n_L = 1.45 and n_H = 2.03, this lower index is less than the index of the prism and manifests a greater angle shift, as indicated in the *Shift* column. With increasing angle of incidence, the stop band of a stack with a higher angle shifts away from λ_0 more rapidly, as compared to a stack with smaller angle shift. For that reason, the MacNeille pair with indices of 1.60 and 1.74 is chosen.

As in §6.3.2.3.1, the number of layers in each stack is determined by the required rejection of the s polarization. It is curious that Table 6-3 states that 38 layers are needed to attenuate the s polarization at 51°, yet the caption to Fig. 6-25 shows that only 29 layers are used in that stack. Why? The answer is that in §6.3.3.3 more layers are added to those 29 that also attenuate the s polarized light at λ_0.

Figure 6-25 hints at the angular range over which the "lowest angle" stack is effective at attenuating T_s. The angle shift to shorter wavelengths is quite evident—at 56° the transmittance at 633 nm is dropping rapidly. The result is that if the angles are to range to 71°, additional stacks must be added to attenuate the s polarized light at angles in excess of 56°. This conclusion is substantiated by Fig. 6-26, which plots the stop band (see §5.2.2 and §5.2.3) of the basic period for s polarization. As mentioned in §5.3.4.1.2, within the stop band spectral region the transmittance decreases steadily as more periods are added to the stack. Figure 6-26 makes it evident that, at angles of 57° and greater, the

stop band has shifted so that it does not encompass 633 nm. Therefore, it is futile to add more periods to this stack if the goal is to attenuate at 633 nm at these higher angles.

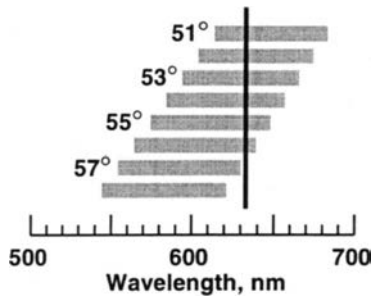


Fig. 6-26 At the indicated angles of incidence (in 1° increments) in glass of index 1.517, spectral width of the stop band for *s* polarized light of the basic period

1.39**B** 1.51**A**

where **A** and **B** are defined in the caption to Fig. 6-25. The vertical line is at $\lambda_0 = 633$ nm.

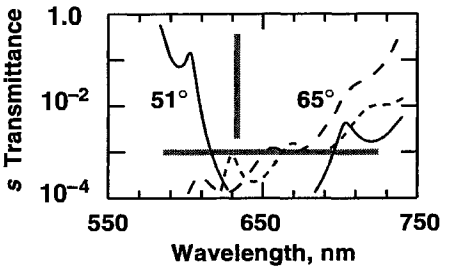


Fig. 6-27 At incidence angles of 51° (*solid*), 59° (*short dash*) and 65° (*long dash*), $T_s(\lambda)$ (on an absorbance scale) of

glass 1.51**A** (1.39**B** 1.51**A**)^{13} 1.39**B** 0.75**A**

0.83**A** 1.26**C** (1.65**A** 1.26**C**)^{7} 0.83**A** *glass*

where the layer indices appear in Table 6-3. Each layer is a multiple of $\lambda_0/4$ in optical thickness at λ_0 of 633 nm. The vertical line is at λ_0. The incidence angles are in *glass* of index 1.517.

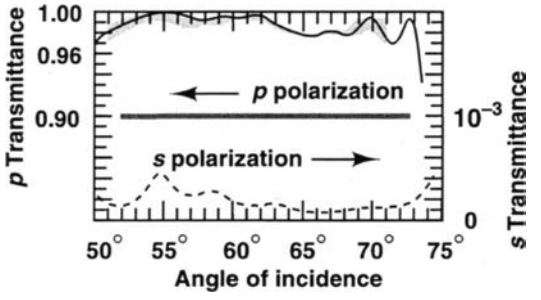


Fig. 6-28 Versus angle of incidence in glass of index 1.517, $T_p(\lambda)$ (*upper panel*) and $T_s(\lambda)$ (*lower panel*) of the 52 layer coating whose refractive index profile appears in Fig. 6-29 and the design in §6.8.1. The *solid* and *dashed* curves are at 633 nm. The shaded area is the envelope of T_p as the wavelength varies by ±1% — from 627 nm to 639 nm.

6.3.3.3 Stacks that function at the higher range of angles

Table 6-3 lists indices of 1.60 (symbol: **A**) and 2.10 (symbol: **C**) — a MacNeille pair for an incidence angle of 57°. When concatenated with the stack captioned in Fig. 6-25 that attenuates at 51° incidence angle, the design becomes

glass 1.51**A** (1.39**B** 1.51**A**)^{13} 1.39**B** 0.75**A**

0.83**A** 1.26**C** (1.65**A** 1.26**C**)^{7} 0.83**A** *glass* .

Figure 6-27 shows that the *s* transmittance in the 633 nm spectral region is relatively low at angles ranging from 51° to 65°. Although not graphed, T_p is relatively high at 633 nm.

The remaining task is to add the "C" stack that contains MacNeille pairs effective at 67°. The indices of 1.74 and 2.34 are the *low* and *high* components, respectively. It is now evident why the symbols **L** and **H** were not used to identify the layers. At 51°, the 1.74 index layer is the *high* component of the MacNeille pair, whereas at 67° it is the *low* index component. Table 6-3 indicates that at 67° the ratio of the *s* effective indices is double that at 51°. The consequence is that the *s* rejection is achieved with only seven layers in the "C" stack, which appears on a separate line below. The design, prior to computer optimization, is

glass 1.51A 1.02(1.36B 1.48A)^{13} 1.36B 0.75A 0.83A 1.26C (1.65A 1.26C)^{7} 0.83A 0.84B (1.25D 1.64B)^{3} *glass* .

Because the stop bands of the "A", "B" and "C" stacks overlap spectrally (for *s* polarized light), the teachings of §5.5.1 are used in constructing the foregoing design. Eighth-wave thickness layers are inserted at the interfaces of the stacks.



Fig. 6-29 Refractive index profile (vs thickness) of the coating captioned in Fig. 6-28.

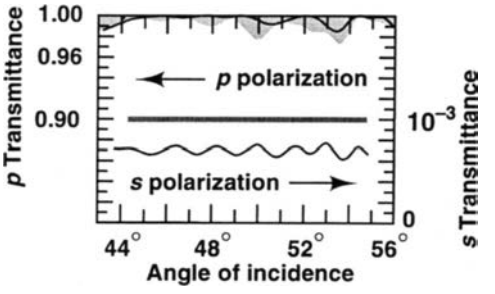


Fig. 6-30 Versus angle of incidence in glass of index 1.73, $T_p(\lambda)$ (*upper panel*) and $T_s(\lambda)$ (*lower panel*) of the coating whose design appears in §6.8.2. The *solid* curves are at 633 nm. The shaded area is the envelope of T_p as the wavelength varies by ±1% — from 627 nm to 639 nm.

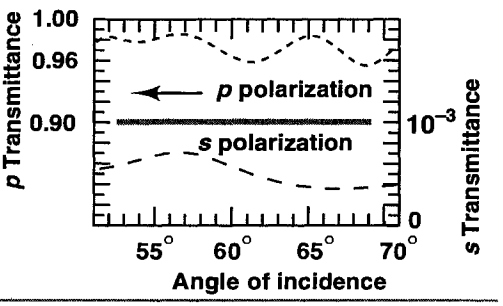


Fig. 6-31 Versus angle of incidence in glass of index 1.517, $T_p(\lambda)$ (*upper panel*) and $T_s(\lambda)$ (*lower panel*) of the 17 layer coating whose design appears in an appendix — §6.8.3. The wavelength is 633 nm.

6.3.3.4 Final result

The foregoing design was optimized using a commercial thin film design package "TF-Calc." [Software Spectra 96]. The targets were a large T_p and low T_s in a spectral band ranging from 613 nm to 653 nm. The range of angles was 51° to 71°. Figure 6-28 illustrates that a reasonably large T_p is maintained, even when the wavelength centering of the coating drifts by ±1%. Figure 6-29 depicts the refractive index profile of the entire coating and §6.8.1 lists the design.

6.3.3.5 Use of higher refractive index glass

Suppose that the foregoing design is deposited upon glass of higher index 1.73. This is the new incident medium. The emergent medium is an optical cement—its index remains at 1.517. Angles of 51° and 71° in a medium of BK-7 translate into incidence angles of 43° and 56° in the higher index glass. The angles of incidence were changed so that L remains constant [see Eq. (2-4a)].

The design was then optimized using a commercial software package [OptiLayer 97]. The computer code added layers. Figure 6-30 depicts the transmittance vs angle of incidence of the final 67 layer coating. The design appears in §6.8.2.

There is a Fresnel reflection of about 2% between an optical cement (of refractive index 1.55) and the glass in the prism (index 1.73). §4.9.7 shows that an antireflection coating can reduce this reflectance to below 0.25% over the visible spectrum and over the complete range of angles.

6.3.3.6 Design produced by computer optimization

The next design was communicated to the author by a purveyor of optical coating software [Noe 96] and is illustrative of computer optimization. The starting design was a single quarterwave optical thickness layer of high index. The computer code then used the "needle" (see §10.2.6.3) and "tunneling" techniques to produce the 17 layer polarizer whose transmittance appears in Fig. 6-31. The coating contains three materials, although only three layers of an intermediate index coating material of index 1.90 are included. The design appears in §6.8.3.

6.3.4 Broadening of the spectral bandwidth

In the situation where the range of angles is relatively small—say, ±1°, most of the methods mentioned in §5.3.6 or §5.5 may be used to broaden the spectral region in which T_s is relatively low. A frequently encountered problem is that the dispersion of the index of the prism and/or the higher index coating materials causes the MacNeille condition [see Eq. (2-39d)] to be violated.

As an example, §5.3.6.1.2 cites a reflector consisting of a quarterwave stack surmounted by a single-cavity Fabry-Perot. That single cavity contains seven layers **H L H L L H L H** and is appended to the quarterwave stack:

glass $(\mathbf{HL})^8$ **H L H L L H L H** *glass* .

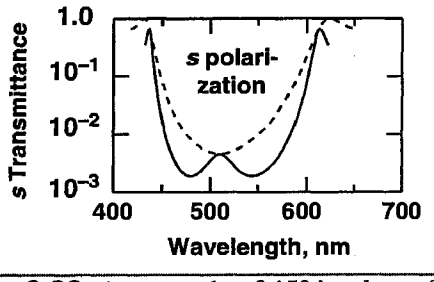


Fig. 6-32 At an angle of 45° in *glass* of index 1.517, $T_s(\lambda)$ of (*solid*)

glass $(\mathbf{HL})^8$ **H L H L L H L H** *glass*

and (*dashed*)

glass $(\mathbf{HL})^8$ *glass* .

The indices of **L** and **H** are 1.38 and 1.705, respectively. Each layer is $\lambda_0/4$ in optical thickness (*matched* at 45°) at λ_0 of 510 nm. T_p is close to 100% and is not plotted.

Figure 6-32 (*solid curve*) shows that the addition broadens the spectral region of low T_s .

6.3.5 Broadening of both the spectral and angular bandwidth

In the spectral region 400 nm to 700 nm, the specification is for a large p transmittance and s transmittance that does not exceed 1%. The incidence angles range from 58° to 64° inside the BK-7 glass. Two designs are considered. One is developed in §6.3.5.1. The other was presented at a conference in 1995 and published later [Li 96]. It is no trivial task to conjure a design that manifests relatively large p transmittance over the entire range of angles.

6.3.5.1 Graded profile design

The concept is that the refractive indices in each stack consist of two sets of MacNeille pairs interleaved with each other, as shown in Fig. 6-33. The **H** layer (n_H = 2.08) and **L** layer (n_L = 1.68) in the basic period, for example, are MacNeille pairs at an angle in the middle angular range of 58° to 64°. The **H** layer and **M** layer (n_M = 1.74) function at a higher range of angles. The starting design contains four such periods to cover the spectral range and the range of angles:

$$1.08(0.5\mathbf{L}\ \mathbf{H}\ \mathbf{M}\ \mathbf{H}\ \mathbf{M}\ \mathbf{H}\ 0.5\mathbf{L})^4\ 1.3(0.5\mathbf{L}\ \mathbf{H}\ \mathbf{M}\ \mathbf{H}\ \mathbf{M}\ \mathbf{H}\ 0.5\mathbf{L})^3$$
$$1.6(0.5\mathbf{L}\ \mathbf{H}\ \mathbf{M}\ \mathbf{H}\ \mathbf{M}\ \mathbf{H}\ 0.5\mathbf{L})^3\ 1.9(0.5\mathbf{L}\ \mathbf{H}\ \mathbf{M}\ \mathbf{H}\ \mathbf{M}\ \mathbf{H}\ 0.5\mathbf{L})^3\ .$$

The layer thicknesses of each of the iterated four periods above are progressively thicker so that each attenuates the *s* polarized light at a longer wavelength.

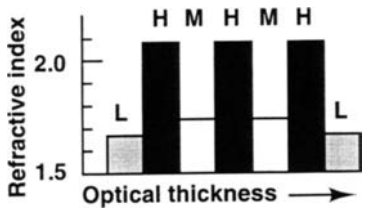


Fig. 6-33 Versus optical thickness, refractive index profile of a basic period of a MacNeille polarizer.

The layer thicknesses of the foregoing starting design were adjusted by an optimization computer code (see §10.2). Figure 6-34 shows the transmittance of the final design containing 98 layers (see §6.8.4). The *s* transmittance is below 0.7%; the *p* transmittance exceeds 94%. In the Li coating, many of the intermediate index layers are relatively thin and are interdispersed between layers of high and low index. Figure 6-35 depicts the spectral transmittance of a coating that is a modification of the Li design (see §6.8.5).

6.3.5.2 Design containing four refractive indices

Although the MacNeille polarizer of Li [96] contains four coating materials and is the result of optimization, it differs from the design in §6.3.5.1.

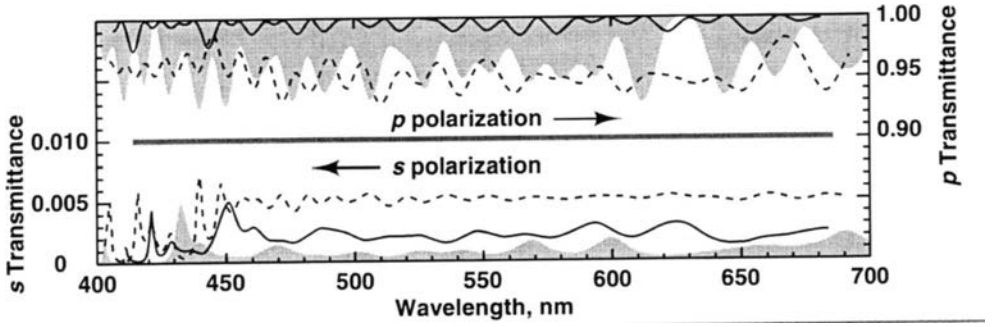


Fig. 6-34 At angles of incidence of 56° (*dashed*), 60° (*solid*) and 64° (*shaded*), $T_p(\lambda)$ (*upper panel*) and $T_s(\lambda)$ (*lower panel*) of a coating containing 98 layers. Its design appears in §6.8.4. The angles of incidence are measured in glass of index 1.517.

6.3.6 Coating that suppresses the *s* reflectance

The coating in this section performs in a manner opposite to those above, in that it transmits the *s* polarized light and reflects the *p*. A *caveat*, however, is that the production of this coating should be undertaken by only the stout hearted. The reason is

that the refractive indices must be controlled with the utmost precision. In fact, this coating might be regarded as merely pedagogical.

6.3.6.1 Design example

The design goal is a p transmittance below 1% at λ_0 of 633 nm. The s transmittance should be close to 100%. The angle of incidence is $\theta_0 = 45°$ in the prism of refractive index $n_0 = 1.60$. Two components of the basic period are a low index layer **L** ($n_L = 1.47$) and high index layer **H** ($n_H = 2.30$) of quarterwave optical thickness, *matched* at θ_0. The 2.30 is a reasonable value for the high index layer in the visible part of the spectrum and §6.3.6.1.1 justifies the choice for this value for the **L** layer.

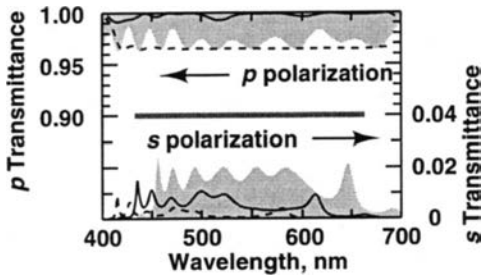


Fig. 6-35 At angles of incidence (in glass of index 1.52) of 50° (*shaded*), 54° (*solid*) and 58° (*dashed*), $T_p(\lambda)$ (*upper panel*) and $T_s(\lambda)$ (*lower panel*) of the coating whose design appears in §6.8.5. Based upon the design of Li [96].

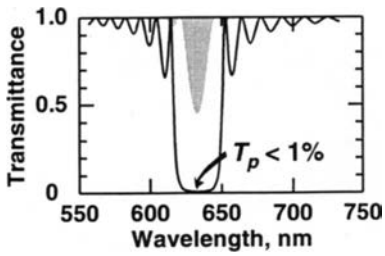


Fig. 6-36 At an angle of incidence of 45° in *glass*, $T_p(\lambda)$ (*solid*) of

glass **S** (0.5**L M H M** 0.5**L**)^{24} **S** *glass*

where the indices of *glass*, **L**, **S**, **M** and **H** are 1.60, 1.47, 1.52, 1.777 and 2.30, respectively. T_s is in excess of 99% and is not plotted. (*Shaded curve*) T_s when the **M** layer index is 1.758. The layers are $\lambda_0/4$ in optical thickness (*matched* at 45°) at λ_0 of 633 nm.

Table 6-4 Effective indices in an *s* transmitting linear polarizer

Symbol	n	N_s	N_p	Description of medium
L	1.470	0.9385	2.302	Low index layer
M	1.777	1.371	2.304	Matching layer: **L** to **H**
H	2.300	2.002	2.642	High index layer
S	1.520	1.015	2.276	Matching layer: **L** to *glass*
glass	1.600	1.131	2.263	Incident, emergent medium

For the refractive index listed in the n column, effective index N_s for s polarization [see Eq. (2-39a)] and effective index N_p for p polarization [see Eq. (2-39b)]. The incidence angle is 45°, as measured in an incident medium of refractive index 1.60.

The s reflectance is suppressed by inserting an **M** layer ($n_M = 1.777$) of intermediate index between each **H** layer and **L** layer. **M** functions as an antireflection coating for s polarization between **H** and **L**. If N_H, N_M and N_L are the effective indices [see Eq. (2-38a)] for the **H**, **M** and **L** layers, respectively, then Eq. (2-23j) requires that

$$N_M = \sqrt{N_H\, N_L} = \sqrt{(2.002)\,(0.9385)} = 1.3707 \ . \qquad (6\text{-}5)$$

The foregoing value is substituted into Eq. (2-38e) and the result is that n_M is 1.777. The optical thickness of **M** is a QWOT, *matched at angle*. As mentioned in §2.6.7 and §5.2.6, the basic period is symmetrical—eighth-wave thicknesses are inserted on the ends. The final design is

glass **S** (0.5**L M H M** 0.5**L**)^q **S** *glass*

where the **S** layers match the effective index of the glass to the effective index of the **L** layers that abut the glass. In other words, the index of the **S** layer is obtained by substituting the effective index of the glass for N_H in Eq. (6-5).

The final task is to find the number of periods, which is $q = 24$ in this example. The transmittance of the 99 layer stack appears in Fig. 6-36. The p transmittance is below 0.01 at the design wavelength.

The next step is to vary the refractive index of the **M** layer by a mere 1%. Figure 6-36 shows that the s transmittance of the coating with the "correct index" **M** layer is close to 100% and that this small drift in the index causes T_s to plunge to 50%. It is for this reason that this type of coating has little to recommend it. It does illustrate, however, that the strong s reflectance may be greatly reduced, as in the nonpolarizing designs of §6.4.

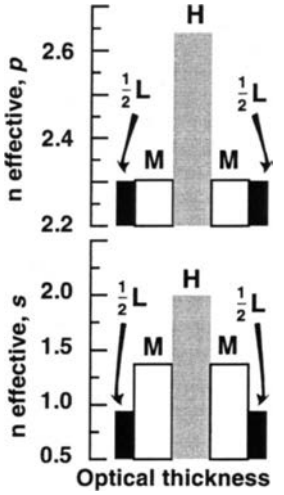


Fig. 6-37 Versus optical thickness, effective index for p polarization (*upper panel*) and s polarization (*lower panel*) of

0.5**L M H M** 0.5**L** .

The angle of incidence, incident medium refractive index and meaning of **L**, **M** and **H** appear in the caption to Fig. 6-36.

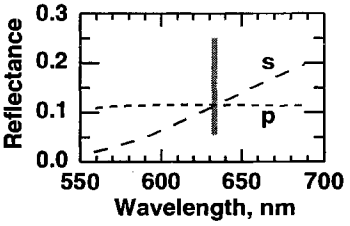


Fig. 6-38 At an incidence angle of 60°, $R_s(\lambda)$ (*long dash*) and $R_p(\lambda)$ (*short dash*) of a nonpolarizing reflector of the design

air 0.98**L** 1.20**H** *glass*

where the refractive indices of **L**, **H** and *glass* are 1.40, 2.40 and 1.50, respectively. The optical thicknesses of **L** and **H** are multiples of $\lambda_0/4$. The vertical line is at $\lambda_0 = 633$ nm. After Cordray [73].

The next step is to vary the refractive index of the M layer by a mere 1%. Figure 6-36 shows that the *s* transmittance of the coating with the "correct index" M layer is close to 100% and that this small drift in the index causes T_s to plunge to 50%. It is for this reason that this type of coating has little to recommend it. It does illustrate, however, that the strong *s* reflectance may be greatly reduced, as in the nonpolarizing designs of §6.4.

6.3.6.1.1 Choice of refractive index

There is a reason for the choice of the refractive indices for the L and M layers. Both Fig. 6-37 and Table 6-4 show that their effective indices (for *p* polarization) are nearly the same. The advantage is that this produces a larger *p* reflectance. Why? In the interior of the stack there are sequences of layers such as MLM. The foregoing three-layer sequence is essentially a single three quarterwave thick layer, as far as the *p* polarized light is concerned. This is evident in Fig. 6-37, which plots the effective indices. The stack then consists of a series of three-quarterwave thickness layers interdispersed between H layers. It is this "high-low" QW-3QW stack that reflects the *p* polarized light.

6.4 Nonpolarizing nonimmersed designs

6.4.1 Designs containing layers of nonquarterwave optical thickness

The advantage of the quarterwave design in §6.4.2 and other sections is that its reflectance is relatively insensitive to changes in layer thickness. For example, Fig. 6-38 depicts the reflectance of a nonpolarizing coating composed of two dielectric layers in which the thicknesses are not quarterwaves [Cordray 73]. The R_p and R_s curves intersect at an oblique angle. For this reason, its reflectance is quite sensitive to both changes in refractive index and perturbations in layer thickness.

6.4.2 Quarterwave layer thickness designs — introduction

The objective is to produce a nonpolarizing design that functions over a relatively narrow bandwidth centered around its tuned wavelength λ_0. As in §6.2, the phase thickness β of each layer is 90°, *matched at all angles.*

The design LHLH (captioned in Fig. 6-19) is nonpolarizing at an incidence angle θ_n of 63°. Its reflectance is a scant 35%, however. So the question remains, are there stacks that manifest a higher reflectance and at an angle of incidence closer to 45°? §6.4.4 shows that the answer is yes, provided infrared coating materials are used. Before delving into this subject, however, §6.4.3 addresses the question: Given a quarterwave stack, does a polarizing angle exist?

6.4.2.1 Analysis using reduced admittance

The performance of the foregoing designs could be described in terms of vector diagrams, as in §6.2.2.1. This approach is quite laborious, however, because lengths of the vectors continually change with angle. Hence, it is no easy task to find the angle at which $|\hat{\rho}_p|$ is equal to $|\hat{\rho}_s|$. For this reason, the performance vs angle of these coatings is best analyzed by using the reduced admittance.

6.4.2.1.1 Existence of a nonpolarizing angle

Graphs of $W(\theta_0)$ appear in Figs. 6-17 and 6-19 for quarterwave stacks. The former stack, whose design is HLH, is not nonpolarizing at *any* angle, whereas the stack LHLH manifests an angle at which it is nonpolarizing. The question is, can it be proven or

disproven that any given design is nonpolarizing at a particular angle? An appendix (see §6.7.1) derives the parameter D for a quarterwave stack. An angle exists at which it is nonpolarizing, if D is negative.

This is illustrated by **LHLHL**—a stack with an odd layer count. Assume that $n_0 = n_s = 1.517$ and n_L and n_H are the same as in Fig. 6-16. D is given by Eq. (6-11d)

$$D = 0.5\ (1.517)^{-2} + 0.5 + 2\ (2.25)^{-2} - 3\ (1.45)^{-2} = -0.315\ . \qquad \text{(6-6a)}$$

The negative value of D implies that an angle θ_n exists at which the coating is nonpolarizing at λ_0—the wavelength at which the layer thicknesses are "tuned" to be quarterwaves.

The right of Eq. (6-6a) only hints at the value of θ_n, however. In general, the larger the absolute value of D, the closer θ_n is to grazing incidence. The exact angle θ_n may be obtained by either of two methods: (1) The left of Eq. (6-11a) is set equal to 1.00 and it is solved as a polynomial in L^2; or (2) an iterative numerical procedure, such as the Newton-Raphson method, may be used.

6.4.3 Analytical solution for a coating containing two refractive indices

Our focus is on a stack with a low index outer layer and even layer count, such as

air **(L H)**^{q} *glass*

where q is the number of periods. Each layer has an optical thickness such that at λ_0 its phase thickness β is 90°, *matched at all angles.*

6.4.3.1 Beamdivider for the near IR

An example is a beamdivider that achieves a $R = 50\%$ and $T = 50\%$ division ratio at 45° incidence at λ_0 of 2.0 μm. Air is the incident medium and the substrate is sapphire of index 1.73. The reflectance of 50% is substituted into Eq. (2-19c), which results in a standing wave ratio of 5.83. The left of Eq. (2-48d) should be the reciprocal of 5.83

$$(5.83)^{-1} = n_s\ n_0^{-1}\ n_L^{2q}\ n_H^{-2q} = U \qquad \text{(6-6b)}$$

and setting $W = 1$, Eq. (2-48e) becomes

$$1 = \frac{\cos\theta_s\ (1 - L^2\ n_L^{-2}\ )^q}{\cos\theta_0\ (1 - L^2\ n_H^{-2}\ )^q} \qquad \text{(6-6c)}$$

where L appears in Eq. (2-4a). Given that θ_0, n_0 and n_s are specified, the designer may choose a value for q. An appendix (§6.7.5) provides solutions to the foregoing equations in terms of the indices of the **L** and **H** layers.

For $q = 2$ (i.e., four layers) a coating on a sapphire substrate is

air **L H L H** *sapphire*

where n_L is 1.737 and n_H is 3.096. The thickness of each layer is $\lambda_0/4$, *matched* at 45°. The remaining question is, how are the foregoing refractive indices produced? Silicon monoxide has an index of 1.95 when it is evaporated slowly and with little oxygen in the chamber. A lower index of 1.737 is obtained by increasing the oxygen pressure as the SiO is deposited. Equivalent layers consisting of zinc sulfide ($n = 2.25$) and silicon ($n = 3.50$) are used to synthesize the **H** layers. An *lhl* simulation is used for one equivalent layer and an *hlh* simulation for the other. As discussed in §5.3.5.3, the inclusion of both *lhl* and *hlh* equivalent layers compensates for the dispersion of the equivalent layer's index. Figure 6-39 depicts the reflectance of a beamdivider that is nonpolarizing at $\lambda_0 = 2.00$ μm.

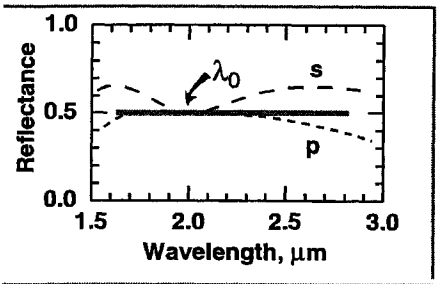


Fig. 6-39 $R_s(\lambda)$ (*long dash*) and $R_p(\lambda)$ (*short dash*) at an incidence angle of 45° of a nonpolarizing beamdivider at $\lambda_0 = 2.0$ μm. Its design is

air **L** 0.233**M** 0.5**H** 0.233**M**

L 0.405**H** 0.174**M** 0.405**H** *sapphire*

where the refractive indices of **L**, **M**, **H** and *sapphire* are 1.737, 2.25, 3.50 and 1.73, respectively. The optical thicknesses of **L** and **H** are multiples of $\lambda_0/4$, *matched* at 45°.

6.4.3.1.1 Analysis using vector addition of amplitudes

Insight into the performance of the aforementioned coating is provided by vector addition of amplitudes. Figure 6-40 illustrates that its vector diagram is similar to that of the four-layer laser reflector in Fig. 5-5 . In both instances, the vectors from the interfaces of the outer layer are in opposite directions. In fact, the vector produced at the air interface is contrary in direction to the vectors from all of the four other interfaces.

Consider the Fresnel amplitude reflection coefficients as the incidence angle shifts from normal to 45°. The largest change is in the magnitude of r_0. Figure 6-40 shows that, with increasing incidence angle, r_0 "grows" for *s* polarization. Because the direction of r_0 is contrary to that of the other vectors, the net effect of this "growth" is to compensate for the growth of the other vectors. In a similar manner, r_0 "shrinks" for *p* polarization as the incidence angle increases. The greater shrinkage of r_0 compensates for the shrinkage of the other vectors. Figure 6-40 shows that the net amplitude reflection coefficient is virtually the same for *p* polarization and *s* polarization.

6.4.3.2 Beamdividers that function at various angles and division ratios

The question remains: Are there combinations of layers that produce a 50%-50% division (between *R* and *T*) at 45°? The refractive indices should be attainable in the visible part of the spectrum. The problem is that the **H** layer index is usually too large to be attainable, as illustrated in Table 6-5 by entry of $n_H = 2.8$ and $n_L = 2.05$ when $q = 4$ (i.e., eight layers). Mahlein [74b] reached the same conclusion. An alternative course is to select another division ratio or different incidence angle. Table 6-5 shows that coatings can be produced in the visible part of the spectrum with a 50%-50% division at 60° or with a $R = 30\%$ division at 45°.

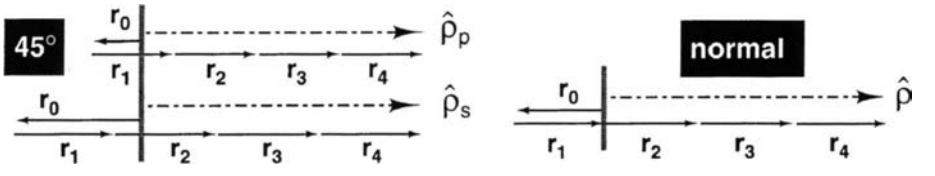


Fig. 6-40 The net amplitude reflection coefficient $\hat{\rho}$ (*dashed*) is found by vector addition of amplitudes for the design

air **L H L H** *sapphire*

where the refractive indices of **L**, **H** and *sapphire* are 1.737, 3.096 and 1.73, respectively. (*Upper left*) θ_0 is 45°, *p* polarized flux. (*Lower left*) θ_0 is 45°, *s* polarized flux. (*Right panel*) Normal incidence. The phase thickness of each layer is 90°, at normal and at 45° incidence. The *heavy shaded* vertical line is the imaginary axis at the origin.

Table 6-5 Nonpolarizing beamdividers

Design	θ_0	n_0	n_s	q	R	T	n_L	n_H	n_M
Two	45°	1.00	1.517	3	30%.	70%	1.77	2.33	unused
Two	45°	1.00	1.517	4	50%.	50%	2.05	2.80	unused
Two	60°	1.00	1.517	4	50%	50%	1.77	2.33	unused
Two	50°	1.00	1.517	4	33%	67%	1.91	2.37	unused
Three	45°	1.60	1.60	4	50%	50%	1.381	2.31	1.60
Three	45°	1.60	1.60	5	75%	25%	1.35	2.374	1.57
Three	45°	1.60	1.60	6	75%	25%	1.38	2.31	1.60
Three	45°	1.60	1.60	7	75%	25%	1.35	2.13	1.545
Three	45°	1.60	1.60	7	75%	25%	1.38	2.20	1.585
Three	45°	1.60	1.60	8	75%	25%	1.45	2.28	1.674

At a wavelength λ_0 and incidence angle θ_0 measured in an incident medium of index n_0, designs of beamdividers that manifest the same reflectance R and transmittance T for both p polarized and s polarized light. For two-component systems the design is

incident $(\mathbf{L}\,\mathbf{H})^q$ *substrate*

and for *three* components

incident $(\mathbf{L}\,\mathbf{M}\,\mathbf{H}\,\mathbf{M})^q$ *substrate* .

Listed above are the index n_s of the substrate, number of periods q, index n_L of the **L** layer, index n_H of the **H** layer and index n_M of the **M** layer (if used).
All layers have an optical thickness of $\lambda_0/4$, *matched* at θ_0 .

6.4.4 Designs containing three refractive indices

Thelen [78] proposed the use of three coating materials in a beamdivider that is nonpolarizing at a tuned wavelength λ_0. The design is

air $(\mathbf{M}\,\mathbf{H}\,\mathbf{M}\,\mathbf{L})^5\,\mathbf{M}$ *glass*

where the **L**, **M** and **H** layers have an optical thickness $\lambda_0/4$, *matched at all angles.* The goal of the example below is to produce an "air to glass" 50%-50% beamdivider.

Given the target reflectance R, angle of incidence, substrate index and incident medium index, the design procedure is similar to that of §6.4.3.1: (1) Compute the standing wave ratio from Eq. (2-19c). This is 5.83 for R = 50%. (2) Equate the reciprocal of step (1) to

$$U = (5.828)^{-1} = \frac{n_M^{4q+2}}{n_H^{2q}\, n_L^{2q}\, n_0\, n_s} \tag{6-7a}$$

and

$$W = 1.00 = \frac{(1 - L^2\, n_M^{-2})^{2q+1}}{(1 - L^2\, n_H^{-2})^q\, (1 - L^2\, n_L^{-2})^q \cos\theta_s \cos\theta_0} \,. \tag{6-7b}$$

(3) Assign a value to the index n_M of the **M** layer. (4) Following the procedure of §6.7.2, solve for n_L and n_H as the number of periods q is varied.

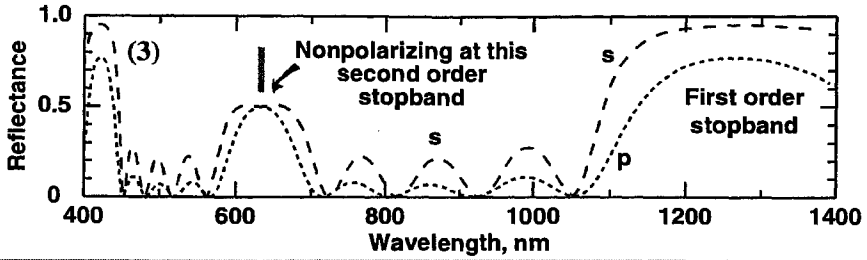


Fig. 6-41 $R_s(\lambda)$ (*long dash*) and $R_p(\lambda)$ (*short dash*) at an incidence angle of 45° of a beamdivider that is nonpolarizing at $\lambda_0 = 633$ nm. Its design is

air (**M H M L**)^5 **M** *glass*

where the refractive indices of **L**, **M**, **H** and *glass* are 1.43, 1.60, 2.25 and 1.517, respectively. The optical thickness of each layer is $\lambda_0/4$, *matched* at 45°.

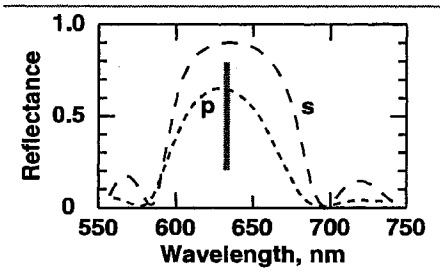


Fig. 6-42 $R_s(\lambda)$ (*long dash*) and $R_p(\lambda)$ (*short dash*) at an incidence angle of 45° of

air 0.421**L** 0.144**H** 0.421**L**
(0.421**L** 0.144**H** 1.421**L** 0.211**H** 0.548**L**
1.211**H**)^5 *glass*

where the refractive indices of *glass*, **L** and **H** are 1.517, 1.442 and 2.26, respectively. **H** and **L** are of optical thickness $\lambda_0/4$ at λ_0 of 633 nm, *matched* at 45°.

Table 6-5 lists results of the aforementioned design method. Figure 6-41 graphs the performance of a 50%*R*-50%*T* reflector. This illustrates that R_s and R_p are equal over a limited spectral bandwidth.

6.4.4.1 Higher order reflection peak

§5.2.3.4 defines *order number* of a stop band as the number of halfwaves that "fit" in the optical thickness of the basic period. The basic period of the four-layer group **MHML** in §6.4.4 is second order at $\lambda_0 = 633$ nm, where it is nonpolarizing and reflects 50%. Figure 6-41 shows that the first-order reflection peak in the IR and the third-order reflection peak in the blue spectral region both are more than 50% and manifest strong polarization splitting.

Why is this important? Suppose that the goal is a 50% reflectance, nonpolarizing coating in the spectral range 400 nm to 700 nm. The starting design consists of several **MHML** stacks, each located at a different wavelength. The problem is the polarization splitting of the third-order reflection peak in the blue spectral region. Any computer optimization that attempts to reduce this polarization splitting in the blue surely affects the reflectance of the design in the 600 nm range.

6.4.4.2 Use of equivalent layers and computer optimization

It is shown in §2.6.7 that an equivalent layer consisting of a group of three layers may be substituted for a layer of intermediate refractive index, such as the **M** layers in the foregoing design.

This substitution achieves limited success, as shown in Fig. 6-42. At the tuned wavelength of $\lambda_0 = 633$ nm, both the *s* and *p* reflectances exceed the target of 50%. An

optimization type of computer code (see §10.2) was unleashed to reduce this spread. The goal was to achieve 50% reflectance for both linear polarizations in the spectral region 613 nm to 643 nm. The number of layers was reduced from 32 to 17. The reflectance of the final result, which is depicted in Fig. 6-43, shows that the R is within 0.5% of the target value.

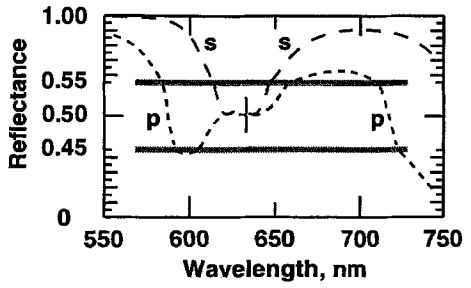


Fig. 6-43 $R_s(\lambda)$ (*long dash*) and $R_p(\lambda)$ (*short dash*) at an incidence angle of 45° in *air* of a nonpolarizing reflector at λ_0 of 633 nm. The design, containing 17 layers, appears in §6.8.6. The incident medium is air and the substrate index is 1.517. The ordinate changes scale at 0.45 and 0.55.

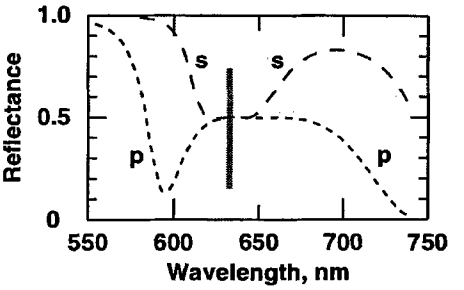


Fig. 6-44 $R_s(\lambda)$ (*long dash*) and $R_p(\lambda)$ (*short dash*) at an incidence angle of 45° in *air* of a nonpolarizing reflector at λ_0 of 633 nm. The design, which contains 18 layers, appears in §6.8.7. The incident medium is *air* and the substrate index is 1.517. The vertical line is at λ_0.

As another example, a high index (n = 2.26) layer of 1.4 μm thickness was the starting design for the *needle optimization* (see §10.2.6.3) using the commercial software package OptiLayer [97]. The resulting coating, whose $T(\lambda)$ appears in Fig. 6-44, is dissimilar to that of the foregoing paragraph.

The important question is, of course, whether either of the aforementioned designs could be produced. There are unanswered questions such as: What is the sensitivity of the reflectance to layer thickness errors and to drift in the refractive index during deposition?

6.5 Nonpolarizing immersed coatings

6.5.1 Introduction

For the reasons stated in §6.4.1, our focus is on quarterwave designs. The resultant coatings are nonpolarizing only at a rather narrow spectral bandwidth centered at λ_0. The broadening of the spectral bandwidth is discussed in §6.5.5.

There are several methods of designing these coatings. The focus of §6.5.2 is a coating that is nonpolarizing at all angles. The designs considered previously are nonpolarizing at only one angle (see §6.5.3).

6.5.2 Nonpolarizing—independent of angle

Consider an all-dielectric stack in which the optical thickness of each layer is $\lambda_0/4$, *matched at all angles*. A useful theorem (that follows from §6.2.1.1.1) is:

> A stack that is nonreflecting at normal incidence is nonpolarizing at all angles.

An equivalent form of the foregoing is: A stack that manifests a reduced admittance of 1.00 at normal incidence is nonpolarizing at all angles. This is true only at a wavelength of λ_0.

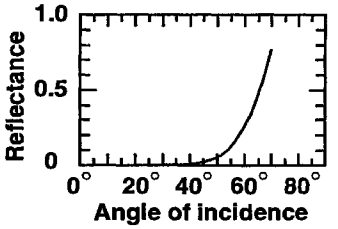


Fig. 6-45 Versus the angle of incidence in *glass* at a wavelength λ_0, $R(\lambda)$ (for both *p* and *s* polarization) of
glass **MHMLMHM** *glass*
where the refractive indices of *glass*, **L**, **M** and **H** are 1.517, 1.45, 1.847 and 2.30, respectively. The optical thickness of each layer is $\lambda_0/4$, *matched at all angles.*

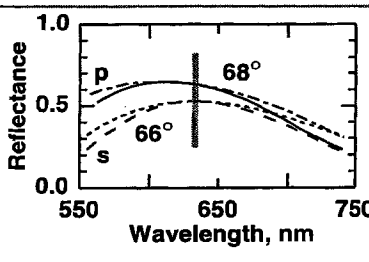


Fig. 6-46 At 66° incidence in glass of index 1.517, $R_s(\lambda)$ (*long dash*) and $R_p(\lambda)$ (*short dash*) of the design captioned in Fig. 6-45. The $R_s(\lambda)$ curve (computed at an incidence angle of 68°) is *solid* and the $R_p(\lambda)$ is *long-short dash*. The optical thickness of each layer is $\lambda_0/4$ at λ_0 of 633 nm, *matched at* 66°.

6.5.2.1 Design procedure

The design procedure is to find a system of quarterwave layers that is nonreflecting at normal incidence, tip it to a certain angle (while maintaining the phase thicknesses of every layer at 90°) and, voilà, the coating manifests equal reflectance for *p* and *s* polarization. Whether it has the desired reflectance at that angle is another matter. As an illustration, Fig. 6-15 depicts $R(\theta_0)$ of a single-layer antireflection coating on a germanium substrate. It reflects R_p and R_s equally at all angles of incidence.

Thus it is necessary to concoct a quarterwave stack in which the reduced admittance at normal incidence (and at $\lambda = \lambda_0$) is 1.00. Suppose that the design
glass **MHMLMHM** *glass*
is selected, with $n_H = 2.30$, $n_L = 1.45$ and a *glass* index of $n_0 = 1.517$. The index n_M of the **M** layer is yet to be determined. From Eq. (2-48d) its reduced admittance is

$$U = n_M^8 n_L^{-2} n_H^{-4} n_0^{-2} \quad . \qquad (6\text{-}7c)$$

Setting U equal to 1.00 and solving for n_M provides

$$n_M = [(1.517)^2 \ (1.45)^2 \ (2.30)^4 \ ]^{1/8} = 1.847 \quad . \qquad (6\text{-}7d)$$

The index of 1.847 is close to that of yttria, yet it is difficult to achieve. The reflectance of the foregoing coating vs angle of incidence appears in Fig. 6-45. It is concluded that the coating is usable if a division ratio of 10% *R* and 90% *T* is needed. If a 50%-50% division is required, however, the coating is impractical because its reflectance is excessively sensitive to changes in the angle of incidence. This is illustrated in Fig. 6-46, which shows that a 2° increase in the angle of incidence (from 66° to 68°) shifts the reflectance from 52% to 67%.

6.5.3 Nonpolarizing at a single angle

Thelen [76] showed that the use of three components produces a nonpolarizing beamdivider by virtue of the fact that $W = 1.00$ at a particular angle. §6.5.3.1 describes a

synthesis design method. §6.5.3.2 utilizes the concept of effective index to provide insight into the behavior of such beamdividers.

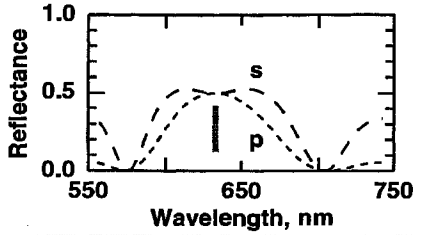


Fig. 6-47 $R_s(\lambda)$ (*long dash*) and $R_p(\lambda)$ (*short dash*) at an incidence angle of 45° (measured in *glass*) of

glass (**HMLM**)^5*glass*

where the refractive indices of **L**, **M**, **H** and *glass* are 1.454, 1.70, 2.37 and 1.517, respectively. Layer optical thicknesses are $\lambda_0/4$ at λ_0 of 633 nm, *matched* at 45°.

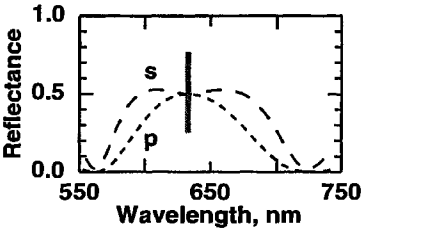


Fig. 6-48 $R_s(\lambda)$ (*long dash*) and $R_p(\lambda)$ (*short dash*) at an incidence angle of 45° (measured in *glass*) of

glass (**HMLM**)^4*cement*

where the refractive indices of **L**, **M**, **H** *glass* and *cement* are 1.429, 1.62, 2.328, 1.73 and 1.517, respectively. Layer optical thicknesses are $\lambda_0/4$ at λ_0 of 633 nm, *matched* at 45°.

6.5.3.1 Synthesis of designs

The goal is a beamdivider with a 50% division between R and T at 45° incidence. Setting W equal to 1.00 makes the beamdivider nonpolarizing. The variable U is obtained from Eq. (2-19c) and equals $3 + \sqrt{8} = 5.83$ (or its reciprocal).

The basic design is

glass (**HMLM**)^q *glass*

where the **L**, **M** and **H** layers have an optical thickness $\lambda_0/4$, *matched at all angles.* The exponent of the stack q is yet to be determined.

Following the approach of §6.4.4, the value of U [from Eq. (2-48d)] is

$$U = \frac{n_s\, n_H^{2q}\, n_L^{2q}}{n_0\, n_M^{4q}} = 5.828 \qquad \text{(6-8a)}$$

and setting W equal to 1.00 establishes that

$$W = \frac{\cos\theta_s \left(1 - L^2 n_H^{-2}\right)^q \left(1 - L^2 n_L^{-2}\right)^q}{\cos\theta_0 \left(1 - L^2 n_M^{-2}\right)^{2q}} = 1.00\ . \qquad \text{(6-8b)}$$

The indices of the substrate and incident medium are retained in the foregoing equations, even though they cancel in most cases because those indices are identical. An appendix (§6.7.3) shows that the foregoing equations may be solved for two of the indices.

6.5.3.1.1 Reflector with 50% reflectance and 50% transmittance

The steps in the design procedure are: Values for n_M and q are assumed. Equations (6-7a) and (6-7b) are then solved for n_L and n_H , using the procedure of §6.7.3. The solution is translated into a design if attainable values for the refractive indices result. Figure 6-47 shows the $R(\lambda)$ of a beamdivider containing 20 layers. The spectral region in which the reflectance attains 50% is centered at $\lambda_0 = 633$ nm. The extent of this region is from $0.95\ \lambda_0$ to $1.05\ \lambda_0$.

The coating whose $R(\lambda)$ appears in Fig. 6-48 illustrates that the 20 layers in Fig. 6-47 can be reduced to 16 if a higher refractive index incident medium is used.

6.5.3.1.2 Reflector with 75% reflectance and 25% transmittance

The design procedure outlined above is quite general and may be used to produce a beamdivider with virtually any division ratio. For example, if a beamdivider with an R of 75% and T of 25% were required, then substitution into Eq. (2-19c) shows that the U should be 13.9. The procedure of §6.7.3 is used to find n_L and n_H, which are the indices of the **L** and **H** layers, respectively. The reflectance of a stack with seven periods appears in Fig. 6-49.

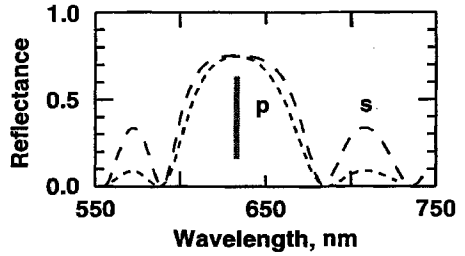


Fig. 6-49 $R_s(\lambda)$ (*long dash*) and $R_p(\lambda)$ (*short dash*) at an incidence angle of 45° (measured in *glass*) of

glass $(\mathbf{H}\,\mathbf{M}\,\mathbf{L}\,\mathbf{M})^7$ *glass*

where the refractive indices of **L**, **M**, **H** and *glass* are 1.413, 1.650, 2.26 and 1.517 respectively. All optical thicknesses are $\lambda_0/4$ at λ_0 of 633 nm, *matched* at 45°.

Table 6-5 lists solutions for several numbers of periods q. The index n_L ranges from 1.35 to 1.45. The span of n_H is from values as low as 2.13 to 2.37 on the high end, depending upon the number of periods.

6.5.3.2 Analysis using effective indices

The foregoing section offers a formal procedure for finding the refractive indices of a nonpolarizing quarterwave stack, but no intuitive view of how such a coating functions. The use of effective indices (see §2.5.3) provides the desired insight.

6.5.3.2.1 Analysis of a stack containing two refractive indices

Consider first the effective indices of an eight-layer "high-low" stack $(\mathbf{H}\,\mathbf{L})^4$, where **L** and **H** represent layers of optical thickness $\lambda_0/4$, *matched at angle*. The angle of incidence is 45° in glass of index 1.73. The refractive indices of **L** and **H** approximate those of silica and titania.

The effective indices appear in Table 6-6 and are graphed in Fig. 6-50. The effective indices for p polarization are quite close together and, consequently, the Fresnel amplitude reflection coefficient is small at each interface between an **H** and an **L** layer. It follows that the p reflectance of the entire stack is quite low.

Table 6-6 Effective indices of the layers of a nonpolarizing beamdivider

Symbol	n	N_s	N_p	Description of medium
L	1.429	0.7381	2.7656	Low index layer
M	1.620	1.0621	2.4711	Intermediate index layer
H	2.328	1.9807	2.7362	High index layer

For the refractive index listed in the n column, effective index N_s for s polarization [see Eq. (2-39a)] and effective index N_p for p polarization [see Eq. (2-39b)]. The incidence angle is 45°, as measured in an incident medium of refractive index 1.73.

On the other hand, there is an immense spread between the effective indices for *s* polarization. The result is that the *s* reflectance approaches 100%, even when just a few periods are used. A consequence is that the stack is strongly polarizing.

6.5.3.2.2 Analysis of a stack containing three refractive indices

§6.5.3.2.1 illustrates the futility of producing a nonpolarizing beamdivider using a quarterwave stack containing two components. The insertion of the **M** layers of intermediate index reduces the polarization splitting. §6.3.6 teaches that the reduction of the *s* reflectance to zero is accomplished by inserting an intermediate index **M** layer of quarterwave optical thickness between each **H**-**L** pair. This **M** layer functions as an antireflection coating for *s* polarization.

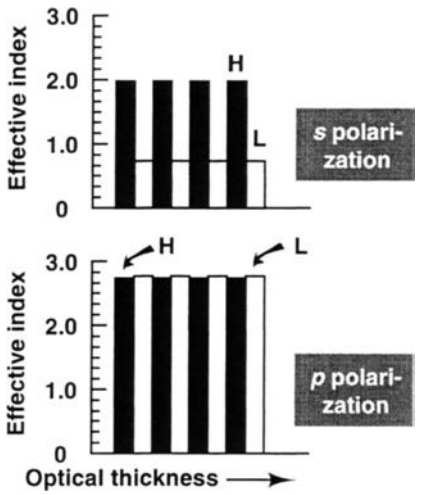


Fig. 6-50 The design of the stack is $(\mathbf{H}\,\mathbf{L})^4$.

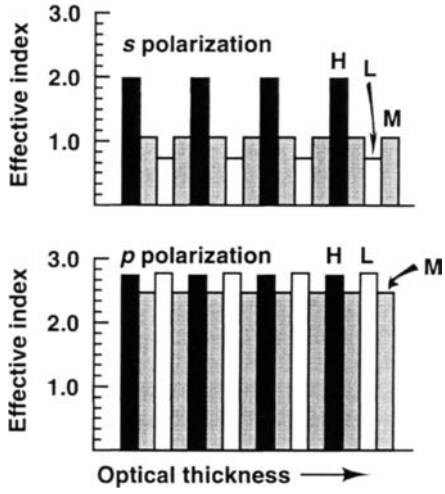


Fig. 6-51 The design of the stack is $(\mathbf{H}\,\mathbf{M}\,\mathbf{L}\,\mathbf{M})^4$.

Versus optical thickness, effective index for *s* polarization (*upper panels*) and *p* polarization (*lower panels*) of the above designs at an angle of incidence is 45° in a medium of index 1.73. The refractive indices of **L**, **M** and **H** are 1.429, 1.62 and 2.328, respectively.

When this coating is used as a reflector with equal *s* and *p* reflectance, however, the objective is to *reduce* the *s* reflectance, not to suppress it completely, as is done in §6.3.6. The **M** layers in the design

glass **H M L M H M L M H M L M H M L M** *cement*

may be viewed as single-layer antireflection coatings that *partially* dereflect (for *s* polarization) the interface between the **L** and **H** layers. The effective indices of the **L**, **M** and **H** layers appear in Table 6-6 and are graphically presented in Fig. 6-51.

Another important effect is that the *p* reflectance is enhanced because the effective index of an **M** layer is *lower* than that of either an **L** layer or an **H** layer, as is evident in the lower panel of Fig. 6-51. That is, even though the *actual* index (n_M = 1.62) of the **M** layers is intermediate to the **L** layer (n_L = 1.429) and the **H** layer (n_H = 2.328), the *effective indices* of both the **H** and **L** layers *exceed* that of the **M** layer. Consider a sequence of layers **H M L M** whose indices vary as "high-intermediate-low-intermediate"

at normal incidence. At 45° those effective indices for p polarization vary as "high-low-high-low." This functions as a quarterwave stack (see §5.2.4.3.1) that boosts the p reflectance.

6.5.4 Tolerances of the refractive index and layer thickness

How do changes in layer thickness and/or refractive index alter the performance of the foregoing designs? §5.1.4.2 points out that the quarterwave stack provides the optimum reflectance for a given number of layers. The multilayer has the feature that, at the tuned wavelength of λ_0, its reflectance R is stationary—i.e., a minimum or maximum. It can be shown that fractional errors (dh/h) in layer thickness affect the reflectance R at λ_0 proportional to $(dh/h)^2$. A consequence is that errors in layer thickness have a relatively small effect upon the reflectance.

Errors in refractive index, however, can affect R at λ_0 quite dramatically. An approximation is used to estimate the effects of such index changes. The R at λ_0 of the seven-period design listed in the caption to Fig. 6-49 serves as an illustration. Its reflectance is 75% and, because the stack is nonabsorbing, its transmittance T ($T \equiv 1 - R$) is 25%. The standing wave ratio V of this coating is 13.9 at λ_0, as cited in §6.5.3.1.2. For a V of this magnitude, the approximation in Eq. (2-19f) may be used

$$T \approx 4/V \ . \tag{6-9a}$$

Suppose that V is increased by dV due to a change of index. What is the change dT that occurs in T? It follows from Eq. (6-9a) that a 1% increase in V (example: 13.9 → 14.0) produces the same fractional change in T (example: 0. 25 → 0.248)

$$dT/T \approx -dV/V \ . \tag{6-9b}$$

It is true, to first order, that

$$dV/V \approx dU/U \ . \tag{6-9c}$$

Although the foregoing neglects the change in the W function (see §2.7.3.3) as the refractive index is varied, it still suffices.

There are 14 M layers in the stack and with $q = 7$ substituted into in Eq. 6-8a, it follows that

$$U \propto n_M^{-28} \ . \tag{6-9d}$$

The foregoing equation shows that a 1% *decrease* in the index of all M layers increases V by 28%. Substituting this fractional change in V into Eq. (6-9b), it is concluded that a 28% fractional decrease in T (from 0.25 to 0.195) is produced by the 1% drift in refractive index. Although the foregoing change is large enough that it undoubtedly violates our assumption of the validity of a linear theory, it nevertheless illustrates that the reflectance of this design is quite sensitive to drift in the refractive indices.

6.5.5 Comparison of beamdividers: all-dielectric and containing silver

An example illustrates that a beamdivider containing a silver layer may be preferable if the range of angles is large. Its performance is compared to that of an all-dielectric beamdivider. At a central angle of 45°, the objective is a reflectance in the range of 0.45 to 0.55 for both p and s polarization in the spectral range 420 nm to 680 nm.

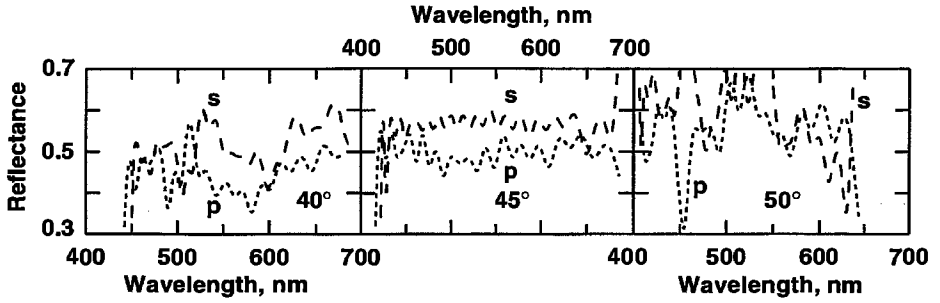


Fig. 6-52 $R_s(\lambda)$ (*long dash*) and $R_p(\lambda)$ (*short dash*) at incidence angles of 40° (*left panel*), 45° (*center panel*) and 50° (*right panel*) of a beamdivider with minimal polarization splitting. The foregoing angles are measured in the glass of index 1.517. The coating design appears in §6.8.8.

6.5.5.1 All-dielectric beamdivider

The starting design consisted of a Fabry-Perot. Two nonpolarizing reflectors — similar to the design shown in Fig. 6-48— were separated by a thick spacer layer. Computer optimization produced the 105 layer beamdivider whose design appears in §6.8.8. The $R(\lambda)$ of the coating appears in Fig. 6-52 at the design angle of 45° and at angles of 45°±5° inside the glass.

The problem is that the coating contains relatively low index layers, such as $n_L = 1.38$. The angle shift in such layers is larger than in higher index layers. Manifestations of the angle shift are evident in the *right panel* of Fig. 6-52. The splitting ratio deviates considerably from 50% at wavelengths in excess of 645 nm. At a 40° angle of incidence, the reflectance plunges to low values in the blue spectral region.

6.5.5.2 Beamdivider containing silver

The starting design was a single silver layer of 25 nm thickness that was sandwiched between stacks of relatively thin layers. The stacks consisted of alternate layers of index 2.25 and 1.62. These higher indices were used in an effort to reduce the angle shift. The computer optimization code was not permitted to alter the thickness of the silver layer. After optimization, 41 layers remained.

Figures 6-53 and 6-54 depict the reflectance and transmittance of this stack at the extreme angles of 40° and 50° (angles measured inside the glass). Although this beamdivider has an *RT* product [see §1.8.1.2(g)] of approximately 0.21, at the higher and lower angles of incidence the beamdivider with the silver-based beamdivider performs better than the all-dielectric design.

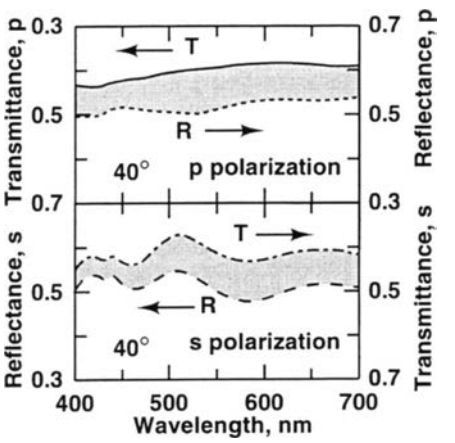


Fig. 6-53 The angle of incidence is 40° in glass of index 1.517.

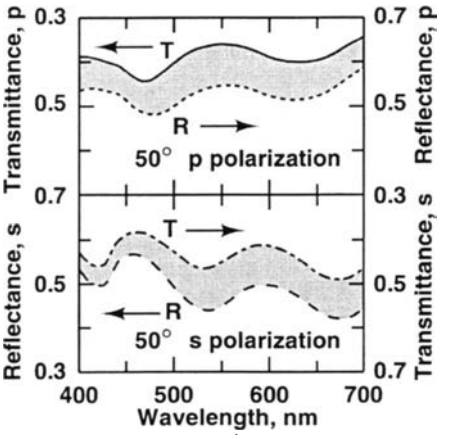


Fig. 6-54 The angle of incidence is 50° in glass of index 1.517.

For Figs. 6-53 and 6-54, $R_s(\lambda)$ (*long dash*), $R_p(\lambda)$ (*short dash*), $T_s(\lambda)$ (*long-short dash*) and $T_p(\lambda)$ (*solid*) of the coating whose design appears in §6.8.9. The transmittance is plotted downwards. The shaded area represents the radiant absorption.

6.6 Miscellaneous topics

This section covers both the practical and theoretical aspects of beamdividers.

6.6.1 Joining prisms

In all of the immersed designs cited previously, the prisms must be joined together with a material (such cement) of sufficiently high refractive index such that total internal reflection does not occur at the prism-cement interface. The cement may, or may not, provide mechanical strength. Methods of joining the prisms are:

(1) A film of viscous liquid is inserted between the prism faces. A mechanical fixture clamps the prisms so that the optical alignment is maintained. Microscope immersion oil may be used as the liquid in the visible part of the spectrum. The slang of the trade is to say that prisms are *oiled* together. This technique is also used in the UV, where glycerin has lower optical absorption than optical cement [Pellicori 64, 66].

(2) A non-oxide chalcogenide glass consisting of a mixture of As, S, Se and Te, as well as a glass former, such as B_2O_3 has been used to bond prisms, such as zinc selenide, in the IR [Cary 64].

(3) Prisms can be optically contacted [Lees 79a, 79b]. Title [88] improved the optical contact by depositing a relatively thick layer on the top of the multilayer stack. That layer was subsequently optically polished to flatness.

(4) Optical cement is widely used to join prisms. The types of cement include: (a) an organic two-component cement, such as an epoxy; (b) an organic material thermal setting cement; (c) Canada balsam, which is a liquid at elevated temperatures and solidifies when it cools; and (d) an organic cement that polymerizes when irradiated with UV. Problems associated with cements are mentioned below.

One problem is that the differential coefficient of thermal expansion of a cement and of a glass, in addition to any mechanical shrinkage of the cement, creates mechanical

stress at the glass-film interface. This, in turn, induces strain birefringence in the glass. Clapham [73] observed that this limited the ratio of the transmittances for the two linear polarizations — the T_p / T_s ratio— to 1000 or less in a MacNeille polarizer (see §6.3.2). This problem is mitigated if Schott SF-57 glass is used, It manifests low strain-induced birefringence.

Another problem may occur when the beamdivider contains silver (see §1.8.2.2.1 or §6.5.5). The silver may be degraded by the oxidants in a two-component cement, such as an epoxy. A UV polymerizing cement is preferable to use in such systems. Still another problem is that the optical cement penetrates into the layers adjacent to it and alters their refractive indices. This may be avoided by depositing a barrier layer, such as a halfwave film of aluminum oxide, to inhibit the diffusion.

6.6.2 Linear retarder with 180° differential phase shift upon reflection

In §6.5.2, there is a reflector that is "nonpolarizing — independent of angle." A serendipitous feature of this design is that it manifests a 180° differential phase shift upon reflection (see §2.3.4). The result is that it is *not only* a nonpolarizing beamdivider, but in addition the multilayer equivalent of a halfwave plate in terms of the reflected light. The "halfwave plate" is more accurately termed a "180° linear retarder" [Shurcliff 62].

The proof that the differential phase shift upon reflection is 180° is seen when U from Eq. (6-7d) is set equal to 1.00 and the amplitude reflection coefficients of the stack for p and s reflectance are written at the "tuned" wavelength λ_0, where the reduced admittance is pure real. Equations. (2-16f) and (2-16g) are rewritten as:

$$\hat{\rho}_s = \frac{1 - W}{1 + W} \tag{6-10a}$$

and for p polarization as

$$\hat{\rho}_p = \frac{1 - W^{-1}}{1 + W^{-1}} \ . \tag{6-10b}$$

The conclusion is that, regardless of whether W exceeds 1.00 or is less than 1.00, the algebraic signs of Eqs. (6-10a) and (6-10b) are opposite.

6.6.3 References to beamdividers and polarizers in other chapters

The edge filter polarizer (see §5.4.3.5.4) is an excellent linear polarizer, albeit over a relatively narrow bandwidth. Beamdividers containing silver layers are the focus of §8.8. The Ward polarizer (see Fig. 1-176) is a bandpass filter and is analyzed in §7.9.7.

6.7 Appendices containing derivations

6.7.1 Condition for the existence of a nonpolarizing angle

Figure 6-30 hints at an approach to find a necessary condition for the existence of an angle θ_n at which a nonabsorbing multilayer is nonpolarizing. The phase thickness of each layer is 90°, regardless of the angle of incidence.

The derivation centers on the slope of the function $W(\theta_0)$ [see Eq. (2-48e)] as the incidence angle θ_0 increases at $\theta_0 = 0$. §6.2.1.1.2 lists the conditions under which $W \rightarrow \infty$ as θ_0 approaches 90°. Hence, if W is less than 1 at angles near zero and eventually becomes infinite, there must be a specific angle θ_n at which W equals 1, which is the nonpolarizing angle. Equation (2-48h) is satisfied at that angle. W is, in fact, less than 1 at angles near zero if its derivative at $\theta_0 = 0$ is negative. Equation (2-48e) is rewritten for the stack with ℓ layers, ℓ being odd

$$W=\frac{(1-a_1L^2)(1-a_3L^2)\ldots(1-a_iL^2)\ldots(1-a_{\ell-2}L^2)(1-a_\ell L^2)}{(1-a_2L^2)(1-a_4L^2)\ldots(1-a_iL^2)\ldots(1-a_{\ell-1}L^2)\sqrt{1-L^2}\sqrt{1-a_sL^2}} \tag{6-11a}$$

where a_i is the reciprocal index squared

$$a_i \equiv n_i^{-2} \tag{6-11b}$$

for the ith layer and a_s is the reciprocal index squared of the substrate. L appears in Eq. (2-4a). The next step is to take the log of Eq. (6-11a) and its derivative D with respect to L^2. It is evaluated at $L = 0$, i.e., normal incidence

$$D \equiv \frac{\partial(\ln W)}{\partial(L^2)}\Big]_{\theta_0=0} = \tfrac{1}{2}n_s^{-2} + \tfrac{1}{2} + \sum_{i_{even}=2}^{\ell-1} n_i^{-2} - \sum_{i_{odd}=1}^{\ell} n_i^{-2} \ . \tag{6-11c}$$

For an odd number of layers, D is proportional to

$$D \propto -a_1 + a_2 - a_3 + a_4 \ldots \pm a_i \ \ldots + a_{\ell-1} - a_\ell + \tfrac{1}{2} + \tfrac{1}{2}a_s \tag{6-11d}$$

and for an even number of layers,

$$D \propto -a_1 + a_2 - a_3 + a_4 \ldots \pm a_i \ \ldots - a_{\ell-1} + a_\ell + \tfrac{1}{2} - \tfrac{1}{2}a_s \ . \tag{6-11e}$$

6.7.2 Nonpolarizing, nonimmersed system — solution for the indices

The goal is to solve Eqs. (6-7a) and (6-7b) for n_L and n_H, where θ_0, n_0, n_s, θ_s, L and n_M are given. The aforementioned equations are rewritten in terms of reciprocal index squared

$$a_H \equiv n_H^{-2}, \qquad a_L \equiv n_L^{-2} \tag{6-12a}$$

which results in

$$1 = \frac{(\,1-L^2\,n_M^{-2}\,)^{2q+1}}{(1-L^2\,a_L)^q\,(1-L^2\,a_H)^q\cos\theta_0\cos\theta_s} \tag{6-12b}$$

and

$$U = n_M^{4q+2}\,a_H^q\,a_L^q\,(n_0\ n_s)^{-1} \ . \tag{6-12c}$$

Taking the qth root of each of the foregoing equations and rearranging produces

$$(1-L^2\,a_L)\ (1-L^2\,a_H) = (\sec\theta_s\ \sec\theta_0)^{1/q}\ (1-L^2\,n_M^{-2})^{(2q+1)/q} \equiv c \tag{6-12d}$$

and

$$a_L\ a_H = n_M^{-(4q+2)/q}\,(U\ n_0\ n_s)^{1/q} \equiv f \ . \tag{6-12e}$$

Further manipulation of Eqs. (6-12d) and (6-12e) results in a quadratic equation for a_L. The solution is completed in §6.7.4.

6.7.3 Three-component immersed system — solution for the indices

This system differs from that of §6.7.2 because it contains an even number of layers and is immersed. The goal is to solve Eqs. (6-8a) and (6-8b) for n_L and n_H, where θ_0, n_0, n_s, θ_s, L and n_M are given. The aforementioned equations are rewritten in terms of reciprocal index squared, from Eq. (6-12a)

$$(1-L^2\,a_L)\ (1-L^2\,a_H) = (\sec\theta_s\ \cos\theta_0)^{1/q}\ (1-L^2\,n_M^{-2})^2 \equiv c \tag{6-13a}$$

and

$$a_L\ a_H = n_M^{-4}\,[\,n_s\,/\,(U\,n_0)]^{1/q} \equiv f \ . \tag{6-13b}$$

The solution is completed below.

6.7.4 Solution for the reciprocal index squared

Given Eqs. (6-12d) and (6-12e), or Eqs. (6-13a) and (6-13b), the solution for the quadratic is

$$a_L = \frac{-c+1+L^4 f+\sqrt{c^2-2c-2cL^4 f+1-2L^4 f+L^8 f^2}}{2L^2} . \tag{6-13c}$$

Equation (6-12a) is inverted to find n_L

$$n_L = (a_L)^{-1/2} \tag{6-13d}$$

$$n_H = \sqrt{a_L / f} . \tag{6-13e}$$

6.7.5 Two-component immersed system — solution for the indices

Equations (6-6b) and (6-6c) are solved for the high and low indices. They are linear in reciprocal index squared — see Eq. (6-12a). κ is defined as

$$\kappa \equiv (n_0 \; n_s^{-1} \; U)^{1/q} = a_H / a_L \tag{6-14a}$$

and z as

$$z \equiv (\cos\theta_0 \sec\theta_s)^{1/q} = (1 - L^2 a_L) \; (1 - L^2 a_H)^{-1} . \tag{6-14b}$$

Then

$$n_L = L \sqrt{\frac{1 - z\,\kappa}{1 - z}} . \tag{6-14c}$$

a_H is then found from

$$n_H = n_L \sqrt{1/\kappa} . \tag{6-14d}$$

6.8 Appendices containing multilayer designs

6.8.1 Design: polarizer with 52 layers

The design related to Figs. 6-28 and 6-29 is:

glass 249A 125B 132A 114B 128A 119B 131A 121B 134A 123B 137A 123B 136A 121B 136A 122B 135A 122B 137A 126B 143A 128B 141A 125B 142A 131B 152A 138B 121A 97.6C 169A 75.4C 185A 63.3C 199A 84.8C 198A 87.5C 204A 65.2C 203A 55.2C 194A 92.4C 153A 47.4B 88.1D 129B 105D 180B 117D 34.9B *glass* .

The refractive indices of *glass*, A, B, C and D are 1.517, 1.60, 1.74, 2.10 and 2.34, respectively. Layer thicknesses are metric in nm.

6.8.2 Design: polarizer with 67 layers

The design related to Fig. 6-30 is:

high-index 0.101D 0.1047C 0.0432B 0.1058D 0.1084B 0.1293D 0.0811B 0.1148D 0.0293C 0.0725B 0.1159D 0.0269C 0.1073A 0.1015D 0.0174C 0.1748A 0.0153D 0.1774A 0.0391C 0.0799D 0.1395A 0.1320D 0.1253A 0.1157C 0.173A 0.0282C 0.2003A 0.0979C 0.1678A 0.0709C 0.0719D 0.1613B 0.1882A 0.0695C 0.1511A 0.0917B 0.1960A 0.0784C 0.1475A 0.0956B 0.1890A 0.0211D 0.2270A 0.069C 0.2014A 0.019C 0.1565A 0.0351D 0.0354C 0.2001A 0.3297B 0.13742A 0.0752C 0.2274A 0.0323B 0.1372A 0.0622C 0.2791A 0.3386B 0.0899A 0.0648D 0.1042A 0.0434D 0.1122A 0.0761D 0.1473C 0.5868A *low-index*

where *high-index* and *low-index* represent the incident medium of refractive index 1.73 and emergent medium of index 1.517, respectively. Layer thicknesses are metric in micrometer. The refractive indices of A, B, C and D are 1.60, 1.74, 2.10 and 2.34, respectively.

6.8.3 Design: linear polarizer

The design related to Fig. 6-31 is:

glass 55.25H 133.07L 73.39H 156.97L 75.62H 147.94L 57.57H 50.11M 160.50L 74.99H 233.40L 54.67M 25.29H 156.82L 93.02M 167.91H 87.41L 33.16H *glass* .

The refractive indices of *glass*, L, M and H are 1.517, 1.45, 1.90 and 2.35, respectively. Layer thicknesses are metric in nm.

6.8.4 Design: polarizer with 98 layers

The design related to Fig. 6-34 is

glass 0.1335L 0.0724M 0.0391H 0.1157L 0.0450H 0.0819L 0.0691H 0.0921L 0.0462H 0.0904L 0.0698H 0.0904L 0.0414H 0.0817L 0.0831H 0.0888M 0.0432H 0.0709M 0.1142H 0.0729L 0.0405H 0.1109L 0.0113M 0.0601H 0.1108M 0.0602H 0.0630L 0.1006H 0.0928M 0.0652H 0.0893M 0.0948H 0.0549L 0.1007H 0.0967M 0.0805H 0.0415M 0.0594L 0.0774H 0.1118L 0.0837H 0.1069M 0.0809H 0.1062M 0.0841H 0.1169L 0.0830H 0.1177M 0.0991H 0.1235M 0.0790H 0.1202L 0.0983H 0.1325M 0.0586H 0.1701M 0.0878H 0.1524L 0.0926H 0.1665M 0.0351H 0.1580M 0.1376H 0.0690L 0.1248H 0.1349M 0.1010H 0.1461M 0.1326H 0.0567L 0.1401H 0.1482M 0.0954H 0.0416L 0.1170M 0.1169H 0.1733L 0.1205H 0.1495M 0.1080H 0.1577M 0.1173H 0.0639L 0.0373M 0.0746L 0.1217H 0.1564M 0.1171H 0.1635M 0.1210H 0.2270L 0.1005H 0.0238L 0.0390H 0.1617L 0.1111M 0.1246L *glass* .

Layer thicknesses are metric in micrometer. The refractive indices of *glass*, L, M and H are 1.517, 1.67, 1.74 and 2.08, respectively.

6.8.5 Design: polarizer with 64 layers

The design related to Fig. 6-35 is based upon the design of Li [96] and is

glass 0.0815A 0.0439B 0.1637A 0.0782B 0.0771A 0.0629C 0.0625D 0.0346A 0.0428D 0.1227A 0.0222C 0.0269D 0.1156A 0.0497D 0.1288A 0.0464D 0.1051A 0.0373D 0.0441B 0.1006A 0.0454C 0.0925D 0.1571A 0.0377D 0.0855A 0.0384D 0.0986C 0.0451A 0.0614B 0.0162C 0.0631D 0.0239C 0.0224B 0.1256A 0.0210B 0.0234C 0.0747D 0.0217C 0.0244B 0.1255A 0.0159B 0.0204C 0.0856D 0.0414B 0.1277A 0.0322B 0.0921D 0.0821B 0.0560A 0.0539B 0.0940D 0.0566C 0.0447A 0.0826B 0.1000D 0.0438C 0.0588A 0.0723B 0.1028D 0.0330C 0.0649A 0.0720B 0.1192D 0.0197A *glass*

where the refractive indices of *glass*, A, B, C and D are 1.52, 1.45, 1.65, 1.84 and 2.35, respectively. Layer thicknesses are metric in μm.

6.8.6 Design: 50%-50% beamdivider

The design related to Fig. 6-43 is

air 0.2396L 0.0478H 0.2762L 0.0964H 0.3639L 0.0749H 0.1214L 0.0670H 0.1043L 0.0547H 0.0862L 0.0492H 0.0871L 0.0548H 0.1018L 0.0628H 0.2630L *glass*

where the refractive indices of *glass*, L and H are 1.517, 1.442 and 2.26, respectively. Layer thicknesses are metric in μm.

6.8.7 Design: 50%-50% beamdivider

The design related to Fig. 6-44 is

air 0.0146H 0.1484L 0.1695H 0.1416L 0.0662H 0.1017L 0.0561H 0.0933L 0.0543H 0.0931L 0.0553H 0.0975L 0.0602H 0.1095L 0.0670H 0.1160L 0.0667H 0.0554L *glass*

where the refractive indices of *glass*, L and H are 1.517, 1.442 and 2.26, respectively. Layer thicknesses are metric in μm.

6.8.8 Design: all-dielectric beamdivider with minimal polarization splitting

The coating related to Fig. 6-52 contains 105 layers and has the design:

glass 0.0163**H** 0.0405**F** 0.0316**H** 0.0981**M** 0.0224**H** 0.0474**F** 0.0226**H** 0.1003**M** 0.0262**H**
0.0517**F** 0.0197**H** 0.1004**M** 0.0633**H** 0.0961**M** 0.0208**H** 0.0561**F** 0.0197**H** 0.0906**M** 0.0155**H**
0.0535**F** 0.0227**H** 0.0919**M** 0.0597**H** 0.0951**M** 0.0299**H** 0.0562**F** 0.0147**H** 0.0953**M** 0.0345**F**
0.0532**H** 0.1051**M** 0.0100**H** 0.1554**F** 0.1329**M** 0.0460**H** 0.0658**F** 0.1098**M** 0.0347**F** 0.0545**H**
0.1050**M** 0.0320**H** 0.0526**F** 0.0215**H** 0.0987**M** 0.0195**F** 0.0605**H** 0.1096**M** 0.1739**F** 0.1578**M**
0.0227**H** 0.0744**F** 0.0199**H** 0.0950**M** 0.0536**H** 0.0206**F** 0.0886**M** 0.0194**H** 0.0603**F** 0.0223**H**
0.0643**M** 0.0619**F** 0.0222**H** 0.0786**M** 0.0578**H** 0.0931**M** 0.0151**H** 0.0446**F** 0.0268**H** 0.0775**M**
0.0388**H** 0.0873**M** 0.0333**H** 0.0394**F** 0.0139**H** 0.0954**M** 0.0247**H** 0.0352**F** 0.0278**H** 0.1876**M**
0.0207**H** 0.0278**F** 0.0365**H** 0.1039**M** 0.0209**H** 0.0385**F** 0.0408**H** 0.1193**M** 0.2092**F** 0.1090**M**
0.0630**H** 0.0514**F** 0.0218**H** 0.1126**M** 0.0562**H** 0.0255**F** 0.1918**M** 0.0657**H** 0.0981**M** 0.0205**H**
0.0659**F** 0.0391**H** 0.1732**M** 0.0790**F** 0.0166**H** 0.0348**F** *glass*

where the refractive indices of *glass*, **F**, **M** and **H** are 1.517, 1.38, 1.68 and 2.29, respectively. Layer thicknesses are metric in μm.

6.8.9 Design: beamdivider with a single silver layer

The design related to Figs. 6-53 and 6-54 contains 41 layers and is:

glass 93.29**M** 12.18**H** 42.80**M** 24.17**H** 27.06**M** 26.52**H** 27.35**M** 26.25**H** 30.04**M** 22.00**H** 28.92**M**
17.00**H** 24.32**M** 17.00**H** 34.62**M** 13.08**H** 39.56**M** 14.48**H** 28.80**M** 61.51**H** 25.00**S**
61.51**H** 28.80**M** 14.48**H** 39.56**M** 13.08**H** 34.62**M** 17.00**H** 24.32**M** 17.00**H** 28.92**M**
22.00**H** 30.04**M** 26.25**H** 27.35**M** 26.52**H** 27.06**M** 24.17**H** 42.80**M** 12.18**H** 93.29**M** *glass*

where the refractive indices of *glass*, **M** and **H** are 1.517, 1.62 and 2.25, respectively. The optical constants of the silver layer **S** appear in Table 8-6. Layer thicknesses are metric in nm.

Chapter 7
All-dielectric bandpass filters

7.1 Introduction

7.1.1 Usage of all-dielectric filters

As mentioned in §1.6, a bandpass filter is highly transmissive in the spectral region of its passband. It rejects both the shorter and longer wavelengths contiguous to its passband, thus allowing it to isolate a limited region of the spectrum. The spectral width of its passband may be as small as a fraction of a nm for a filter used in fiber optical communications. A filter that isolates the FLIR band (7.5 μm to 13 μm) in the IR has a passband that is as wide as 4 μm.

Table 7-1
Bandpasses: All-dielectric and those containing metal layers

	Type of filter	
Attribute	All-dielectric	Contains metal layers
Fractional bandwidth	Can be as small as a few tenths of a percent.	Relatively broad, usually a few percent or greater.
Maximum transmittance in passband	Can be quite large, as, for example 99%.	Limited, due to the absorption of the metal layers.
Absorbance in offband spectral region	4 to 5, over a limited spectral bandwidth .	Can be as large as 6 — over a quite broad spectral bandwidth.
Spectral width of offband rejection region	Small. For example, ±12% of λ_0 .	Can be quite broad — many hundreds of nm.

Attributes of bandpass filters themselves — exclusive of any additional blocking filters.

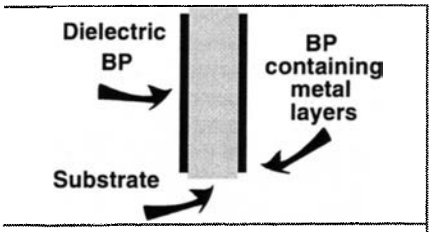


Fig. 7-1 A filter consists of a substrate coated on one side with an *all-dielectric bandpass* and on the opposite side with a *blocking filter* containing metal layers.

A bandpass may be constructed two ways: (1) with only nonabsorbing layers — termed an *all-dielectric* coating; or (2) with combinations of nonabsorbing and absorbing metal layers. Bandpasses containing absorbing materials (i.e., metal films) are reviewed in Chap. 8. Chapter 7 focuses on all-dielectric designs. Under what circumstances should each type of filter be used? Although there are inherent dangers in making sweeping generalizations, Table 7-1 lists the attributes of each filter. The advantages of the all-dielectric bandpass are its relatively high peak transmittance and, if required, a narrow spectral bandwidth. A desirable property of the bandpass containing metal layers is that it offers an immense amount of rejection — termed *blocking* in the trade — over a relatively wide spectral region. Blocking is illustrated in §1.6.5. An example is that the silver layers in the coating whose transmittance is shown in Fig. 8-29 attenuate quite effectively in the near IR.

Figure 7-1 shows a substrate coated on one side with an all-dielectric bandpass and on the opposite side with a filter containing metal layers. With the exception that it is lacking a relatively large passband transmittance, it combines the best of both worlds. It manifests both a narrow spectral bandwidth (provided by the all-dielectric) and substantial offband rejection over a broad range of wavelengths (supplied by the metal-dielectric coating).

7.1.2 Filter classification

Figure 7-2 portrays the design of the three classes of all-dielectric bandpasses:

(a) a mutual-passband type;
(b) a single-cavity (i.e., the Fabry-Perot) filter; or
(c) a multiple-cavity system.

The optical performance of each is quantified by parameters such as *resolution, shape* and *bandwidth,* as discussed in §1.6.4.

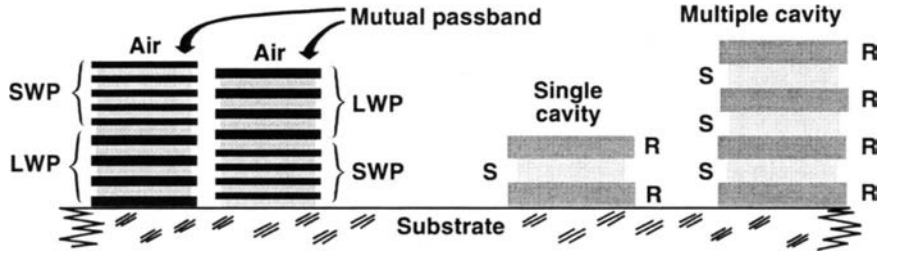


Fig. 7-2 Construction of all-dielectric bandpasses. (*Left*) A shortwave pass and longwave pass are cascaded to form a mutual-passband filter. (*Center*) Single-cavity bandpass consists of a spacer *S* sandwiched between multilayer reflectors *R*. (*Right*) Multiple-cavity bandpass contains halfwave thickness spacers *S* interdispersed between multilayer reflectors *R*.

When should each type of filter be used? The mutual-passband is preferred when the spectral bandwidth is rather broad—as, for example, 10% or greater. The Fabry-Perot is rarely used because of its poor offband rejection. The multiple-cavity filter offers the designer the most flexibility because of the many ways that its bandwidth and offband rejection may be altered. The focus of this chapter is on the multiple-cavity bandpass.

7.1.3 Design methods

7.1.3.1 Introduction

Of the filters mentioned in §7.1.2, the design of a mutual-passband filter is discussed in §7.1.4. The Fabry-Perot is used in §7.2 to illustrate methods of varying the spectral bandwidth of a bandpass.

7.1.3.2 Multiple-cavity design methods

There are at least three approaches to the design of multiple-cavity filters: (1) the method of *effective interfaces* [Smith 58]; (2) the *periodic-structure* technique; and (3) the procedure based on *microwave filter* design. The *effective interfaces* method is based upon partitioning (see §2.6.4) and is useful in the design of a two-cavity bandpass. The focus in this book is on the last two methods.

A reader who desires to study bandpass design in depth will undoubtedly concoct bandpasses using several of the above methods and will conclude which method is the most useful. My own experience has taught me: (1) the *periodic structure* method is conceptually simpler and easier to learn; (2) there is considerably more art—if not guesswork—in using the *periodic structure* to produce a filter with a given spectral bandwidth; (3) the complexity of the *microwave* design procedure makes it more difficult to master; and (4) the *microwave* is more logical and rationally guides the user to the next step.

An example illustrates the previous paragraph. Suppose a bandpass is designed in which the semiwidth is adequate but the spectral bandwidth at the 1% transmittance level is too large. How is this rectified? I have observed two optical coating engineers [Robbins 85, Swaby 85] whose use of the *periodic-structure* design method for several decades has endowed them with a superb intuitive and analytical grasp of this method. They know instinctively the procedure used to modify any given bandpass. That next step in the *periodic-structure* design method may not be so obvious to the novice, however. An increase in the number of cavities (see §7.6.3) is the solution provided by the *microwave* design method.

7.1.4 Mutual-passband filter

This bandpass is fabricated by depositing a highpass and a lowpass edge filter on the same substrate. Each edge filter has a stop band, which is the spectral region of low transmittance. The gap between those stop bands—the overlapping passbands—constitutes the *mutual-passband.* The advantage of this method is that the width of the passband can be adjusted quite arbitrarily by moving one or both of the edge filters to either shorter or longer wavelengths.

When should this type of filter be used? Many designers prefer this filter when the spectral bandwidth exceeds 10%, since it is difficult to design multiple-cavity filters with such a large spectral bandwidth. However, when the spectral bandwidth drops below 7%, it becomes increasingly difficult to control the spectral bandwidth of the mutual-passband filter. Then it is preferable to use the multiple-cavity type of filter.

Epstein [52] designed an archetypal mutual-passband filter that used equivalent layers to increase the transmittance in the passband region. This same approach is used to design the bandpass whose spectral transmittance is shown in Fig. 7-3 [Ufford 72]. Its spectral bandwidth of 18% is quite wide. The edge filter (see §5.4) adjacent to air reflects in the blue part of the spectrum. The *core stack* (see §5.4.2.1) consists of ten quarterwave (near 450 nm) layers. The $T(\lambda)$ of this long wavelength pass filter was optimized—*air* to *air*, as shown in the caption to Fig. 7-4.

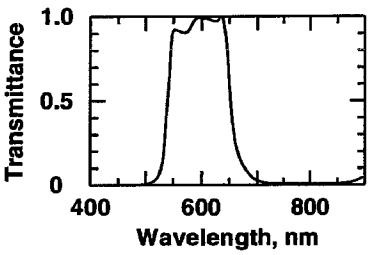


Fig. 7-3 $T(\lambda)$ of the mutual-bandpass filter
air 0.234H 0.455L 0.461H 0.447L
0.416H (0.45L 0.45H)^5 0.447L
0.461H 0.455L 0.234H 0.991H
0.682L 0.841H (0.762L 0.762H)^3
0.762L 0.819H
0.768L 0.922H 0.943L *glass* .
The symbols are captioned in Fig. 7-4.

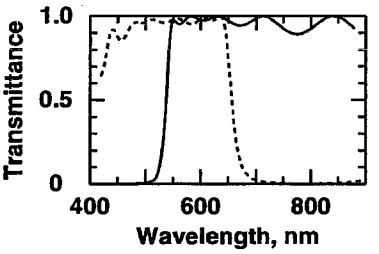


Fig. 7-4 $T(\lambda)$ of a shortwave pass filter (*dashed*)
air 0.991H 0.682L 0.841H (0.762L
0.762H)^3 0.762L 0.819H 0.768L
0.922H 0.943L *glass*
and a longwave pass filter (*solid*)
air 0.234H 0.455L 0.461H 0.447L
0.416H (0.45L 0.45H)^5 0.447L
0.461H 0.455L 0.234H *air*
where *glass*, L and H represent materials of refractive indices 1.52, 1.38 and 2.25, respectively. The optical thickness of L and H is $\lambda_0/4$ at λ_0 of 1.0 μm.

The high-frequency-pass multilayer stack reflects the red; its *core stack* is centered at 762 nm. The mutual passband extends from approximately 530 nm to 640 nm. Those annoying dips in the transmittance on the short-wavelength side of the passband, as well as the offband rejection, could be improved by optimizing the entire filter.

There is little point in devoting more space to the design of this filter, since it is relatively straightforward. The production of this filter is not without its problems, however. The optical monitoring (see §9.6.2) of the thicknesses of the layers as the filter is deposited can be a problem because the coating contains layers with quarterwave optical thicknesses that are widely separated—from 450 nm to 762 nm. The dispersion of the high-index materials is noticeable at short wavelengths and must be taken into account when the optical monitoring scheme is set up. Another problem is that even a small amount of absorption in the dielectric layers reduces considerably the transmittance at the edges of the passband.

7.1.5 Periodic structures without halfwave layers

There is a unique type of bandpass that contains no *cavities*—layers of halfwave thickness. It has a rather wide passband, the width of which is difficult to alter. This bandpass is a periodic structure in which the passband is the gap between adjacent stop bands. A *stop band* (see §5.2.2) may be viewed as a spectral region where the transmittance decreases as more basic periods are added to the stack. An example is the coating

air (0.667L 0.667H 0.667L)^q *substrate* ,

where q is the number of times the basic period is repeated. The optical thicknesses of the layers are a fraction of $\lambda_0/4$, where λ_0 is the center wavelength of the first-order stop band. The coating is similar to the one depicted in the middle panel in Fig. 5-24.

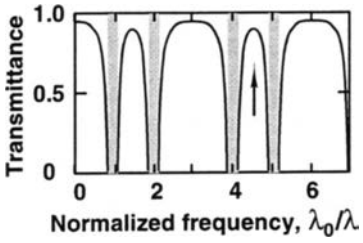


Fig. 7-5 Lower envelope of the transmittance (*solid*) and stop bands (*shaded*) of the basic period
air **L' H' L'** *glass* .
The refractive indices of **L'**, **H'** and *glass* are 1.45, 2.30 and 1.70, respectively. The optical thickness of each layer is $\lambda_0/6$. The arrow designates a passband.

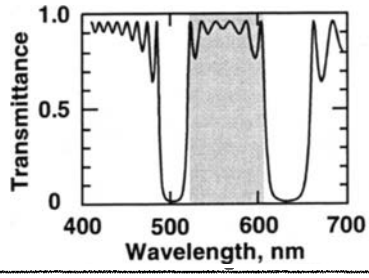


Fig. 7-6 $T(\lambda)$ of the bandpass
air $(\mathbf{L'\,H'\,L'})^7$ *glass*
where the refractive indices of *air*, **L'**, **H'** and *glass* are 1.00, 1.45, 2.30 and 1.52, respectively. The optical thickness of each layer is 420 nm. The shaded region designates its passband.

Figure 7-5 plots the spectral locations of the stop bands vs normalized frequency. The lowest order stop band occurs at a frequency of 1.00. One possibility is to produce a bandpass by locating the passband center at a frequency of 4.5. The stop bands (that provide the blocking) are contiguous to that center — at frequencies of 4.00 and 5.00.

The coating includes 15 alternate layers of SiO_2 (*index* 1.45) and TiO_2 (*index* 2.30). Suppose that passband center (center on a *frequency* scale) is to be located at 560 nm. Then the fourth-order stop band should be centered at a wavelength of $\lambda_4 = 560 \times 4.5/4 = 630$ nm. In other words, the optical thickness of each layer is a fullwave at 630 nm. Figure 7-6 shows the transmittance — the passband is shaded. The fifth-order stop band is at $\lambda_5 = 500$ nm, which is approximately four-fifths of 630 nm.

This example illustrates that the spectral bandwidth of this type of coating is rather inflexible, since the ratio of λ_4 to λ_5 is set at 5:4. A slight narrowing of the spectral bandwidth would be provided, of course, if the low-index layer **L** has an index of 1.38, rather than 1.45, since this would increase the stop band width (see §5.2.4).

The layer thicknesses are somewhat greater than those of a mutual-passband filter (see §7.2), which is an alternative design approach. Notwithstanding these limitations, Rock [Kelly 56] used this design quite successfully as an immersed beamdivider that transmitted the green and reflected the blue and red. Its transmittance appears in Fig. 1-181.

Although the foregoing example uses the gap between the fourth-order and fifth-order stop bands, it is clear that there are other options. A quite broad passband would be obtained, for example, if the gap between the first-order and second-order stop bands were used.

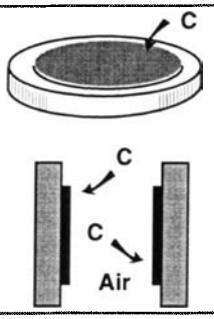


Fig. 7-7 (*Upper panel*) Showing a plate of a Fabry-Perot with the reflector (an optical coating C). (*Lower panel*) In cross section, assembly of the Fabry-Perot with two plates facing each other and with air as the spacer. The thicknesses of the air and the optical coating are exaggerated for purposes of illustration.

7.2 Control of the spectral bandwidth

The objective is to produce a multiple-cavity filter with a *prescribed* bandwidth. The last paragraph of §2.11.4.2 lists three methods of altering the spectral bandwidth:

(1) changing the thickness of a spacer;
(2) altering the reflectance of the reflectors that surround a spacer; and
(3) adjusting the *phase dispersion* of the reflectors contiguous to a spacer.

The single-cavity filter—the Fabry-Perot—serves as an example. Figures 2-96 and 2-97 depict the Fabry-Perot as a cavity surrounded by reflectors, which could be termed the "walls of the cavity." An air-spaced Fabry-Perot is used as an example in §7.21, §7.22 and §7.23 to illustrate control of the spectral bandwidth. The center wavelength is at the YAG laser line, $\lambda_0 = 1060$ nm. The upper panel of Fig. 7-7 shows one of the glass plates of the Fabry-Perot, which is coated with a multilayer reflector.

The lower panel depicts (in cross section) the final assembly of the Fabry-Perot bandpass, in which two coated plates are separated by a thickness of air. The design of the reflector is represented by the three layers of alternating high and low index, **HLH**. **H** and **L** represent layers of optical thickness $\lambda_0/4$.

Resonance (and the accompanying maximum transmittance) occurs when the optical thickness of the air is a halfwave, or integer multiple thereof. The design of the final assembly is

glass **H L H** (spacer layer of air) **H L H** *glass*

or

glass **H L H 2A H L H** *glass*

where **A** represents a layer of quarterwave optical thickness of air (refractive index 1.0003). The symbol **2A** represents two contiguous layers of quarterwave optical thickness, which coalesce into a halfwave optical thickness layer. The terms *spacer* or *cavity* are used interchangeably to describe it.

A final comment is that although the Fabry-Perot—the single-cavity bandpass filter—is not widely used because of the wide spread (spectrally) of its passband, it is still useful as an illustration.

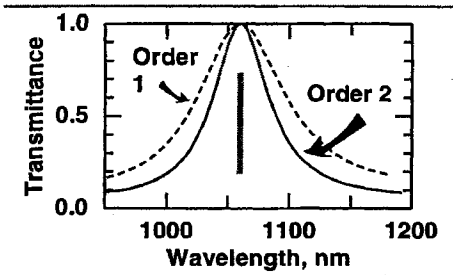


Fig. 7-8 $T(\lambda)$ of a Fabry-Perot with a first-order spacer (*dashed*)
glass **H L H** 2**A H L H** *glass*
and with a second-order spacer (*solid*)
glass **H L H** 4**A H L H** *glass*
where *glass,* **L**, **H** and **A** represent materials of refractive indices 1.51, 1.45, 2.30 and 1.0003, respectively. **L**, **H** and **A** are layers of optical thickness λ_0 /4 at λ_0 of 1060 nm.

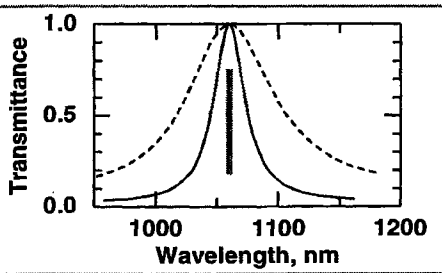


Fig. 7-9 $T(\lambda)$ of a Fabry-Perot with a five layer reflector (*solid*)
glass **H L H L H** 2**A H L H L H** *glass*
and with a three-layer reflector(*dashed*)
glass **H L H** 2**A H L H** *glass*
where *glass,* **L**, **H** and **A** represent materials of refractive indices 1.51, 1.45, 2.30 and 1.0003, respectively. **L**, **H** and **A** are layers of optical thickness λ_0 /4 at λ_0 of 1060 nm.

7.2.1 Spectral bandwidth control — alteration of the thickness of the spacer

Figure 2-96 illustrates that the order number may be visualized as the number of halfwaves of λ_0 that "fit" into the cavity. A thicker cavity produces a higher order number. Inspection of Eq. (2-83h) indicates that the Q (i.e., reciprocal bandwidth) is linearly dependent upon the spacer cavity thickness. The dashed curve in Fig. 7-8 represents the transmittances of single-cavity bandpasses with first-order spacers.

The meaning of the symbols **H**, **A** and **L** appears in §7.2. If the separation between the plates is doubled, the design with the second-order spacer is

glass **H L H** 4**A H L H** *glass* .

Its transmittance appears as the solid curve in Fig. 7-8. Table 7-2 shows that increasing the spacer thickness from first order to second order decreases the spectral bandwidth from 91 nm to 57 nm.

7.2.2 Reflectance of the reflectors

An immediate question is, how is the reflectance of a reflector measured? The answer is that the reflectance is determined from inside of the spacer, which is air. This "naked" reflector is depicted in the upper panel of Fig. 7-7. Consider the reflector used in §7.2.1. Figure 5-1 illustrates that at λ_0 the reflectance of the reflector containing three quarterwave layers **HLH** is 63%. The addition of another **HL** pair to form **HLHLH** boosts the reflectance to 84%. Using the resulting bandpass has the design

glass **H L H L H** 2**A H L H L H** *glass*

and its transmittance is graphed in Fig. 7-9. Table 7-2 shows that the higher reflectance reduces the spectral bandwidth from 91 nm to 31 nm. There is little point in using Eq. (2-83h) to find the fractional bandwidth because single-cavity filters are rarely used.

Table 7-2 Single-cavity bandpass—altering its spectral bandwidth

Modification	Design	*R*	BW	Ratio	Ref.
None — *original design*	**H L H 2A H L H**	63%	91 nm	1.00	§7.2.1
Double spacer thickness	**H L H 4A H L H**	63%	57 nm	0.63	§7.2.1
Reflector: Increase reflectance	**(H L)^2 H 2A (H L)^2 H**	83%	31 nm	0.34	§7.2.2
Reflector: Increase faz dsprs'n	**(M L)^3 M 2A (M L)^3 M**	63%	57 nm	0.63	§7.2.3.1
Reflector: Increase faz dsprs'n	**H L 3H 2A 3H L H**	63%	62 nm	0.68	§7.2.3.4.1
Reflector: Increase faz dsprs'n	**H 3L H 2A H 3L H**	63%	73 nm	0.80	§7.2.3.4.2

Modifications of a single-cavity bandpass (a Fabry-Perot) that reduce its bandwidth BW. "faz dsprs'n" is an abbreviation of "phase dispersion." **L**, **H**, **M** and **A** are layers of optical thickness λ_0 /4 at λ_0 of 1060 nm and refractive indices of 1.45, 2.30, 1.826 and 1.0003, respectively. The refractive index of the incident medium and emergent medium is 1.51. *R* is the reflectance of the reflector, as viewed from the spacer. *Ratio* is the BW divided by the spectral bandwidth of the original design.

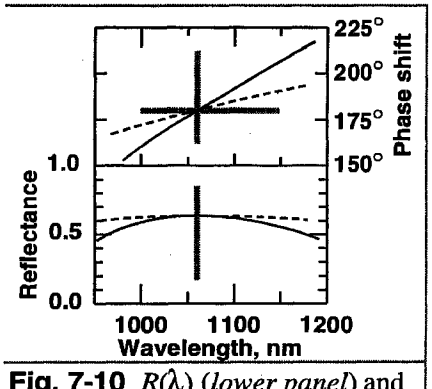


Fig. 7-10 *R*(λ) (*lower panel*) and phase shift upon reflection (*upper panel*) of a three-layer reflector (*solid*)
air **H L H** *glass*
and a seven-layer reflector(*dashed*)
air **M L M L M L M** *glass*
where *glass*, **L**, **H** and **M** represent materials of refractive indices 1.51, 1.45, 2.30 and 1.826, respectively. **L**, **H** and **M** are layers of optical thickness λ_0 /4 at λ_0 of 1060 nm.

7.2.3 Phase dispersion

What is phase dispersion? It is the change (with wavelength) of the phase shift upon reflection of the reflectors that bound a spacer. As discussed in §7.6.2.5, the phase dispersion has a profound effect upon the spectral bandwidth. This section illustrates that control of the phase dispersion alters the spectral bandwidth.

Two examples are included in §7.2.3. In the first, the reflectors of a single-cavity bandpass have identical radiant reflectance. However, due to the fact that the phase dispersion of the reflectors is greater, the spectral bandwidth of the bandpass is reduced.

In the second example, the phase dispersion is altered by introducing layers of three quarterwave optical thickness in the vicinity of the spacer layers.

7.2.3.1 Phase-dispersion narrowing of the passband — first example

Figure 7-10 depicts the reflectance of a three-layer quarterwave stack reflector **HLH** in which there is a substantial difference in the refractive indices of **L** and **H**—indices of 1.45 and 2.30 are used. The ratio of the refractive indices of the **M** and **L** layers in a seven-layer reflector **MLMLMLM** is smaller. The refractive indices are 1.826 and 1.45. The graph of the reflectance in the lower panel of Fig. 7-10 shows that the reflectance at the "tuned wavelength" of 1060 nm is essentially the same.

The upper panel of Fig. 7-10 graphs the phase shift upon reflection. Although the phase shifts are identical (i.e., 180°) at λ_0, the slope of the phase shift curve is greater for the seven-layer reflector. Ergo, the dispersion of the phase shift upon reflection is greater, which, in turn, narrows the spectral bandwidth of the Fabry-Perot, as depicted by the transmission curve in Fig. 7-11.

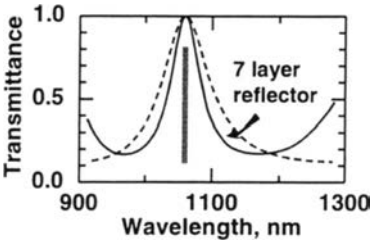


Fig. 7-11 $T(\lambda)$ of a Fabry-Perot with a three-layer reflector (*dashed*)
glass **H L H 2A H L H** *glass*
and with a seven-layer reflector(*solid*)
glass **(M L)^3 M 2A (M L)^3 M** *glass*
where *glass,* **L**, **H**, **M** and **A** represent materials of refractive index 1.51, 1.45, 2.30, 1.826 and 1.0003, respectively. **L**, **H**, **M** and **A** are layers of optical thickness $\lambda_0/4$ at λ_0 of 1060 nm.



Fig. 7-12 Standing wave electric field in a resonant cavity: (*A*) At resonance at a wavelength λ_0 in a standard-length cavity; (*B*) off resonance at a wavelength of 0.75 λ_0 in a standard-length cavity; (*C*) off resonance at a wavelength λ_0 in a cavity whose length is 1.3 times the standard-length cavity; and (*D*) off resonance at a wavelength of 0.75 λ_0 in a cavity whose length is 1.3 times the standard-length cavity.

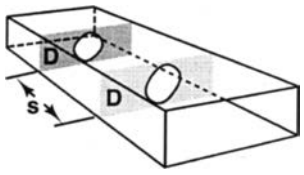


Fig. 7-13 Microwave analog of a single-cavity optical filter in which two iris diaphragms *D* in a waveguide function as partial reflectors. The distance *s* is approximately one-half of the guide wavelength.

7.2.3.2 Phase dispersion — mechanical analogy

As mentioned above, the transmittance of a single-cavity filter peaks when resonance occurs within the cavity (see Fig. 2-96). Figures 2-97 and 7-12 show the standing wave electric fields both at resonance and off resonance. At the tuned wavelength, a halfwave (or multiple thereof) "fits" into the cavity. At other wavelengths (off resonance), the foregoing condition does not hold and, as a consequence, the transmittance drops off.

There are two ways of achieving the off-resonance condition: (1) the wavelength (i.e., frequency) is changed; or (2) the wavelength remains constant and the cavity length is altered by physically moving one of the walls. Both of these scanning methods are used in the microwave and optical spectral regions. As an illustration, Fig. 7-13 shows the microwave analog of the single-cavity filter, consisting of two reflectors *D* (iris diaphragms) separated by a segment of waveguide. At resonance, the separation is approximately $\lambda_0/2$, where λ_0 is the guide wavelength.

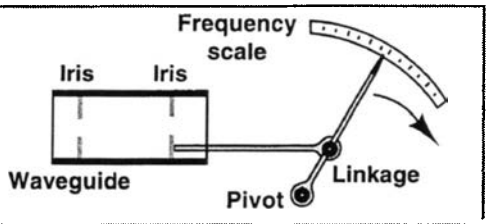


Fig. 7-14 Illustration of phase-dispersion narrowing of a resonance system. A microwave cavity is excited by the power from an analog sweep frequency generator. A mechanical linkage alters the physical cavity length simultaneously as the frequency is changed.

Suppose that one of the irises is linked mechanically to the dial that alters the frequency of the microwave oscillator which, in turn, feeds microwave energy into the cavity, as depicted in Fig. 7-14. The result is that a rotation of the dial causes the physical location of the iris to move, thus causing *both* the cavity length and wavelength to change simultaneously. The result is that as the oscillator is tuned to higher frequencies (i.e., shorter wavelengths), the cavity length simultaneously gets *longer* and, hence, the transmittance of the system drops even faster than it would if the linkage were absent (see panel *C* of Fig. 7-12). The result is that the transmittance vs frequency curve is narrowed.

How is this related to *phase-dispersion narrowing* in multilayers at *optical* frequencies? The analogy is that the phase shift upon reflection (see §2.3.2) of the reflectors on either end of the cavity changes with wavelength. Because the resonance depends upon the phase shift of the reflector (see §5.6.4), a change in phase shift with wavelength causes the effective cavity length to change as well. Although it is possible to concoct multilayers that have a phase dispersion that *broadens* the resonance line, in most cases the phase dispersion of the multilayer reflectors at the ends of the cavity tends to *narrow* the transmittance of the filter—hence, the term *phase-dispersion narrowing*.

7.2.3.3 Phase dispersion of a multilayer treated mathematically

A mathematical analysis of the effect of phase dispersion on the Q of the cavity is now considered [see Eq. (2-83h)]. Figure 7-10 depicts the phase shift upon reflection δ vs wavelength for the three-layer reflector and seven-layer reflector containing layers of quarterwave optical thickness. Because the sign of the derivative (phase with respect to wavelength) $d\delta/d\lambda$ is positive, the sign of $d\delta/d\sigma$ (i.e., derivative with respect to wave number σ) is negative. A consequence is that the phase dispersion adds algebraically to the $2\pi n h$ term in Eq. (2-83h). The result is that as the phase dispersion increases the Q, a narrowing of the spectral passband width occurs.

7.2.3.4 Modification of the phase dispersion

§5.6.2 lists several methods of altering the phase dispersion. The magnitude of the phase dispersion is increased by: (1) using a quarterwave stack for a reflector that has a smaller difference in the refractive indices of the **H** and **L** layers that constitute the stack; and (2) inserting one or more layers of three (or five, seven, etc.) quarterwave optical thickness in the reflectors.

7.2.3.4.1 Phase-dispersion narrowing of the passband — second example

Figure 7-15 depicts the phase shift upon reflection of reflectors, each containing three layers, that are identical, except for the thickness of the outer layer

air 3**H L H** *glass* and *air* **H L H** *glass*.

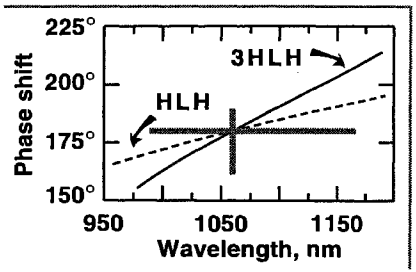


Fig. 7-15 Phase shift upon reflection of the stack (*solid*)
air 3H L H *glass*
and (*dashed*)
air H L H *glass*
where *glass,* L and H represent materials of refractive index 1.51, 1.45 and 2.30, respectively. L and H are layers of optical thickness λ_0 /4 at λ_0 of 1060 nm.

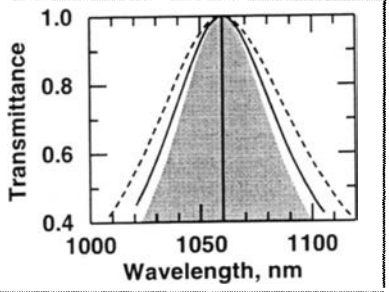


Fig. 7-16 $T(\lambda)$ of a Fabry-Perot (*dashed*)
glass H L H 2A H L H *glass*
and with a reflector that furnishes phase-dispersion narrowing (*shaded*)
glass H L 3H 2A 3H L H *glass*
and with less narrowing *(solid)*
glass H 3L H 2A H 3L H *glass*
where *glass,* L, H and A represent materials of refractive index 1.51, 1.45, 2.30 and 1.0003, respectively. L, H and A are layers of optical thickness λ_0 /4 at λ_0 of 1060 nm.

containing two coating materials. §7.5 illustrates that three or more coating materials may be used in the bandpass.

Although the reflectances and the phase shifts upon reflection are identical at λ_0, Fig. 7-15 shows that there is a larger slope of the phase shift upon reflection of the coating with the 3H layer on its exterior. This enhanced phase dispersion of the reflector narrows the spectral bandwidth, as shown by the shaded curve in the Fabry-Perot transmittance curve in Fig. 7-16.

glass H 3L H 2A H 3L H *glass* .

The solid curve in Fig. 7-16 shows that there is phase-dispersion narrowing, but a smaller amount, as compared to the bandpass in which the 3H layer is adjacent to the spacer.

A final comment is that the phase dispersion of the reflectors of the *phase conjugate filter* (see §7.9.1) is quite large and accounts, in large part, for the narrow spectral bandwidth of that device.

7.2.3.4.2 Phase-dispersion narrowing of the passband — third example

Suppose that the narrowing of the coating in §7.2.3.4.1 is excessive and a smaller narrowing of the passband is desired. The three quarterwave thickness layer has the maximum effect when it is located adjacent to the spacer, as in §7.2.3.4.1. If the three quarterwave thickness layer is moved farther away from the spacer, it still produces a narrowing of the passband, but less of it. An example is where the layer that is *nearly* adjacent to the spacer is converted from a quarterwave optical thickness to a three quarterwave optical thickness, as in the design H 3L H.

7.2.3.5 Summary

Table 7-2 summarizes the change in spectral bandwidth produced by the three control methods set forth in §7.2. Other sections in Chap. 7 apply the same principles to multiple-cavity filters.

7.3 Periodic-structure bandpass design method

§7.3 provides an overview of the periodic-structure design method. In §7.4, bandpasses are designed

7.3.1 Periodic-structure method of bandpass design — overview

Assuming that the bandpass is surrounded by optical cement of the same refractive index as that of the substrate, the bandpass consists of two components: an *antireflection* **AR** and the *periodic structure*

optical cement **AR (periodic structure) AR** *substrate* .

The *periodic structure* is, by definition, a *basic period* that is repeated many times.

An example is a bandpass centered at 633 nm. Assume that layers of refractive index 1.35 and 2.30 are used in the coating. Some design rules follow:

- The spectral width of the passband is controlled by the number of layers in the *basic period* of the periodic structure.
- The offband rejection (at wavelengths other than the passband) is controlled by the number of layers in the basic period of the periodic structure and by the number of times that basic period is repeated.
- The passband transmittance is controlled by the antireflection (AR stack) deposited on the exterior of the periodic structure.

Each is discussed below.

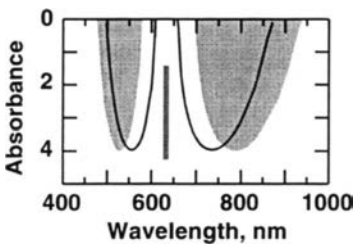


Fig. 7-17 $T(\lambda)$ of "five layers in the basic period" (*solid*)
glass (**H L H L H**)^6 *glass*
and "seven layers in the basic period" (*shaded*)
glass (**H L H L H L H**)^4 *glass* .

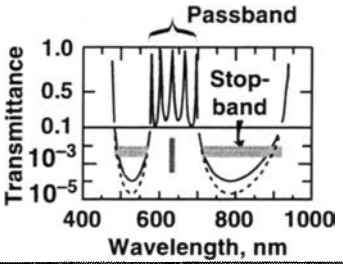


Fig. 7-18 $T(\lambda)$ of the bandpass (*solid*)
glass (**H L H L H**)^6 *glass*
and (*dashed*)
glass (**H L H L H**)^7 *glass* .
The scale of the ordinate changes from linear to log at 0.1.

For both Fig. 7-17 and \Fig. 7-18, the refractive indices of *glass*, **H** and **L** are 1.52, 2.30 and 1.35, respectively. **L** and **H** represent layers of optical thickness $\lambda_0/4$ at λ_0 of 633 nm.

7.3.1.1 Spectral width of the passband

The passband spectral width is controlled by the number of layers in the basic period. Figure 7-17 graphs the transmittance (on an absorbance scale) of the periodic structure consisting of the five layer basic period (**H L H L H**)^6 and a seven-layer basic period (**H L H L H L H**)^4. The spectral width of the passband of the "seven-layer basic period" is narrower than that of the "five layer basic period." The selection of the basic period is investigated further in §7.4.1.

7.3.1.2 Offband rejection of the bandpass

Given the five layer basic period cited above, Fig. 7-18 illustrates that the maximum offband rejection increases from a transmittance of 10^{-4} to 2×10^{-5} as the number of basic

periods increases from six to seven. The offband rejection is quantified by the parameter D_{max} in Table 7-3.

7.3.1.3 Effect of the antireflection layer upon passband transmittance

For the "naked" periodic structure, Fig. 7-18 illustrates that the transmittance in the passband spectral region manifests violent oscillations and has a relatively low average value. A layer of quarterwave optical thickness and (fictitious) refractive index 3.18 is appended to each end. The result is that passband transmittance increases to greater than 95%, as depicted in Fig. 7-19.

7.3.2 Periodic-structure method of bandpass design — procedure

The goal is the design of a bandpass with its passband centered at a tuned wavelength λ_0 with a prescribed spectral bandwidth and offband rejection. The procedure below follows the publication of Thelen [68b]:

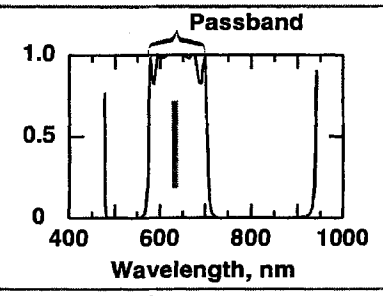


Fig. 7-19 $T(\lambda)$ of the bandpass *glass* **M** (**H L H L H**)^6 **M** *glass* where the refractive indices of *glass*, **L**, **M** and **H** are 1.52, 1.35, 3.18 and 2.30, respectively. **M**, **L** and **H** represent layers of optical thickness $\lambda_0/4$ at λ_0 of 633 nm.

(1) Select the available coating materials and establish their refractive indices.

(2) Examine a variety of *basic periods*. Each stack contains an odd number of layers of optical thicknesses $\lambda_0/4$ and is symmetrical about the center layer, as illustrated in §2.6.7.5.4. §5.3.7 shows that each stack has a passband at λ_0 and stop bands at both shorter and longer wavelengths.

(3) For each of the stacks chosen in step (2), determine the bandwidth BWS between the stop bands adjacent to λ_0.

(4) Select a *basic period* with a BWS slightly larger (perhaps 1%) than the target bandwidth of the bandpass under design. §7.3.2.1 and §7.5.1.1 illustrate this technique.

Table 7-3 Optical properties of basic periods of periodic structures

			D_{max}		
Basic period	n_e	BWS in %	$q=2$	$q=3$	$q=4$
HLHLH	6.7	22	1.2	1.9	2.6
HLHLHLH	11.4	12	1.9	2.9	3.9
HLHLHLHLH	19.3	6.6	2.7	4.0	5.4
HLHLHLHLHLH	33.0	3.7	3.5	5.2	6.9

Properties of periodic structures in which the basic period (column one) is repeated:

air (basic period)^q *air* .

L and **H** represent layers of optical thickness $\lambda_0/4$ and refractive indices of 1.35 and 2.30, respectively. The BWS is the fractional width (in percent) between the stop bands contiguous to λ_0. Its equivalent index n_e is computed at λ_0. The *maximum absorbance* D_{max} is the lowest transmittance occurring at wavelengths longer or shorter than λ_0.

(5) Cascade a sufficient number of the (identical) *basic periods* in tandem to provide the offband rejection (i.e., low transmittance) that is needed. The resulting multilayer is termed the *periodic-structure stack.* This is similar to the procedure of §5.3.7.1 in which the desired reflectance determined the number of periods.
(6) Find the equivalent index at λ_0 of the *periodic-structure stack* using Eqs. (2-47f) or (2-47g).
(7) Regarding the equivalent index in step (6) as a pseudo-substrate, design an AR (designated as AR_1) that matches the index of the *periodic-structure stack* to the index of the incident medium.
(8) Using the equivalent index cited in step (6) as a pseudo-incident medium, design an AR (designated as AR_2) that matches the index of the *periodic-structure stack* to the index of the substrate. The bandpass is now complete and takes the form

incident $\mathbf{AR_1}$ **(periodic-structure stack)** $\mathbf{AR_2}$ *substrate*.

§7.4 develops coatings in which the *periodic-structure stack* contains only two materials. Three material designs are considered in §7.5.

7.3.2.1 Determination of the BWS

The first step is to compute the *trace* [using Eq. (2-41e)] of the characteristic matrix of each *periodic-structure stack.* The wavelengths at which its magnitude equals 2.00 determine the locations of the stop bands. It is also possible to find the phase thickness β_e of the equivalent layer [see Eq. (2-47a)]. An imaginary β_e implies the onset of a stop band.

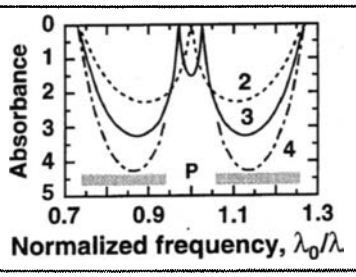


Fig. 7-20 Transmittance (on absorbance scale) of the basic period *air* $(\mathbf{HLHLHLH})^q$ *air* for $q = 2$ (*short dash*), 3 (*solid*) and 4 (*long-short dash*). **H** and **L** are layers of optical thickness $\lambda_0/4$ and refractive indices of 2.30 and 1.35, respectively. *P* is the passband and the shaded areas are stop bands.

7.4 Filter design using two components

The requirements for the bandpass are: (1) it is cemented—the incident medium and the substrate index are 1.52; (2) the layers are composed of cryolite ($n = 1.35$) and zinc sulfide ($n = 2.30$); (3) its spectral bandwidth should be close to 3%; and (4) the transmittance in the offband region (i.e., rejection) should be well below 10^{-3}. In a second example, the same materials are used to fabricate a filter with 7% spectral bandwidth.

7.4.1 Selection of the periodic structure

Both Fig. 7-18 and Fig. 7-20 graph transmittance as the number of periods is varied. As compared to Fig. 7-18, Fig. 7-20 uses seven layers in the basic period and frequency on the abscissa. Figure 7-20 graphs the absorbance vs the number of periods q for the coating

air $(\mathbf{HLHLHLH})^q$ *air*.

The spectral location of the stop bands is shown. The passband *P* is nested between the two adjacent stop bands. The transmittance drops monotonically in the stop band spectral regions. The maximum offband rejection D_{max} exceeds 4.0 for four periods. As in Fig.

7-18, the transmittance in the passband spectral region is quite poor, which is to be expected because the ARs have yet to be added to the bandpass.

Table 7-3 lists the BWS and offband rejection of *periodic-structure stacks* in which the number of basic periods is varied. The eleven layer stack has a BWS of 3.7%, which fulfills the requirement that the BWS should be larger than the target spectral bandwidth of 3%. A bandpass with two periods provides a maximum offband rejection of 3.5 absorbance units, which exceeds the requirement of 3 units.

7.4.1.1 Equivalent index of the periodic-structure stack

Equation (2-47g) is used to find the equivalent index n_e of one period of this *periodic-structure stack*, which (at λ_0) is

$$n_e = n_H^6 \, n_L^{-5} = (2.30)^6 \, (1.35)^{-5} = 33 \; . \qquad (7\text{-}1)$$

It is the task of the ARs adjacent to each side of the *periodic-structure stack* to match the foregoing index (i.e., the pseudo-substrate, $n_s = 33$) to the incident medium of index $n_0 = 1.52$. Both a single-section AR and a three-section AR are used.

Table 7-4 Optical properties of bandpasses

AR sections	Layers	BW	Centiwidth, %	Shape factor	Maximum absorbance
1	29	2.82%	5.66	2.00	4.7
2	45	3.06%	5.50	1.80	6.0
3	45	3.32%	5.50	1.66	6.95
1*	29	1.92%	3.76	1.96	4.9

BW is the full spectral width of the passband at $T = 0.50$ and the *centiwidth* its full width at $T = 0.01$. The *shape factor* is S_2 [see §1.6.4(f)]. The *maximum absorbance* is the smallest transmittance occurring at wavelengths longer than the passband. In the designs, the symbols **L** and **H** have the same meaning as in the caption to Table 7-3. *Cement* and *substrate* represent media of refractive index 1.52. The design (29 layers) of the bandpass with the single-section AR is

Cement **H L H L H** (**H L H L H L H L H L H**)^2 **H L H L H** *substrate*.

The design of the bandpass with the two-section AR is

Cement **H** 0.384**L** 0.21**H** 0.384**L HH L H** 1.3**L** 0.358**H** 1.3**L HL H** (**H L H L H L H L H L H**)^2
H L H 1.3**L** 0.358**H** 1.3**L HL HH** 0.384**L** 0.21**H** 0.384**L H** *substrate*.

The design of the bandpass with the three-section AR is

Cement **HH L H L HH L HL H L H L H** (**H L H L H L H L H L H**)^2
H L H L H L H L HH L H L HH *substrate*.

The design of the bandpass with the single-section AR (tagged with an asterisk) is

Cement **H L H LLL H** (**H LLL H L H L H LLL HH**)^2 **H LLL H L H** *substrate*.

7.4.2 Use of a single-section AR

From Eq. (2-23j) the index n_2 of the section of the AR is

$$n_2 = \sqrt{n_0 \, n_s} = \sqrt{(1.52)\,(33)} = 7.1 \; . \qquad (7\text{-}2)$$

Its optical thickness is a quarterwave or odd multiple thereof. The foregoing value is clearly beyond the range of available coating materials and must, therefore, be synthesized with an equivalent layer. At λ_0, the sequence of layers **H L H L H** has an equivalent index [see Eq. (2-47g)] of

$$n_e = n_H^3\ n_L^{-2} = (2.30)^3\ (1.35)^{-2} = 6.7 \qquad (7\text{-}3)$$

which patently falls short of the *desideratum* of 7.1. How much residual reflectance occurs if this lower index of 6.7 is used, rather than the correct value? From Eq. (2-48b), the reduced admittance y of a single layer of index 6.7 deposited on a substrate of index 33 (with an incident medium index of 1.52) is

$$y = (6.7)^2\ (33)^{-1}\ (1.52)^{-1} = 0.895\ . \qquad (7\text{-}4)$$

The substitution of the foregoing into Eqs. (2-48a) and (2-12c) shows that the resulting reflection loss *at this interface* is only 0.3%, which is acceptably low.

Two periods are used in the periodic-structure stack. The final bandpass design is

cement **HLHLH...(HLHLHLHLHLH)^2...HLHLH** *glass*

where the dots separate the five-layer AR *stacks* from the 21 layer *periodic-structure stack*. The transmittance appears in Fig. 7-21. Table 7-4 lists its optical performance. Although its maximum offband rejection of 4.7 absorbance units is adequate, it is unfortunate that its spectral bandwidth of 2.7% is smaller than the *desideratum* of 3%.

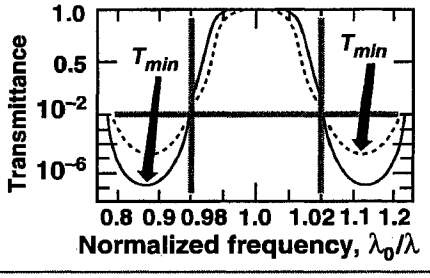


Fig. 7-21 $T(\lambda_0/\lambda)$ of bandpasses whose designs are captioned in Table 7-4 as the "bandpass with the single-section AR" (*dashed*) and the "bandpass with the three-section AR" (*solid*). The abscissa changes scale at 0.98 and at 1.02. The scale of the ordinate changes from linear to logarithmic at 0.01. The log of the reciprocal of T_{min} is the *maximum absorbance* cited in the caption to Table 7-4.

7.4.3 Use of a three-section AR

A three-section AR is interposed between the *periodic-structure stack* and the surrounding media. The spectral bandwidth over which the AR must function is small, which suggests that a maximally flat AR (see §4.5.1) should be used. Each section of the AR has a phase thickness of 90° or an *odd* multiple thereof. According to Table 4-3, the index of the section adjacent to the incident medium is

$$n_1 = n_0^{7/8}\ n_s^{1/8} = (1.52)^{7/8}\ (33)^{1/8} = 2.2 \qquad (7\text{-}5)$$

and the index of the section next to the substrate is

$$n_3 = n_0\ n_s\ n_1^{-1} = 1.52 \times 33 / 2.2 = 22.8\ . \qquad (7\text{-}6)$$

A single **H** has an equivalent index (which is its actual index) at λ_0 of 2.3 and the sequence of nine layers **H L HL HL HL H** manifests an equivalent index of

$$n_e = n_H^5\ n_L^{-4} = (2.30)^5\ (1.35)^{-4} = 19.4, \qquad (7\text{-}7)$$

which is a bit lower than the required index of 22.8. The index of the center section appears in Eq. (7-2). This section is designed in §7.4.2. The final bandpass is

[H][HLHLH][HLHLHLHLH](HLHLHLHLHLH)^2[HLHLHLHLH][HLHLH][H]

where the incident medium and substrate have an index of 1.52. The brackets in the previous design separate the boundaries of the sections of the ARs. The transmittance appears in Fig. 7-21.

7.4.4 Comparison of bandpasses

Table 7-4 summarizes the attributes of the bandpasses of §7.4.2 and §7.4.3. The latter contains 16 more layers than the former. The manifestations of the additional layers are: (1) the spectral bandwidth widens from 2.8% to 3.3%; and (2) the maximum offband rejection increases by at least two absorbance units.

The additional layers also make the spectral shape of the passband "squarer," as manifested by a smaller *centiwidth* in Table 7-4, which is the full width of the passband when the transmittance is 1%. Table 7-4 shows that the *shape factor* [see §1.6.4(f)] decreases from 2.0 to 1.66 as more layers are added.

7.4.5 Use of a two-section AR

Both aforementioned bandpasses manifest adequate offband rejection but patently fail to meet the *desideratum* (see §7.4) of a 3% spectral bandwidth. A solution is suggested by the spectral bandwidths in Table 7-4. The filter with a single-section AR has too narrow a spectral bandwidth, whereas the spectral bandwidth of the triple-section AR is too large. Because these bracket the *desideratum*, is it possible that a bandpass with two-section ARs would hit the mark?

As in §7.4.3, a maximally flat AR is used. According to §4.5.1, the refractive index of the outer section should be

$$n_1 = n_0^{3/4}\ n_s^{1/4} = (1.52)^{3/4}\ (33)^{1/4} = 3.28 \qquad (7\text{-}8)$$

and for the inner section

$$n_2 = n_0^{1/4}\ n_s^{3/4} = (1.52)^{1/4}\ (33)^{3/4} = 15.3\ . \qquad (7\text{-}9)$$

The equivalent indices of the stacks **H L H** and **H L H L H L H L H** are 3.92 and 19.4, respectively. Each exceeds the respective target value in Eqs. (7-8) and (7-9). One design technique is to replace the center layer with a film of intermediate index **E** (of *index* n_E) so that the nine layer stack becomes **H L H L E L H L H**. Equation (2-47g) gives the equivalent index n_e of this stack as

$$n_e = n_H^4\ n_L^{-4}\ n_E\ . \qquad (7\text{-}10)$$

The solution of Eq. (7-10) with its left equal to 15.3 results in an **E** layer index of 1.81.

A similar calculation using Eq. (2-47f) is made for the outer section of the AR. The layer sequence **H E H** is to have an index of 3.28. A calculation similar to that in the previous paragraph results in a center layer index of 1.61. The **E** layers are themselves simulated by three-layer groups (see §2.6.7.1) that have an equivalent phase thickness of 90° and an equivalent index of either 1.81 or 1.61, for the inner and outer sections, respectively. Table 7-4 lists the properties of the bandpass, which contains 45 layers. The spectral bandwidth of 2.8%, as shown by the transmittance curve in Fig. 7-22, is quite close to the goal of 3%.

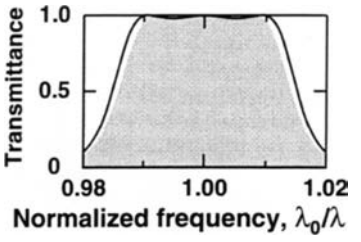


Fig. 7-22 $T(\lambda_0/\lambda)$ of bandpasses whose designs are captioned in Table 7-4 as the "bandpass with the single-section AR" (*solid*) and the "bandpass with the two-section AR" (*shaded*).

Figure 7-23 depicts the basic designs of these bandpasses, in which the cascaded periodic structures (the periodic-structure stacks) are surrounded by antireflection coatings — the ARs. The number of sections in the ARs progresses from one to three.

7.4.6 Phase-dispersion narrowing of the passband

An example illustrates that phase dispersion narrows the passband. Rather than the 2.8% spectral bandwidth achieved in §7.4.5, suppose that 2% spectral bandwidth is requisite. The most severe narrowing of the spectral bandwidth occurs when each of the quarterwave layers contiguous to a halfwave thickness layer is replaced by a film of three quarterwave thickness (symbol: **T**). The modified coating is

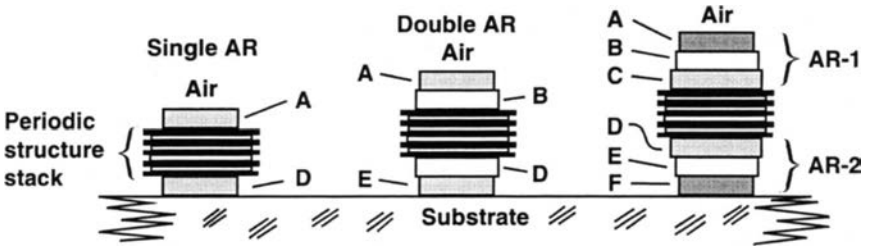


Fig. 7-23 The construction of three bandpasses using the periodic-structure method, in which a *periodic-structure stack* is sandwiched between antireflection multilayers of one, two and three sections.

cement **H L H T H**...**(H T H L H L H L H T H)**^{2}...**H T H L H** *glass*.

Figure 7-24 and Table 7-4 show that the spectral bandwidth is narrowed considerably — to 1.92%. A spectral bandwidth closer to 2% is achieved by discarding the two outermost three quarterwave layers, as in

cement **H L H L H**...**(H T H L H L H L H T H)**^{2}...**H L H L H** *glass*.

In other words, the use of fewer three quarterwave thickness layers decreases the amount of phase-dispersion narrowing.

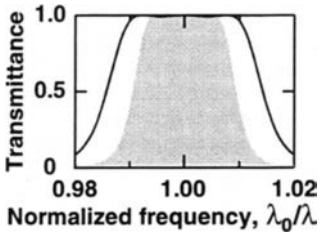


Fig. 7-24 $T(\lambda_0/\lambda)$ of bandpasses, the designs of which are captioned in Table 7-4 as the "bandpass with the single-section AR" (*solid*) and the "bandpass with the single-section AR (tagged with an asterisk)" (*shaded*).

7.4.6.1 Consequences of phase-dispersion narrowing

Although phase dispersion quite effectively reduces the spectral bandwidth of a filter, the designer should keep in mind the other consequences: (1) thicker layers must be deposited, which are, of necessity, more difficult to control; (2) the width of the offband spectral region over which the transmittance is low (i.e., the spectral width of the *blocking* region) is reduced; and (3) ripple may be introduced into the passband.

7.4.7 Bandpass with 7% spectral bandwidth

The environment of this bandpass differs slightly from that of previous examples, in that the incident medium is air and the substrate is soda-lime glass. The first step is to select a basic

period from Table 7-3. Clearly, the combination **H L H L H L H L H** is not acceptable because its BWS of 6.6% fails to exceed that of our *desideratum* of 7%. The use the sequence of seven layers **H L H L H L H** for the basic period is mandated. Table 7-3 shows that its BWS is quite large — 12%.

The equivalent index of the basic period is 11.4 at $\lambda = \lambda_0$. The use of a single-section AR — similar to the coating of §7.4.2 — produces a bandpass with excessive ripple in the transmittance at the edges of the passband. A three-section AR reduces the ripple.

Assume that there is air on both ends of the bandpass. The periodic-structure stack is the pseudo-substrate of index 11.4. Each of the three sections of the A R (assumed to be a maximally flat design) has a phase thickness of 90° or an *odd* multiple thereof. Following the procedure of §7.4.3, the section next to air should have an index

$$n_1 = n_s^{1/8} = (11.4)^{1/8} = 1.36 , \qquad (7\text{-}11)$$

which is close enough to that of an **L** layer.

The index of the center section of the AR should be

$$n_2 = \sqrt{n_s} = \sqrt{11.4} = 3.38 , \qquad (7\text{-}12)$$

which is lower than the equivalent index of the three-layer group **H L H** of 3.91. The design for the center section is then **H E H**, where **E** is an equivalent layer with 90° phase thickness. The index of the equivalent layer is obtained from Eq. (2-47f)

$$n_e = n_H^2 \ (3.38)^{-1} = (2.30)^2 \ (3.38)^{-1} = 1.57 . \qquad (7\text{-}13)$$

The index of the section next to the periodic-structure stack — the pseudo-substrate — should be

$$n_3 = n_s^{7/8} = (11.4)^{7/8} = 8.41 , \qquad (7\text{-}14)$$

which exceeds the equivalent index of **H L H L H**. The stack is modified to **H L H E H L H**, where **E** is an equivalent layer of 90° phase thickness. Solving Eq. (2-47f) for n_e leads to

$$n_e = n_H^4 \, n_L^{-2} \ (8.41)^{-1} = (2.30)^4 \ (1.35)^{-2} \ (8.41)^{-1} = 1.83 . \qquad (7\text{-}15)$$

The periodic-structure stack contains three periods. Because the foregoing design goes "air-to-air" (i.e., an index of 1.00 for both the incident and emergent medium), the admittance of air is simulated with glass overcoated with a single **L** layer. The final bandpass is captioned in Fig. 7-25 and is symmetrical when the aforementioned **L** layer (next to the glass) is omitted. The problem is that its spectral bandwidth of 10.4% exceeds the *desideratum* of 7%.

The spectral bandwidth is reduced by inserting higher order spacers (see §7.2.1) or by phase-dispersion narrowing (see §7.2.3). The latter method is used. A restriction is that each of the three sections of the periodic-structure stack should contain layers of quarterwave (or odd multiple thereof) optical thickness. If **T** represents a layer of three quarterwave optical thickness and index 2.30, three possibilities for the periodic-structure stack design are:

T L H L H L T...	**T L H L H L T...**	**T L H L H L T**	(one)
H L H L H L H...	**T L H L H L T...**	**H L H L H L H**	(two)
T L H L H L T...	**H L H L H L H...**	**T L H L H L T**	(three).

There is little use in computing the spectral bandwidth of each of these combinations. The last has the requisite 7% spectral bandwidth, as is evident from its $T(\lambda_0/\lambda)$ curve in Fig. 7-25. The maximum offband rejection of 5.2 absorbance units is adequate.

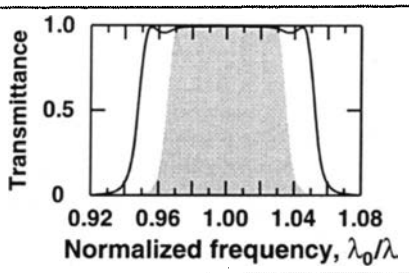


Fig. 7-25 $T(\lambda_0/\lambda)$ of (*shaded*) *air* **L** 1.231**H** 0.494**L** 2.231**H L** 1.348**H** 0.271**L** 1.348**H L H** (3**H L H L H L** 3**H**) (**H L H L H L H**) (3**H L H L H L** 3**H**) **H L** 1.348**H** 0.271**L** 1.348**H L** 2.231**H** 0.494**L** 1.231**H** *Substrate.*
Substrate, **L** and **H** represent materials of refractive indices 1.52, 1.35 and 2.30, respectively. **L** and **H** represent layers of optical thickness $\lambda_0/4$. For the *solid curve*, all four 3**H** layers in the foregoing bandpass are replaced with **H**.

7.5 Periodic structures containing three materials

The design procedure of §7.3.2 is modified so that the *periodic-structure stack*

incident $\mathbf{AR}_1$
(periodic-structure stack)
$\mathbf{AR}_2$ *substrate*

is allowed to contain more than just two coating materials.

7.5.1 Design procedure

The procedure below uses the concept of BWS [see §7.3.2 step (3)]. Other procedures for bandpass design have been published by Thelen [63, 89] and Jacobs [81].

7.5.1.1 Selection of the periodic-structure stack

In order to achieve a bandpass with a reasonable amount of offband rejection, a rule of thumb is that the *periodic-structure stack* should contain at least seven layers. Jacobs [81] published a computer code that permutes the three materials used in the coating. Table 7-5 lists many of them for a sequence of nine layers. **H**, **L** and **M** represent the high, low and intermediate index layers, respectively. The bandwidth BWS between the stop bands increases from 2.9% for the nine-layer sequence **H L H L H L H L H** to as much as 18% for stacks that are essentially useless because of their poor offband rejection (see §7.5.3). Table 7-5 lists the equivalent index at λ_0. A few guidelines for the selection of a *periodic-structure stack* are:

(1) Select a *periodic-structure stack* that has a BWS that exceeds the bandwidth of the final filter by an amount 0.5% to 1.5%.
(2) Avoid *periodic-structure stacks* in which an **M** layer is sandwiched between an **L** and **H**. Eschew a design with the sequence **H M L** or **L M H**, for example.
(3) The outer layer should be an **H** layer if the goal is a bandpass in which there is minimal shift of the passband wavelength with increasing angle of incidence.
(4) Amongst those *periodic-structure stacks* having a BWS close to the desired value, the one with the largest equivalent index provides the best offband rejection.

Table 7-5 Optical properties of periodic structures

Design	n_e	Recip.	BWS	Design	n_e	Recip.	BWS
LHLHLHLHL	0.08	13.0	2.91	LHLMLMLHL	0.25	3.9	8.29
HLHLHLHLH	98.9	98.9	2.91	HMLHMHLMH	1.52	1.5	8.50
MHLHLHLHM	0.11	8.8	3.23	LMHLMLHML	7.47	7.5	8.66
LHLHMHLHL	0.09	10.7	3.51	MHLMLMLHM	0.37	2.7	9.09
HMHLHLHMH	66.9	66.9	3.75	MLMHLHMLM	0.82	1.2	9.28
MLHLHLHLM	29.8	29.8	3.88	MLMLHLMLM	8.97	9.0	9.31
MHLHMHLHM	0.13	7.2	3.88	HLMHLHMLH	2.73	2.7	10.2
LHMHLHMHL	0.11	8.8	4.04	HMHMLMHMH	20.4	20.4	10.3
HLHMHMHLH	66.9	66.9	4.16	MLHMLMHLM	9.08	9.1	10.7
MHMHLHMHM	0.16	6.0	4.45	MLMHMHMLM	1.00	1.0	11.0
LHMHMHMHL	0.13	7.2	4.83	HLMHMHMLH	3.32	3.3	12.1
HLHLMLHLH	54.3	54.3	5.06	HLMLMLMLH	16.3	16.3	12.2
LMHLHLHML	13.6	13.6	5.11	LMHMLMHML	4.15	4.2	13.6
MHMHMHMHM	0.20	4.9	5.32	LHLMHMLHL	0.56	1.8	14.4
HMHMHMHMH	45.2	45.2	5.32	MHMLHLMHM	1.82	1.8	15.0
MLHMHMHLM	20.1	20.1	5.50	LMLMLMLML	0.84	1.2	15.1
LMLHLHLML	0.25	3.9	5.82	MLMLMLMLM	4.92	4.9	15.1
HMHLMLHMH	36.7	36.7	6.45	HMLMHMLMH	9.20	9.2	15.2
MLHLMLHLM	16.4	16.3	6.67	LHMLHLMHL	1.23	1.2	15.3
LMLHMHLML	0.31	3.2	6.94	MHLMHMLHM	0.83	1.2	15.6
HMLHLHLMH	1.25	1.3	7.14	LHMLMLMHL	0.67	1.5	16.2
LMHMHMHML	9.20	9.2	7.16	MHMLMLMHM	1.00	1.0	17.5
HLMLHLMLH	29.8	29.8	7.46	HMLMLMLMH	4.15	4.2	17.9
HLHMLMHLH	30.2	30.2	8.26	LMLMHMLML	1.87	1.9	18.3

For basic periods of periodic structures, equivalent index n_e and its reciprocal *Recip.* (if it is fractional). BWS is the full bandwidth between the stop bands adjacent to λ_0, expressed as a percent of λ_0. **L**, **M** and **H** represent layers of optical thickness $\lambda_0/4$ and refractive indices 1.85, 2.25 and 4.1, respectively.

7.5.2 Design example—7% spectral bandwidth

The requirements for the bandpass are: (1) the substrate is bulk germanium ($n_s = 4.00$); (2) the materials in the coating are silicon monoxide for **L** ($n_L = 1.85$), zinc sulfide for **M** ($n_M = 2.25$) and germanium for **H** ($n_H = 4.1$); (3) the spectral bandwidth should be close to 7%; (4) the wavelength center is at 4 μm; and (5) the transmittance in the offband region (i.e., the rejection) should be below 10^{-3}.

The *periodic-structure stack* **H L M L H L M L H** (see Table 7-5) has a BWS of 7.46% and equivalent index of 29.8 at λ_0. Hopefully, the ARs will narrow the spectral bandwidth to the target value of 7%. A trial design uses two periods for the *periodic-structure stack* and single-section ARs. The resultant spectral bandwidth is only 5.7% — considerably less than the 7% *desideratum*.

The spectral bandwidths of the bandpasses in §7.4.2 and §7.4.5 broaden as the number of the sections in the ARs progresses from one to three, which leads the designer to use a three-section AR. The steps of §7.4.2 are followed and a three-section maximally flat AR is designed for each end of the *periodic-structure stack*.

The first step is to design an AR_1 that matches the *periodic-structure stack* to air. The index n_1 of the section adjacent to the incident medium should be

$$n_1 = n_0^{7/8}\ n_s^{1/8} = (1.00)^{7/8}\ (29.8)^{1/8} = 1.53 \qquad (7\text{-}16)$$

and is simulated with an equivalent layer with 270° phase thickness. The index n_2 of the center section is

$$n_2 = \sqrt{n_0\, n_s} = \sqrt{(1.00)\,(29.8)} = 5.5\ , \qquad (7\text{-}17)$$

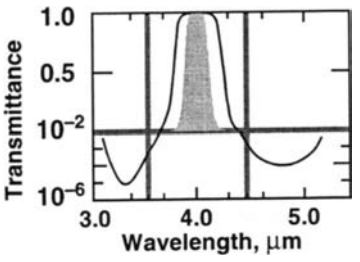


Fig. 7-26 $T(\lambda)$ of the bandpass (*solid*)
air 1.41**L** 0.14**H** 1.41**L** 1.36**H** 0.22**L** 2.36**H L H L H** (**H' L M L H L M L H'**)^2 **H L H L** 2.3**H** 0.321**L** 1.3**H L M** *Ge*
where *Ge*, **L**, **M**, **H'** and **H** represent materials of refractive indices 4.0, 1.85, 2.25, 4.10 and 4.10, respectively. **L**, **M**, **H'** and **H** represent layers of optical thickness 1.0 μm. For the *shaded* curve, the **H'** layers are replaced by 3**H**. The abscissa changes scale at 3.7 μm and 4.3 μm. The ordinate changes scale from linear to logarithmic at 0.01.

which is simulated by the three-layer sequence **H E H**, where the **E** layer has an equivalent index of 3.06 and phase thickness of 90°. The index n_3 of the section of the AR next to the *periodic-structure stack* (which is a pseudo-substrate) should be

$$n_3 = n_0^{1/8}\ n_s^{7/8} = (1.00)^{1/8}\ (29.8)^{7/8} = 19.5\ . \qquad (7\text{-}18)$$

From Eq. (2-47g), the five layer sequence **H L H L H** has an equivalent index of 20.1, which is quite close to the target.

The next step is to design AR_2 that matches the *periodic-structure stack* to the substrate. The substrate is coated with a zinc sulfide layer (*index* $n_M = 2.25$) of quarterwave optical thickness at λ_0. The normalized admittance Y_4 of this pseudo-substrate is

$$Y_4 = n_M^2\ n_s^{-1} = (2.25)^2\ (4.00)^{-1} = 1.27\ . \qquad (7\text{-}19)$$

The index n_1 of the section of the AR adjacent to the *periodic-structure stack* (which is a pseudo-incident medium) is

$$n_1 = n_0^{1/8}\ n_s^{7/8} = (1.27)^{1/8}\ (29.8)^{7/8} = 20.3\ . \qquad (7\text{-}20)$$

This section may be approximated with the five layer sequence **H L H L H**, which has an equivalent index of 20.1. The index n_2 of the center section of the AR is

$$n_2 = \sqrt{n_0\ n_s} = \sqrt{(1.27)\,(29.8)} = 6.19 \qquad (7\text{-}21)$$

which is simulated by the three-layer sequence **H E H**, where the **E** layer has an equivalent index of 2.71 and phase thickness of 90°. The index n_3 of the section of the AR contiguous to the **M** layer coated on the germanium substrate is

$$n_3 = n_0^{1/8}\ n_s^{7/8} = (29.8)^{1/8}\ (1.27)^{7/8} = 1.88\ . \qquad (7\text{-}22)$$

This is quite close to the index of an **L** layer.

The transmittance of the final bandpass containing 34 layers appears in Fig. 7-26. The $T(\lambda)$ is asymmetrical about $\lambda_0/\lambda = 1$ because of the nonquarterwave thickness layers. The spectral bandwidth is 6.9%, which falls a bit short of the goal of 7%. It is difficult to widen the spectral bandwidth to 7% without increasing the indices of either the **L** or **M** layers, or decreasing the index of **H**. Easier said than done! The transmittance outside the passband is below 10^{-3}, to wavelengths as short as 3.3 μm and as long as 5.1 μm.

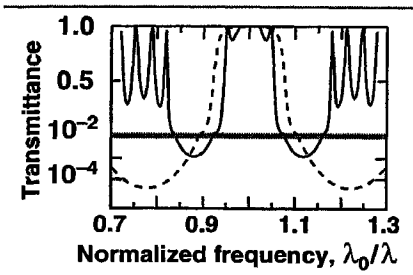


Fig. 7-27 $T(\lambda_0/\lambda)$ of the seven-cavity bandpass (*dashed*)
air 1.247**L** 0.5**M** 3.25**L H MM H M H MM H L H L L H L H M M H M H MM H** 2.354**L** 0.225**H** 0.354**L** *ZnSe*
and (*solid*)
air **H** (**H L M L M L M L H**)^4 1.294**H** 0.327**L** 1.294**H** *ZnSe*
where **L**, **M**, **H**, *air* and *ZnSe* represent materials with indices of 1.85, 2.25, 4.1, 1.00 and 2.4, respectively. **L**, **M** and **H** represent layers of optical thickness $\lambda_0/4$. The ordinate changes scale from linear to logarithmic at 0.01.

7.5.3 Design example — a bandpass with inferior offband rejection

The next bandpass serves to illustrate that, although the foregoing procedure is quite effective when the spectral bandwidth is relatively narrow, there may be problems when the spectral bandwidth is relatively wide.

The goal is a bandpass with approximately 10% spectral bandwidth. The incident medium is air and the substrate is zinc selenide (index 2.40). The first task is to select a periodic-structure stack from Table 7-5. The sequence **H L M L M L M L H** manifests a BWS of 12% and an equivalent index of 16.3. Single-layer ARs are used.

A layer of index $\sqrt{16.3} = 4.04$ matches the periodic-structure stack to air. A section of index $\sqrt{16.3 \times 2.4} = 6.26$ matches the periodic-structure stack to the substrate. The foregoing index is simulated with a three-layer sequence **H E H**, where **E** is an equivalent layer with an index of 2.69 and phase thickness of 90°.

Given the layer sequence **H-E-H** (deposited upon a zinc selenide substrate and with a pseudo-incident medium index of 16.3), its reduced admittance y is

$$y = n_H^4 n_E^{-2} n_0^{-1} n_s^{-1} = (4.1)^4 (2.69)^{-2} (16.3)^{-1} (2.4)^{-1} = 0.99 \qquad (7\text{-}23)$$

and is close to unity. This shows that the index of the **E** layer is computed correctly.

Four periods are used in the periodic-structure stack. The transmittance of the 35 layer bandpass is depicted in Fig. 7-27. Although its spectral bandwidth of 10.8% is close to the *desideratum*, there are annoying "wabbit ear" transmittance dips at the edges of the passband. These could be removed by using more layers in the ARs, although this is not done. Shown also in Fig. 7-27 is the transmittance of a 25 layer bandpass designed by the microwave method. Its maximum offband rejection is 4.4 absorbance units, as compared to 2.9 for the bandpass designed in §7.5.3. In other words, the bandpass derived from the microwave method contains ten fewer layers and manifests superior offband rejection.

Why does the bandpass derived from the periodic-structure method perform less than optimally? A hint is provided by examining the basic period **H L M L M L M L H**. Only two of the **L** layers (of index 1.95) are contiguous to a high-index **H** layer (thus producing a large Fresnel reflection coefficient). All of the other **L** layers are adjacent to an **M** layer (index 2.25), which, because of the relatively small index difference between the **L** and **M** layers, produces a relatively small Fresnel reflection coefficient and, hence, a less-than-optimal offband rejection.

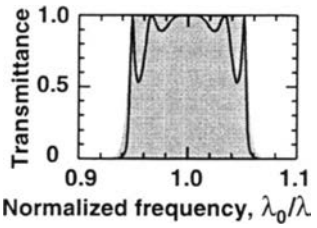


Fig. 7-28 $T(\lambda_0/\lambda)$ of (*solid*) *glass* **H L H (H L H L H L H)**^{6} **H L H** *glass* and (*shaded*) *glass* 1.276**H** 0.405**L** 2.276**H L** 1.431**H** 0.121**L** 1.431**H L** 2**H L** 1.203**H** 0.553**L** 1.203**H L** 2**H L H L H L** 2**H L H L H L** 2**H L** 1.203**H** 0.553**L** 1.203**H L** 2**H L** 1.431**H** 0.121**L** 1.431**H L** 2.276**H** 0.405**L** 1.276**H** *glass* where **L**, **H** and *glass* represent materials with indices of 1.35, 2.30 and 1.52, respectively. **L** and **H** represent layers of optical thickness $\lambda_0/4$.

7.5.4 Use of higher order spacers to narrow the spectral bandwidth

§7.2.1 illustrates that the use of a higher order spacer layer narrows the spectral bandwidth of a single-cavity bandpass. The same is true of a multiple-cavity bandpass, with the *caveat* that the equivalent index of the periodic-structure stack at λ_0 should not be perturbed. Consider the basic period of the bandpass developed in §7.5.2, which is **H L M L H L M L H**. Suppose that a halfwave of **H** (i.e., two QWOTs) were added to each end of the stack. The stack evolves to **3H L M L H L M L 3H**, which may be rewritten as **2H H L M L H L M L H 2H**. The latter emphasizes that an absentee layer **2H** is added to each end of the stack. According to Eq. (2-47g), these absentee layers do not alter the equivalent index of the periodic-structure stack at λ_0. The spectral transmittance curve in Fig. 7-26 illustrates that the passband spectral width is narrowed and also that addition of the **2H** layers does not alter spectrally the passband shape.

7.5.5 Improvement of the average transmittance of the passband

A problem associated with the periodic-structure design method is that the AR positioned at the ends of the periodic-structure stack functions, in theory at least, only at the center wavelength λ_0. A consequence is that when the passband spectral bandwidth is broad, "wabbit ear" dips in transmittance occur at the edges of the passband, as is evident in the solid curve in Fig. 7-28. These dips may be removed by optimization.

As an example, consider as a periodic-structure stack the seven-layer basic period appearing in Table 7-3. Its n_e at λ_0 is 11.4. The index of the single-section AR that matches the equivalent index of the periodic-structure stack to the incident medium of *glass* (index 1.52) is $\sqrt{11.4 \times 1.52} = 4.2$. A single AR of the design **H L H** has an equivalent index of 3.9, which is reasonably close to the *desideratum*. The solid curve in Figure 7-28 shows that the design

glass **H L H (H L H L H L H)**^{6} **H L H** *glass*

manifests severe "wabbit ear" dips. The solution is to modify the refractive indices of the outer sections of the periodic-structure stack. The foregoing design is rewritten

glass **H C H (H A)**^{3} **H (H B)**^{3} **H (H L)**^{3} **H (H L)**^{3} **H (H B)**^{3} **H (H A)**^{3} **H H C H** *glass*

where the refractive indices of **A**, **B** and **C** are variables. The refractive indices of **L** and **H** are not altered. The optical thicknesses of all layers remain constant. The refractive indices n_a, n_b and n_c were initially set at 1.35, 1.35 and 1.7, respectively.

An optimization computer code then increased the passband transmittance in the frequency region $0.95 \leq \lambda_0/\lambda \leq 1.00$. The refractive indices n_a, n_b and n_c were altered to 1.558, 1.405 and 1.656, respectively. In other words, after optimization, the bandpass contained layers with fictitious indices.

The next step is to replace those fictitious indices with real indices. The **C** layer is replaced by an equivalent layer of refractive index 1.656 and 90° phase thickness. Given that n_a is 1.558, at λ_0 the equivalent index [see Eq. (2-47g)] of that seven-layer sequence **H A H A H A H** is

$$n_e = n_H^4\, n_A^{-3} = 7.40 \ . \tag{7-24}$$

The next task is to replace the aforementioned stack with one containing only **H** and **L** layers, as well as a single equivalent layer **E** of refractive index n_E. That is, the stack design is now **H L H E H L H.** n_E is obtained from

$$n_e = n_H^4 n_L^{-2} n_E^{-1} = (2.30)^4\,(1.35)^{-2}\, n_E^{-1} = 7.40 \ . \tag{7-25}$$

Solving Eq. (7-25) for n_E results in 2.07. An equivalent layer with 90° phase thickness and this index is generated using the procedure of §2.11.3. It is inserted into the design, which appears in the caption to Fig. 7-28. The foregoing procedure is repeated for the **B** layer in the stack. The design of the resulting bandpass is captioned in Fig. 7-28.

A final comment relates to the selection of the wavelengths at which the transmittance is increased. The refractive indices are adjusted in a stack consisting of layers of only quarterwave optical thickness. Inspection of the *R/T* polynomial in §2.7.4.2 for such a multilayer shows that there is symmetry in the transmittance curve about the point where the phase thickness β is 90°. Thus it is necessary to adjust the transmittance at wavelengths shorter than the passband center, or wavelengths longer than the passband center, but not both simultaneously.

7.5.6 A bandpass viewed as a multiple-cavity filter

The multiple-cavity filter is defined as follows. If **R** represents a multilayer reflector and **S** a spacer (of halfwave optical thickness), the bandpass is constructed:

$\mathbf{R_0\ S\ R_1}$	(single cavity—Fabry-Perot)
$\mathbf{R_0\ S\ R_1\ S\ R_2}$	(two cavity)
$\mathbf{R_0\ S\ R_1\ S\ R_2\ S\ R_3}$	(three cavity)
$\mathbf{R_0\ S\ R_1\ S\ R_2\ S\ R_3\ S\ R_4}$	(four cavity)
$\mathbf{R_0\ S\ R_1\ S\ R_2\ S\ R_3\ S\ R_4\ S\ R_5}$	(five cavity)

and so on.

Inspection of the bandpasses of the foregoing sections shows that the addition of the ARs to the *periodic-structure stacks* has two effects: (1) the passband transmittance is increased and the spectral bandwidth altered; and (2) the offband rejection is enhanced. Figure 7-21 illustrates that the offband rejection is enhanced by adding more halfwave thickness layers—i.e., the addition of more cavities.

Bandpass with single-section AR	Bandpass with three-section AR
H L H L HH **L H L H L H L H L HH** **L H L H L H L H L HH** **L H L H**	**HH** **L H L HH** **L H L H L H L HH** **L H L H L H L H L HH** **L H L H L H L H L HH** **L H L H L H L HH** **L H L HH**

Consider the bandpasses captioned in Table 7-4. In the box above, the 29 layer bandpass is written on separate lines and the cavities of halfwave optical thickness are

underlined. Two contiguous quarterwave layers, as, for example **HH**, constitute a halfwave optical thickness layer. It is quite evident that it is a three-cavity filter.

The bandpass with three-section ARs is essentially a seven-cavity filter. Table 7-4 shows that the three-cavity filter has a maximum offband rejection of 4.7 absorbance units, whereas the maximum absorbance of the seven-cavity filter is 6.95. It is evident that the multilayer ARs contribute to the offband rejection. The conclusion is that the addition of even more cavities should enhance the offband rejection. A logical way of doing this is to use the microwave design procedure.

7.5.7 "Cavity"—possible meanings of this word

To some writers, a *cavity* is a Fabry-Perot, that is, a single-cavity filter. As illustrated in Fig. 7-29, the 29 layer "Bandpass with the single-section AR" (appearing in the caption to Table 7-4) is viewed as a chaining of three Fabry-Perot filters. Each Fabry-Perot consists of the nine layer sequence **HL H L HH L H L H**. Each of the outer Fabry-Perot filters is "linked" to the center filter by means of a single quarterwave thickness *L* layer. This concept is used by Cushing [98, 99a, 99b, 00a, 00b, 00c].



Fig. 7-29 The 29 layer "Bandpass with the single-section AR" appearing in the caption to Table 7-4 is viewed as three Fabry-Perot filters (i.e., cavities) separated by single *L* layers (arrowed, C).

The reason that §7.5.7 is included is to avoid confusion when the word *cavity* appears in the literature. In this book, *cavity* always refers to a single-layer whose optical thickness is a halfwave (or multiple thereof). *Cavity* does not refer to a group of layers that constitute a Fabry-Perot filter.

7.5.8 Microwave and periodic-structure design methods—a comparison

The "microwave bandpass design method" that is the focus of §7.6 is used to produce a bandpass that simulates the performance of the bandpass whose design appears in the caption to Fig. 7-19. The latter is produced by the "periodic-structure method." In the subsequent discussion, it is termed the "periodic BP." The former is tagged as the "microwave BP." §7.5.8 compares bandpasses produced by both methods. Their spectral properties are not dissimilar.

7.5.8.1 Bandpass designed by the periodic-structure method

The **M** layer in the design captioned in Fig. 7-19 is replaced by the two-layer sequence **E L**, where **E** is an equivalent layer. At the tuned wavelength of 633 nm, Eq. (2-48c) shows that the reduced admittance *y* produced by the **M** layer (of refractive index 3.18) is

$$y = n_M^2 n_H^{-1} \; n_0^{-1} = \; (3.18)^2 \, (2.30)^{-1} \, (1.52)^{-1} = 2.89 \qquad (7\text{-}26)$$

where n_H is the refractive index of the high-index spacer that is contiguous to the **M** layer and n_0 is the refractive index of the *glass* incident medium. When the refractive index of the **E** layer is 1.87, the reduced admittance of the aforementioned two-layer sequence is

$$y = n_E^2 n_L^{-2} \, n_H \; n_0^{-1} = \; (1.87)^2 \, (1.35)^{-2} \, (2.30) \, (1.52)^{-1} = 2.90 \, , \qquad (7\text{-}27a)$$

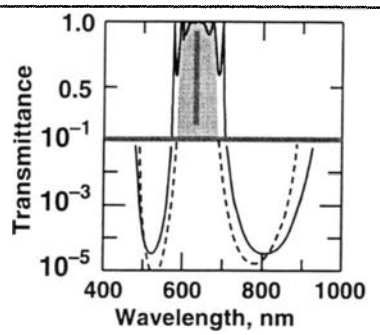


Fig. 7-30 $T(\lambda)$ of the periodic-structure bandpass (*solid*)
glass 0.277**L** 0.405**H** 1.277**L** **H** (**H L H L H**)^6 **H** 1.277**L** 0.405**H** 0.277**L** *glass*
and a bandpass produced by the microwave design method (*shaded* and *dashed*)
glass 0.293**L** 0.373**H** 1.293**L** 2**H L H** **L** 2**H** 1.147**L** 0.67**H** 1.147**L H** 2**L** (**H L**)^2 2**H** (**L H**)^2 2**L H** 1.147**L** 0.67**H** 1.147**L** 2**H L H L** 2**H** 1.293**L** 0.373**H** 0.293**L** *glass*
where **L**, **H** and *glass* represent materials with indices of 1.35, 2.30 and 1.52, respectively. **L** and **H** represent layers of optical thickness $\lambda_0/4$ at λ_0 of 633 nm.

which is close to the value in Eq. (7-26). The **E** layer is replaced by an equivalent layer. The transmittance of the completed bandpass appears as the solid curve in Fig. 7-30. The coating contains 31 layers, seven of which are the high-index spacers of halfwave optical thickness.

7.5.8.2 Bandpass designed by the microwave method

§7.6.2.3 defines the maximum standing wave ratio V_m of the prototype bandpass. §7.6.5 illustrates that the offband rejection of the bandpass depends upon V_m. A V_m of 1.58×10^7 was chosen and a seven-cavity bandpass was designed. Its transmittance appears as the *shaded* (and *dashed*) curve in Fig. 7-30.

7.5.8.3 Comparison of designs

In the comparison of the aforementioned bandpasses, one is identified as the "periodic BP" (designed by the procedure in §7.3.2) and the other is termed the "microwave BP" (designed by the procedures in §7.6.5.5 and §7.6.5.6).

7.5.8.3.1 Spectral transmittance

For the bandpasses designed in §7.5.8.1 and §7.5.8.2, Fig. 7-30 illustrates that the offband rejection of both coatings is essentially the same.

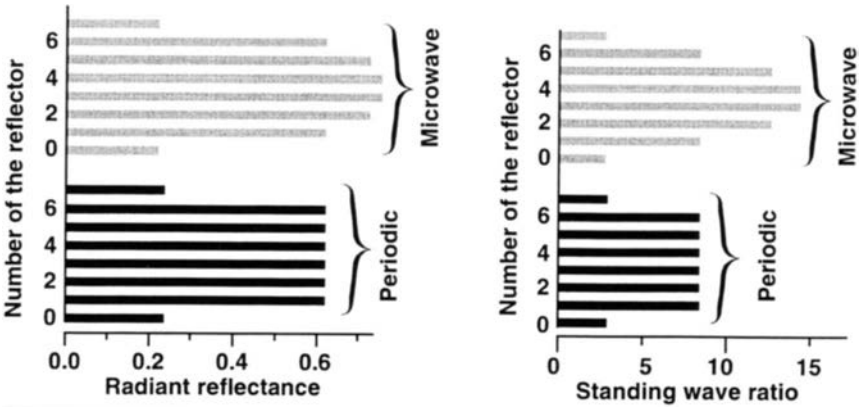


Fig. 7-31 (*Left panel*) Reflectance is plotted versus reflector number. (*Right panel*) Standing wave ratio is plotted versus reflector number.
For both of the panels above, isolation of the reflectors of the bandpasses whose designs appear in the caption to Fig. 7-30.

The passband spectral region of the "periodic BP" contains some transmittance dips and is broader. Its broader spectral bandwidth would shrink if computer optimization or other methods were used to eliminate those dips.

7.5.8.3.2 Reflectance of the isolated reflectors

Each of the designs cited in §7.5.8.3.1 contains seven cavities, which implies that there are eight reflectors. It is possible to *isolate* each of the reflectors in a given design, as explained in §7.6.5.1 and illustrated in Fig. 7-49. For each of these isolated reflectors, the left panel of Fig. 7-31 plots the radiant reflectance and the right panel graphs the standing wave ratio.

Consider first the "periodic BP." Since the basic period of **H L H L H** is repeated six times, each inner reflector consists of a three-layer sequence **L H L** surrounded by halfwave optical thickness spacers of high-index material. It is not surprising that the left panel of Fig. 7-31 shows that the reflectances of the inner reflectors are identical. This equal reflectance of all of the inner reflectors also produces the dips in the transmittance in the passband, as illustrated in Fig. 7-30.

In the case of the bandpass produced by the microwave design method, the left panel of Fig. 7-31 illustrates that the reflectances of the reflectors taper from a maximum for the reflectors at the center of the stack to a minimum for the outer reflector. It is this tapering of the reflectances that produces a maximally flat spectral shape of the passband in Fig. 7-30.

A final comment relates to the reflectance of the reflectors, as plotted in the left panel of Fig. 7-31. Since the maximum reflectance does not exceed 72%, the reflectance plot is adequate. Such a plot would be deceiving, however, if the reflectance were to approach 99.99%, as it does for a wavelength division multiplexing bandpass (see §7.8). It is for this reason that the standing wave ratio V is plotted in the right panel of Fig. 7-31. V (rather than reflectance) is used exclusively in the microwave design method.

7.6 Microwave design method

This section elaborates on a design method that is based upon the microwave prototype that evolved in the early 1960s. As compared to the periodic-structure design method that is the focus of previous sections of this chapter, the microwave design method permits the following:

- The offband rejection of a bandpass may be estimated before it is designed.
- Bandpasses with a Chebyshev spectral shape passband are possible.
- The refractive indices of the spacer layers need not be the same.

7.6.1 Attributes of a bandpass

The attributes of a conventional bandpass are surveyed in §1.6.4 and in §1.6.7 for a Wavelength Division Multiplexing (WDM) bandpass. Some of these attributes are reiterated below and in Table 7-6:

- Spectral bandwidth
- Shape, alias shape factor, alias inverse rejection ratio
- Offband rejection
- Spectral shape of the passband.

A detailed discussion of these attributes is deferred until an important design parameter — the maximum standing wave ratio (V_m) —is defined in §7.6.2.

7.6.1.1 Spectral bandwidth of the passband

Table 7-6 reiterates the two conventions for measuring the spectral bandwidth, depending whether the bandpass is *conventional* or used for WDM

Table 7-26 Optical properties of a bandpass filter

Attribute	Conventional bandpass	Bandpass for WDM
Bandwidth	Measured at 50% T level	Measured at the –0.5 dB T level
Shape factor	Usually important	Almost always important
Offband rejection	Usually important	Ignored

A comparison of the attributes of a conventional bandpass and a bandpass used for Wavelength Division Multiplexing WDM.

7.6.1.2 Shape of the passband

The shape [enumerated in §1.6.4(f)] is an inverse measure of how closely the passband spectral shape resembles a rectangle. It is determined by the number of cavities, as illustrated in Fig. 1-87.

> The greater the number of cavities, the more closely the passband transmittance resembles a rectangle.

This "rectangularness" is evident in the 17 cavity bandpass mentioned in §7.6.3.3.

7.6.1.3 Offband rejection

The offband rejection is the minimum transmittance in the offband spectral region, as identified by T_{min} in Fig. 7-32. It is quite important in a conventional bandpass that is "completely blocked." T_{min} may be translated into an absorbance D_{max} by

$$D_{max} = -\log_{10}(T_{min}) \ . \qquad (7\text{-}27b)$$

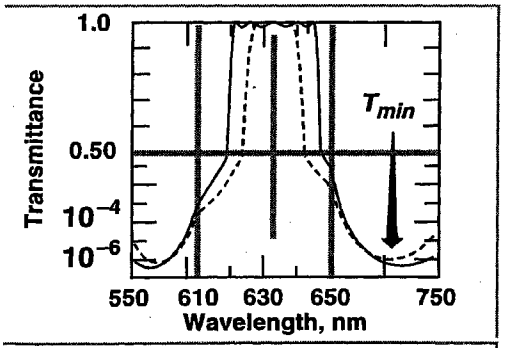


Fig. 7-32 $T(\lambda)$ of a bandpass with Chebyshev response (*solid*) and with maximally flat response (*dashed*). The designs appear in §7.10.3.6. The ordinate changes scale from linear to log at 0.5. The scale of the abscissa changes at 610 nm and 650 nm.

> When a bandpass with a given number of cavities may be characterized by the parameter V_m [maximum standing wave ratio, defined in Eq. (7-35)], its T_{min} becomes smaller as V_m increases.

7.6.2 A prototype multiple-cavity filter

The keystone of microwave bandpass design (see [Matthaei 64]) is the prototype bandpass, which is the focus of §7.6.2. It consists of halfwave lengths of waveguide of alternating high and low impedances and may, or may not, be fabricated as such because the required impedances are sometimes too large. Nevertheless, the prototype furnishes a model upon which an actual filter is constructed.

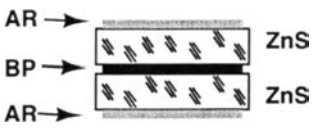


Fig. 7-33 A prototype bandpass BP is sandwiched between slabs of zinc sulfide with antireflection coatings AR on their exterior surfaces.

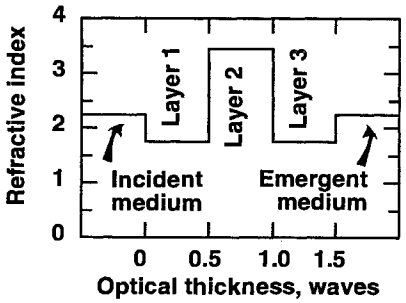


Fig. 7-34 Versus optical thickness, refractive index profile of the three-cavity prototype bandpass filter captioned in Fig. 7-35.

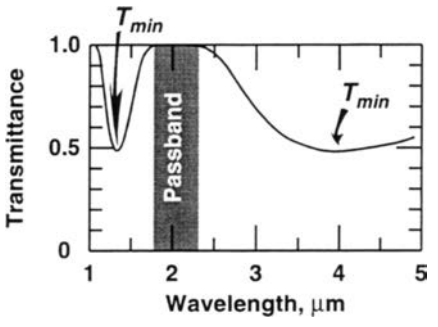


Fig. 7-35 $T(\lambda)$ of a three-cavity prototype bandpass (*dashed*)
ZnS **L H L** *ZnS*
where the media *ZnS*, **L**, and **H** have refractive indices of 2.26, 1.75, and 3.44, respectively. The optical thickness of each layer is $\lambda_0/2$ at λ_0 of 2.0 μm. The shaded region is the passband.

The passband of the prototype filter discussed here has a maximally flat spectral shape. The R/T ratio of such a filter containing q cavities is

$$R/T = F = F_{max}\ (\sin\beta)^{2q} \qquad (7\text{-}27c)$$

where

$$F_{max} \equiv (V_m - 1)^2\ (4V_m)^{-1}\ . \qquad (7\text{-}27d)$$

An approximation is valid for $V_m >> 1$

$$F_{max} \approx V_m/4\ . \qquad (7\text{-}28a)$$

V_m is the maximum standing wave ratio, which is discussed in §7.6.2.3.

7.6.2.1 Example of a prototype

The example is a prototype that is producible, although there would be little point in fabricating it because of its poor offband rejection. Figure 7-33 shows that the coating is sandwiched between two plane parallel slabs of hot pressed zinc sulfide, (*ZnS*) whose refractive index is 2.26. Antireflection coatings are deposited upon the exterior surfaces of the identical slabs. The bandpass itself is deposited upon one of the slabs and the other slab is optically contacted to it, as suggested by Li [99]. The result is that both the incident medium and substrate manifest a refractive index of 2.26.

Figure 7-34 graphs the refractive index profile of the three-layer prototype coating. The optical thickness of each layer is $\lambda_0/2$ (at λ_0 of 2.0 μm) and its design is

ZnS **L H L** *ZnS* .

The caption to Fig. 7-35 lists the refractive index of the outer layers as 1.75 (corresponding to SiO deposited with generous amounts of oxygen) and the center high-index layer is silicon of refractive index 3.44.

The transmittance graph in Figure 7-35 shows that the passband occurs in the spectral region centered at λ_0 = 2.0 μm. The dispersions of the refractive indices are ignored in the calculation. The minimum offband transmittance (designated as T_{min} in

Fig. 7-35) is slightly below 50% and occurs at one-half the wavelength of the passband center and also at double the wavelength of the passband center.

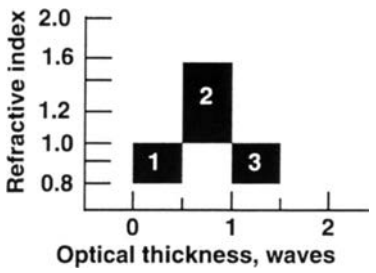


Fig. 7-36 The refractive index profile captioned in Fig. 7-34 is normalized so that the refractive index of the incident medium is 1.00.

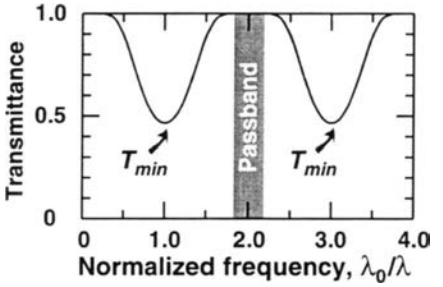


Fig. 7-37 The caption to Fig. 7-35 obtains, with the exception that the abscissa is normalized frequency and the optical thickness of each layer is $\lambda_0/2$ at a normalized frequency of 2.00.

The reflectors in the prototype are produced by the discontinuity in the refractive indices (see Fig. 7-34) at each of the four interfaces—the Fresnel reflections (see §2.2.3). This index discontinuity is evident in Fig. 7-34, which plots the index vs optical thickness for the bandpass.

7.6.2.1.1 Normalized refractive index vs optical thickness

The prototype design of an actual bandpass filter uses an index of 1.00 (i.e., air) as the incident and emergent medium. §2.4.3.4.1 mentions that the refractive index of each layer of a nonabsorbing stack may be divided by a constant and, yet, the radiant transmittance is unaltered, provided the optical thickness of each layer is unchanged. When each refractive index of the **L H L** design captioned in Fig. 7-35 is divided by 2.26 (the refractive index of the ZnS), the incident and emergent media of the normalized design manifest a refractive index of 1.00. The resulting plot of refractive index versus optical thickness appears in Fig. 7-36. Although this coating is hypothetical, it nevertheless introduces the concept of the prototype design in which both the incident medium and emergent medium are air.

7.6.2.1.2 Plot of transmittance vs normalized frequency

The plot of transmittance vs wavelength in Fig. 7-35 compresses the short wavelength spectral region and elongates the longer wavelengths. As shown in Fig. 7-37, a plot of transmittance vs normalized frequency is symmetrical about the center frequency of the passband, which is $\lambda_0/\lambda = 2.00$

How is this normalization accomplished? At a normalized frequency $\lambda_0/\lambda = 1.0$, the optical thicknesses of each layer in the prototype is a quarterwave, that is, $\lambda_0/4$. In this scheme, the lowest frequency passband of the bandpass occurs at a normalized frequency $\lambda_0/\lambda = 2.0$. The lowest frequency transmittance minimum (T_{min} in Fig. 7-37) appears at a normalized frequency $\lambda_0/\lambda = 1.0$. The definition of λ_0/λ differs from that Fig. 7-20 and other Figs. where the periodic-structure design method is used to produce the bandpass.

7.6.2.1.3 Spectral shape of the passband — analysis with vectors

It is evident in Fig. 7-37 that the spectral shape of the maximally flat passband is "square topped." The vector addition of amplitudes offers insight that helps us to understand the reason this happens. Figure 7-38 depicts the vector addition of amplitudes analysis of the three-layer prototype cited in the caption to Fig. 7-35. There are four vectors that represent the amplitude reflection from each of the four interfaces. The final column in Table 7-7 shows that the Fresnel amplitude reflection coefficients at each interface alternate in sign.

At the passband center wavelength of 2.00 μm (see Fig. 7-35), each layer has a halfwave optical thickness, which corresponds to a phase thickness β of 180° and a double phase thickness 2β of 360°. The left panel of Fig. 7-38 shows that the four vectors are collinear and that they cancel each other. At the frequency of the passband center, the net reflectance is zero, which corresponds to 100 percent transmittance.

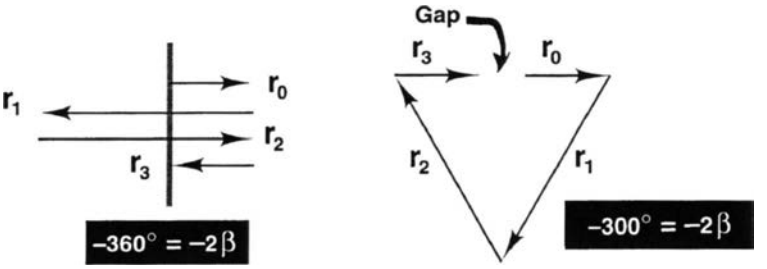


Fig. 7-38 Vector addition of amplitudes to compute the net amplitude reflection coefficient of the bandpass whose design is captioned in Fig. 7-35. (*Left panel*) at a wavelength of 2.00 μm (passband center) (*Right panel*) at a wavelength of 2.40 μm . *Gap* is the length of the vector that represents the net amplitude reflection coefficient, at this wavelength.

The right panel in Fig. 7-38 graphs the vectors at the offband wavelength of 2.4 μm. Since the phase thickness is 180° at a wavelength of 2.00 μm, at a wavelength of 2.4 μm the phase retardance β is

$$\beta = 180° \times 2.00 / 2.40 = 150° \qquad (7\text{-}28b)$$

and the "double phase thickness 2β" is 300°. The vectors are positioned in such a manner that the tip of the last vector (which is r_3) is quite close to the origin. The "Gap" in Fig. 7-38 indicates the length of the vector that represents the magnitude of the net amplitude reflection coefficient, which is 0.088. The radiant reflectance $R = (0.088)^2 = 0.008$, which corresponds to a transmittance of 99.2%.

The behavior of the vectors mentioned in the previous paragraph occurs only if the vector lengths are correct. The refractive indices of the prototype mentioned in the caption to Fig. 7-35 are chosen such that the spectral transmittance of the passband manifests a maximally flat shape. This assures that the passband shape is maximally flat, which, in turn, produces the proper Fresnel coefficients at each interface.

7.6.2.2 Standing wave ratio at each of the interfaces

The generic design of a three-cavity bandpass is $\mathbf{R_0 S R_1 S R_1 S R_0}$ where **R** is a reflector and **S** is a layer of halfwave optical thickness. Table 7-7 shows the construction of the maximally flat bandpass whose index profile appears in Fig. 7-34.

It is evident that there are three cavities—layers of halfwave optical thickness. But, where are the reflectors? The answer is that the reflectances of $\mathbf{R}_0$ and $\mathbf{R}_1$ are provided by the index mismatch between the layers—the Fresnel reflections. At each interface, the Fresnel amplitude reflection coefficients r_i and standing wave ratios V_i are listed.

Table 7-7 Maximally flat three-cavity prototype bandpass

Medium or layer	Refractive index	At i^{th} interface, V_i	At i^{th} interface, Fresnel reflection
Incident medium	$2.26 = n_0$		
		$1.29 = V_0$	$+0.127 = r_0$
Layer 1	$1.75 = n_1$		
		$1.97 = V_1$	$-0.326 = r_1$
Layer 2	$3.44 = n_2$		
		$1.97 = V_2$	$+0.326 = r_2$
Layer 3	$1.75 = n_3$		
		$1.29 = V_3$	$-0.127 = r_3$
Emergent medium	$2.26 = n_s$		
		$V_m = 6.45$, the product	

Table 7-7 shows that the indices of the incident medium and outer layer are $n_0 = 2.26$ and $n_1 = 1.75$, respectively. At that interface, the Fresnel amplitude reflection coefficient r_0 is

$$r_0 = \frac{n_0 - n_1}{n_0 + n_1} = \frac{2.26 - 1.76}{2.26 + 1.76} = 0.127 \qquad (7\text{-}29)$$

and the standing wave ratio (see §2.3.2) V_0 is

$$V_0 \equiv n_0 / n_1 \equiv 2.26 / 1.75 = 1.29 \ . \qquad (7\text{-}30)$$

Table 7-7 cites the indices of the outer layer and inner layer as $n_1 = 1.75$ and $n_2 = 3.44$, respectively. At the interface between these layers, the Fresnel amplitude reflection coefficient r_1 is

$$r_1 = \frac{n_1 - n_2}{n_1 + n_2} = \frac{1.75 - 3.44}{1.75 + 3.44} = -0.326 \ . \qquad (7\text{-}31)$$

If V_1 is the standing wave ratio (see §2.3.2) at that interface, since $n_2 > n_1$

$$V_1 \equiv n_2 / n_1 \equiv 3.44 / 1.75 = 1.97 \ . \qquad (7\text{-}32)$$

Microwave designers usually specify the properties of a reflector in terms of its standing wave ratio, whereas in optical physics (see Seeley [61]) the Fresnel reflection coefficient was used. It is clear that given V_1, r_1 can be found—and vice versa. Where possible, this book uses standing wave ratio in the design procedures.

7.6.2.3 Maximum standing wave ratio

The maximum standing wave ratio V_m is an important parameter because it influences both the offband rejection and the spectral bandwidth of a bandpass. By definition, it is the product of the standing wave ratios at the interfaces of each cavity of the bandpass.

An example is the prototype in Table 7-7. There are three layers and, hence, four interfaces at which Fresnel reflections take place. The standing wave ratios at these interfaces are numbered from zero: V_0, V_1, V_2 and V_3. V_1 (which also equals V_2) appears in Eq. (7-32). V_0 (which, due to symmetry equals V_3) is found in Eq. (7-30).

By definition, the maximum standing wave ratio V_m is the product of the standing wave ratio V at each of the four interfaces

$$V_m \equiv V_0\ V_1\ V_2\ V_3 = 1.29 \times 1.97 \times 1.97 \times 1.29 = 6.45 \quad . \tag{7-33}$$

The minimum transmittance T_{min} in Fig. 7-35 is related to V_m. For a nonabsorbing system, Eq. (2-19f) converts V_m into T_{min} as

$$T_{min} = 4\ (V_m + V_m^{-1} + 2)^{-1} = 4\ (6.45 + 6.45^{-1} + 2)^{-1} = 0.47 \tag{7-34}$$

which is arrowed in Figs. 7-35 and 7-37. If the bandpass were redesigned with a larger V_m, it would manifest a narrower spectral bandwidth and increased offband rejection.

Table 7-8 Design of a five-cavity prototype bandpass

Medium or layer	Refractive index	At i^{th} interface, V_i	At i^{th} interface, Fresnel reflection
Incident medium	$1.00 = n_0$	$3.14 = V_0$	$-0.517 = r_0$
Layer 1	$3.14 = n_1$	$16.0 = V_1$	$+0.882 = r_1$
Layer 2	$0.196 = n_2$	$31.9 = V_2$	$-0.939 = r_2$
Layer 3	$6.256 = n_3$	$31.9 = V_3$	$+0.939 = r_3$
Layer 4	$0.196 = n_4$	$16.0 = V_4$	$-0.882 = r_3$
Layer 5	$3.14 = n_5$	$3.14 = V_5$	$+0.517 = r_3$
Emergent medium	$1.00 = n_s$		
		$V_m = 2.58\ 10^6$, $\eta = 6.41$	

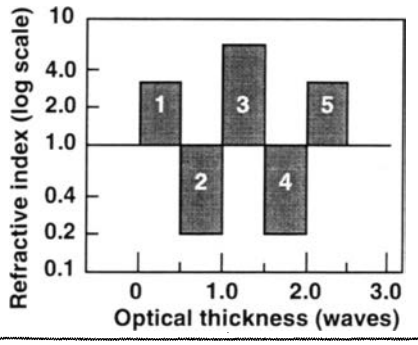


Fig. 7-39 Refractive index (plotted on a logarithmic ordinate) vs optical thickness (in waves at λ_0, where λ_0 is the passband center wavelength) of the five layer prototype filter whose $T(\lambda)$ appears in Fig. 7-40.

For the generic bandpass consisting of q cavities in which the standing wave ratio at the ith interface is V_i, the maximum standing wave ratio V_m is defined

$$V_m \equiv \prod_{i=0}^{q} V_i \quad . \tag{7-35}$$

V_m is often a large number and is better expressed in terms of its decadic logarithm η

$$\eta \equiv \log(V_m) \quad . \tag{7-36a}$$

7.6.2.4 A five-cavity prototype

Figure 7-39 depicts the refractive index profile (on a logarithmic ordinate scale) of a five-cavity prototype. How was this produced? Computer software is used, as discussed in §7.10.1.7. The spectral transmittance of the prototype and multilayer simulation of the prototype are compared in the next section.

Table 7-9 Optical properties of prototype and actual bandpass

Attribute	Prototype	Actual bandpass
T_{min}	Nearly 6 absorbance units	4 absorbance units
Frequency of T_{min}	At 0.5 of passband frequency	At 0.88 passband frequency
Passband width	6.6%	20.4%

For the prototype bandpass defined in the caption to Fig. 7-40 and the multilayer bandpass whose design appears in §7.10.3.5.1, minimum offband transmittance T_{min} , frequency at which T_{min} occurs and the passband spectral width.

For the prototype, Table 7-8 lists the standing wave ratio and amplitude reflection coefficient of each of the six reflectors, as well as the refractive indices of each layer.

7.6.2.5 Comparison of the prototype with the multilayer that simulates it

Figure 7-40 plots on a frequency scale the absorbance of a five section prototype and an actual multilayer bandpass that simulates it. The design of the latter appears as the first entry in §7.10.3.5.1. The differences in the offband rejection and the passband spectral widths are summarized in Table 7-9. These differences may be explained by several facts:

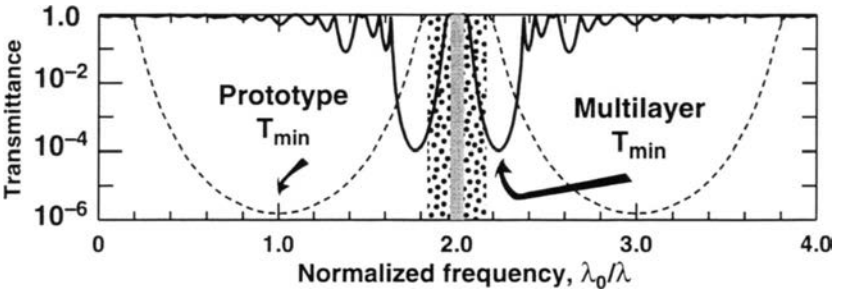


Fig. 7-40 Versus normalized frequency, transmittance (on a log scale) of the prototype bandpass(*dashed*) of the design
air **A B C B A** *air*
where the refractive indices of *air*, **A**, **B** and **C** are 1.00, 3.14, 0.196 and 6.256, respectively. The optical thickness of each layer is $\lambda_0/4$. The passband spectral region is dotted. The *solid* curve is the transmittance of the multilayer bandpass design appearing as the first entry in §7.10.3.5.1. Its passband is shaded.

(1) The reflectors of a prototype are derived from Fresnel reflections (i.e., reflections at the interfaces between the layers), which (by definition) reflect at all frequencies. These reflectors are highly reflective at a normalized frequency of 0.5—ergo, T_{min} occurs at a frequency of 0.5. In other words, the prototype is a quarterwave stack at a frequency of 1.0, where each layer is of quarterwave optical thickness.

(2) The reflectors of a multilayer bandpass are quarterwave stacks, which reflect over a limited spectral bandwidth (see §5.2.4.3). At a normalized frequency of 0.82 or lower, little (or no) reflectance is produced by these quarterwave stacks. The consequences are that T_{min} occurs at a normalized frequency of 1.78 and

that the transmittance is reasonably large at normalized frequencies less than 1.64 and at normalized frequencies in excess of 2.36.

(3) The reflectors of the prototype are derived from Fresnel reflections, which (by definition) manifest no phase dispersion (see §7.2.3). The spectral bandwidth of 20.4% of the prototype depends only upon its V_m (defined in §7.6.2.3). If the passband transmittance simulates a Chebyshev polynomial, the amount of passband ripple influences the spectral bandwidth.

(4) The reflectors of the actual bandpass are quarterwave stacks that introduce phase dispersion (see §7.2.3), which, in turn, narrows the spectral width of the passband by a factor of three.

7.6.2.5.1 Minimum offband transmittance

The minimum offband transmittance T_{min} of the prototype is at the quarterwave point, which is one-half of the frequency (i.e., double the wavelength) of the passband. This occurs because the six reflections at the interfaces between the sections of the prototype (illustrated in Fig. 7-39) are independent of frequency.

Contrast this with the actual multilayer bandpass. Inspection of its design in §7.10.3.5.1 shows that the reflectors consist of layers of quarterwave optical thickness. According to §5.2.4.3, the spectral region of the stop band of the quarterwave stack reflectors extends from a normalized frequency of 0.89 to 1.11. A consequence of this limited spectral bandwidth over which the reflectance is relatively large, is that the T_{min} of the actual bandpass is several decades higher than that of the prototype.

7.6.2.5.2 Spectral bandwidth

Paragraph (4) of §7.6.2.5 mentions that the phase dispersion of the multilayer stack produces a narrower spectral passband width, as compared to a prototype. §7.6.2.5.2 quantifies this narrowing. The starting point is the inverse fractional bandwidth Q of a single-cavity bandpass, as cited in Eq. (2-83h). The foregoing equation contains the derivative of the phase shift upon reflection $d\delta/d\sigma$ of the reflectors that surround the cavity. Let us assume that a spacer of low refractive index n_L is used. According to Eq, (5-16b), the slope of the phase shift upon reflection is

$$\frac{d\delta}{d\sigma} \approx \frac{-\pi \lambda_0 n_L}{n_H - n_L} \tag{7-36b}$$

where the material dispersion of the high-index layer is neglected. The m in Eq. (5-16b) is the order number of the stop band, which is assumed to be unity. The parameter "$2 n_c h_c$" (i.e., double the optical thickness of the spacer layer) is equal to "$m \lambda_0$ ", where λ_0 is the wavelength of the passband center and m is the order number of the spacer. Combining the foregoing variables, Eq. (2-83h) becomes

$$Q \approx \left[\pi m \lambda_0 - \frac{-\pi m \lambda_0 n_L}{n_H - n_L} \right] \sigma_0 \sqrt{R_m} (1 - R_m)^{-1} \tag{7-36c}$$

or

$$Q \approx \pi \left[m + \frac{n_L}{n_H - n_L} \right] \sqrt{R_m} (1 - R_m)^{-1} . \tag{7-36d}$$

The conclusion is that the phase dispersion narrows the passband spectral width by the *shrink factor*, which is

$$shrink\,factor = [\, m + n_L / (n_H - n_L)\,]^{-1} \qquad (7\text{-}36e)$$

where m is the order number of the spacer layer. That is, the spectral bandwidth of the multilayer equals the spectral bandwidth of the prototype multiplied by the *shrink factor*.

Although the foregoing *shrink factor* strictly applies to a single-cavity bandpass, empirical studies indicate that it obtains, as well, for a multiple-cavity bandpass. As an example, suppose that a WDM bandpass is constructed with layers of refractive indices $n_L = 1.47$ and $n_H = 2.065$. Assuming that the spacer thicknesses are first order, a numerical value for the *shrink factor* is

$$shrink\,factor = [1 + 1.47/(2.065 - 1.47)\,]^{-1} = [1 + 1.47/0.595]^{-1} = (3.47)^{-1} \qquad (7\text{-}36f)$$

which shows that the multilayer bandpass is approximately 3.5 times smaller than that of the prototype.

If second-order spacers ($m = 2$) were used, the relative spectral bandwidth of the prototype is $(2.00)^{-1}$ and the right of Eq. (7-36e) is equal to 1/4.47. The phase-dispersion narrowing *shrink factor* is then the ratio $2/4.47 = (2.24)^{-1}$. That is, the spectral bandwidth of the multilayer bandpass is 2.24 times smaller than the prototype.

Although it is not obvious, the foregoing equation obtains when a high-index spacer layer is used. This follows when the numerator and denominator of Eq. (5-16c) is multiplied by the product $n_H n_L$ and n_0 equals n_H.

Table 7-10 Comparison of bandpasses with 3, 4 and 5 cavities

Cav	η	BW	Rel	BW25	Shape	Lyrs	Tkns	Fig.	Curve
3	21.3	0.08	1.0	0.266	3.3	147	33.6	7-41	shaded
4	21.3	0.682	8.5	1.564	2.3	144	33.4	7-41	dash
5	21.3	2.96	37.1	4.78	1.61	147	34.0	7-41	solid

Comparison of bandpasses: η appears in Eq. (7-36a). *BW* is the passband full spectral width (in nm) at a transmittance level of 89.1% (i.e., –0.5 dB) and *BW25* is the width (in nm) at a transmittance level of –25 dB. *Shape* is the ratio of BW25/BW and is proportional to the *shape factor* that appears in Eq. (1-11d). *Lyrs* is the layer count. *Rel* is the relative spectral width, setting the width of the three-cavity bandpass to 1.00. *Tkns* is the metric thickness of the coating in μm. *Fig.* is the figure in which the spectral transmittance curve appears.

7.6.3 Choice of the number of cavities—introduction

The following generalizations may be made about the performance of the bandpass as the number of cavities is varied:

- For a given V_m [defined in Eq. (7-35)], the spectral bandwidth becomes broader as the number of cavities increases. This is illustrated in §7.6.3.1.
- For a given spectral bandwidth, the slope at the edges of the passband (as, for example, where the transmittance is 50%) increases linearly with the number of cavities.
- A consequence of the foregoing sentence is that a bandpass with a large number of cavities has a passband whose spectral shape is more rectangular, as compared to a bandpass with fewer cavities.

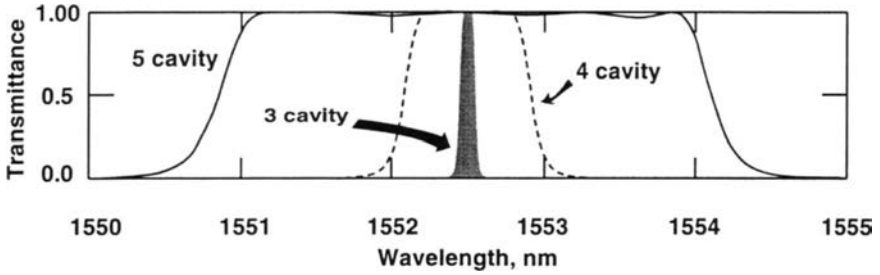


Fig. 7-41 Transmittance of bandpasses with three cavities (*shaded*), four cavities (*dashed*) and five cavities (*solid*). In each case, the η [see Eq. (7-36a)] is 21.3. The designs appear in §7.10.3.1.

7.6.3.1 Spectral bandwidth of a bandpass containing 3, 4 or 5 cavities

For a bandpass with a given maximum standing wave ratio V_m [defined in Eq. (7-35)], its spectral bandwidth broadens as the number of cavities grows larger. Figure 7-41 depicts the transmittances of bandpasses with three, four and five cavities. The V_m of each bandpass is the same, namely $\eta = 21.3$. Table 7-10 shows that the spectral width of the four-cavity bandpass is in excess of eight times that of the three cavity. The spectral width of the five cavity is 37 times greater than the spectral width of the three cavity.

7.6.3.2 Spectral slope influenced by the number of cavities

The first task is to define the spectral slope. It is approximated by the first derivative of the transmittance T with respect to wavelength

$$slope \approx \frac{dT}{d\lambda} \ . \tag{7-37}$$

Table 7-11 Spectral properties of coarse WDM bandpasses

Cav	η	BW	BW25	Shape	Lyrs	Tkns	Order	Fig.
9	14.4	12.94	16.7	1.23	176	43.9	2	7-43 solid
17	47.58	12.18	12.94	1.062	331	97.2	3	7-43 shaded

Comparison of bandpasses with the number of cavities shown in the Cav column. η appears in Eq. (7-36a). *BW* is the passband full spectral width (in nm) at a transmittance level of 89.1% (i.e., –0.5 dB) and *BW25* is the width (in nm) at a transmittance level of –25 dB. *Shape* is the shape factor appearing in Eq. (1-11d). *Lyrs* is the layer count. *Tkns* is the metric thickness of the coating in μm. *Order* is the largest order number of any of the spacers. *Fig.* is the figure in which the spectral transmittance curve appears.

The following "rule of thumb" is useful:

> For a given spectral bandwidth, the slope of the spectral transmittance at the passband edge is proportional to the number of cavities.

Figure 7-41 shows that the slope of the three cavity is greater than that of the five cavity. Since the spectral bandwidths are different, however, this in no way invalidates the aforementioned "rule of thumb."

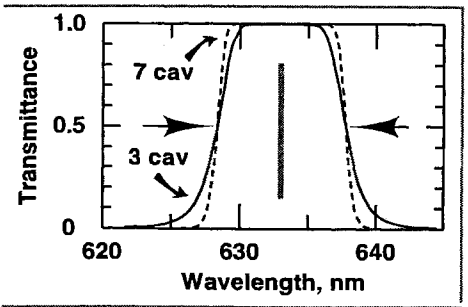


Fig. 7-42 Spectral transmittance of a three-cavity (*solid*) and a seven-cavity (*dashed*) bandpass. The designs appear in §7.10.3.7.1 and 7.10.3.7.2.

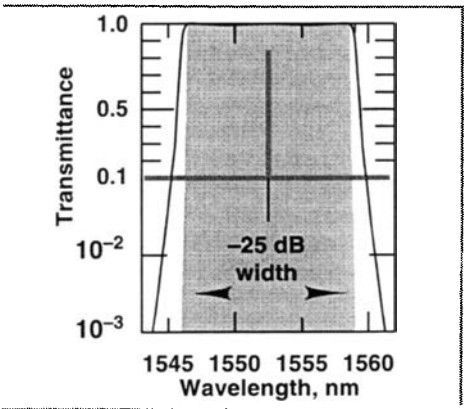


Fig. 7-43 Spectral transmittance of coarse WDM bandpasses containing nine cavities (*solid*) and seventeen cavities (*shaded*). The designs appear in §7.10.3.8. The ordinate scale changes from linear to log at 0.10.

Figure 7-42 depicts the transmittances of a three-cavity bandpass and a seven-cavity bandpass whose spectral bandwidths (measured at the 50% transmittance level) are nearly identical. It is quite evident that at the 50% level, the slope of the seven-cavity bandpass is considerably larger than the slope of the three-cavity bandpass. This, in turn, causes the passband spectral shape to be more rectangular. In other words, the shape factor (see the comment in the next paragraph) is small. Table 7-10 lists the shape factor for the three-cavity, four-cavity and five-cavity systems. The shape factor is 3.3 for the three cavity and decreases to 1.61 for the five cavity.

7.6.3.3 The coarse WDM bandpass—broad passband width

Unlike the conventional WDM bandpass that typically has a spectral bandwidth comparable to a nanometer, the coarse WDM manifests a bandwidth (measured at the –0.5 dB transmittance level) that ranges from 3 nm to 13 nm. Since the spectral bandwidth grows as the cavity count increases, a nine-cavity bandpass and a seventeen-cavity bandpass are choices for this coarse WDM. Figure 7-43 depicts the transmittances. Table 7-11 lists their "minus one-half dB" spectral bandwidths as 16 nm and 13 nm, respectively. The shape factor for the seventeen-cavity bandpass is an amazingly small value of 1.062. This ultra-rectangular transmissive spectral shape of the seventeen cavity may be attributed to the immense slope, as mentioned in §7.6.3.2.

7.6.4 Control of the spectral bandwidth of a multiple-cavity bandpass

§7.2 illustrates that there are three methods of controlling the spectral bandwidth of a single-cavity bandpass:

- adjusting the "reflectance of the reflectors" contiguous to the spacer;
- changing the order number of the spacers (see §7.6.4.2);
- altering the phase dispersion of the reflectors.

Each of these methods is applied in §7.6.4.2 to alter the spectral bandwidths of a three-cavity bandpass and a seven-cavity bandpass. In each case, the passband is centered at 633 nm. The bandpass is deposited on glass and covered with an optical cement of the same refractive index.

Table 7-12 Properties of a single-cavity bandpass

Design of bandpass	Reflector design	R	V_0	V_m	Ref.
H L H 2A H L H	*air* **H L H** *glass*	63%	8.81	77.6	§7.2
(H L)^2 H 2A (H L)^2 H	*air* **(H L)^2 H** *glass*	84%	22.12	492	§7.2.2

For the isolated reflectors listed in the *Reflector design* column and at a wavelength λ_0, the standing wave ratio V_0 as well as its reflectance R are found. The meanings of the design symbols appear in the caption to Fig. 7-9. V_m is the square of V_0. *Ref* is the section in which the single-cavity bandpass is discussed.

7.6.4.1 Re-examination of the reflectors of the Fabry-Perot

The purpose of §7.6.4.1 is to illustrate that:

> The adjustment of the maximum standing wave ratio V_m (defined in §7.6.2.3) is equivalent to altering the reflectance of the reflectors of the bandpass.

In Table 7-2, the first entry cites a bandpass of the design **H L H 2A H L H**. The first step is to find the standing wave ratio of the three-layer reflector **H L H**. *Air* is the incident medium and *glass* (index of refraction 1.51) is the emergent medium. From Eq. (2-48c), the standing wave ratio is

$$V_0 = \frac{n_H^4}{n_L^2 n_s n_0} = \frac{(2.30)^4}{(1.45)^2 (1.51) (1.0003)} = 8.81 \ . \qquad (7\text{-}38)$$

Since there are two identical reflectors, the maximum standing wave ratio of the bandpass is the square of the above, $V_m = (8.81)^2 = 77.7$, which appears in Table 7-12. It is evident in this Table that the standing wave ratio V_0 increases as more layers are added to the reflector. That is, any increase in V_m decreases the spectral bandwidth.

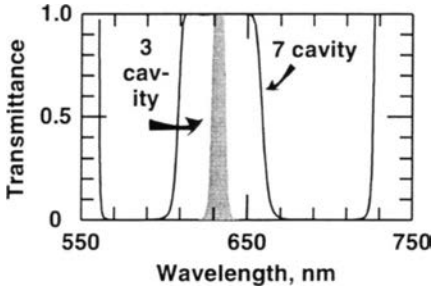


Fig. 7-44 Spectral transmittances of a bandpass with three cavities and a bandpass with seven cavities. The designs appear in §7.10.3.7.1 and §7.10.3.7.3.

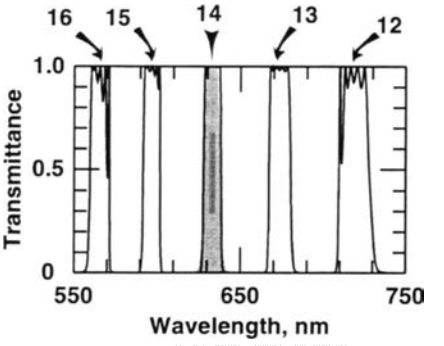


Fig. 7-45 Spectral transmittance of a seven-cavity bandpass with spacer layers of order 14. The order numbers of the passbands are arrowed. The design is in §7.10.3.7.4.

Table 7-13 Bandpasses with three cavities and seven cavities

Cav	η	BW	Lyrs	Tkns	Order	Fig.	Comment
3	7.16	9.44	55	5.5	1	7-44	See also Fig. 7-42
7	7.0	50.37	59	6.4	1	7-44	Broad spectral bandwidth
7	15.9	9.41	123	12.7	1	7-42	Phase-dispersion narrowing used
7	7.0	10.43	59	24.8	14	7-45	High order spacer layers used
7	15.9	10.48	123	12.4	1	--	No phase-dispersion narrowing

Comparison of bandpasses: *Cav* is the number of cavities. η appears in Eq. (7-36a). *BW* is the passband full spectral width (in nm) at a transmittance level of 0.5 (i.e., –3 dB). *Lyrs* is the layer count. *Tkns* is the metric thickness of the coating in μm. *Order* is the largest order number of any of the spacers. *Fig.* is the figure in which a spectral transmittance curve appears.

7.6.4.2 Bandpasses with identical V_m values and unequal spectral bandwidths

The curves in Fig. 7-41 (see §7.6.3.1) illustrate that the spectral bandwidth increases as the cavity count progresses from three to five. In this section, a three-cavity bandpass and a seven-cavity bandpass are compared. In the first two designs, the same V_m is used for each bandpass and the order number of the spacers is the same. Figure 7-44 illustrates that the spectral bandwidth of the seven-cavity bandpass is at least five times as large as the bandwidth of the three-cavity bandpass.

7.6.4.3 Bandpasses with equal spectral bandwidth — adjustment of V_m

The goal is to make equal the spectral bandwidths of a three-cavity bandpass and a seven-cavity bandpass. §7.6.4 lists three methods of shrinking the spectral bandwidth of the seven cavity. Its spectral bandwidth is decreased by increasing the $\log(V_m)$ [defined in Eq. (7-36a)] from 7.0 to 15.9. As mentioned in §7.6.4.5, phase-dispersion narrowing may be introduced to adjust (ever so slightly!!) the spectral bandwidth.

7.6.4.4 Use of high-order spacers

The next approach is not a practical one and is included only for pedagogical reasons. The approach is to modify the design of the seven-cavity bandpass whose transmittance curve appears in Fig. 7-44 by increasing the thickness of each spacer. The spacer order number (see §7.2.1) is increased from 1 to 14. The V_m of this design remains at 7. The transmittance of the resultant bandpass appears in Fig. 7-45. Adjacent to the passband where the order number is 14 are spurious passbands of order 13 and 15. If this coating were shipped to a customer, it would be necessary to block these spurious (and unwanted) passbands by separate coatings.

7.6.4.5 The use of phase dispersion to narrow the spectral bandwidth

Consider a trial version of the seven-cavity bandpass whose transmittance curve is similar to the one appearing in Fig. 7-42. This coating contains layers of only halfwave and quarterwave optical thickness. In the diagram below, the seven spacer layers are sequentially numbered:

$$\underset{\downarrow}{1}\quad\underset{\downarrow}{2}\quad\underset{\downarrow}{3}\quad\underset{\downarrow}{4}\quad\underset{\downarrow}{5}\quad\underset{\downarrow}{6}\quad\underset{\downarrow}{7}$$

$$(\mathrm{H\,L})^3 2\mathrm{H\,L}\,(\mathrm{H\,L})^7 2\mathrm{H}\,(\mathrm{L\,H})^8\,\mathrm{L\,H}\,2\mathrm{L}\,(\mathrm{H\,L})^9\mathrm{H}\,2\mathrm{L}\,(\mathrm{H\,L})^9\mathrm{H}\,2\mathrm{L}\,(\mathrm{H\,L})^9 2\mathrm{H\,L}\,(\mathrm{H\,L})^7 2\mathrm{H}\,(\mathrm{L\,H})^3 .$$

The problem is that the spectral bandwidth (measured at the 50% transmission level) of this seven cavity is too broad. The final entry in Table 7-13 lists its spectral bandwidth as 10.48 nm, as compared to the 9.44 nm spectral bandwidth of the three cavity. The conclusion is that the spectral bandwidth of the seven cavity needs to be narrowed slightly. This is accomplished by altering only two of the quarterwave **H** layers in the bandpass. As depicted below (by the arrows), the **H** layers that are adjacent to spacers number 3 and number 5 are increased in thickness to 3**H**— a three quarterwave:

$$\qquad\qquad\qquad\underset{\downarrow}{\qquad}\qquad\qquad\qquad\underset{\downarrow}{\qquad}$$

$$(\mathrm{H\,L})^3 2\mathrm{H\,L}\,(\mathrm{H\,L})^7 2\mathrm{H}\,(\mathrm{L\,H})^8\,\mathrm{L}\,3\mathrm{H}\,2\mathrm{L}\,(\mathrm{H\,L})^9\mathrm{H}\,2\mathrm{L}\,(\mathrm{H\,L})^9 3\mathrm{H}\,2\mathrm{L}\,(\mathrm{H\,L})^9 2\mathrm{H\,L}\,(\mathrm{H\,L})^7 2\mathrm{H}\,(\mathrm{L\,H})^3 .$$

As illustrated in §7.2.3, the use of three quarterwaves adjacent to the spacer layers produces phase-dispersion narrowing of the spectral bandwidth. The third entry in Table 7-13 shows that this phase dispersion reduces the spectral bandwidth to 9.41 nm, as compared to the 9.44 nm bandwidth of the three cavity.

7.6.4.6 Comparison of the designs

Table 7-13 and Fig. 7-46 compare the layer sequences of the three-cavity bandpass and the seven-cavity bandpass. The following may be concluded:

> The layer count (and hence, the optical thickness) of a bandpass scales as the log(V_m).

As an example, Table 7-13 shows that the three-cavity bandpass and broadband seven-cavity bandpass have values of η [i.e., log(V_m)] that are nearly the same. The layer counts in those bandpasses of 55 and 59, respectively, are nearly the same.

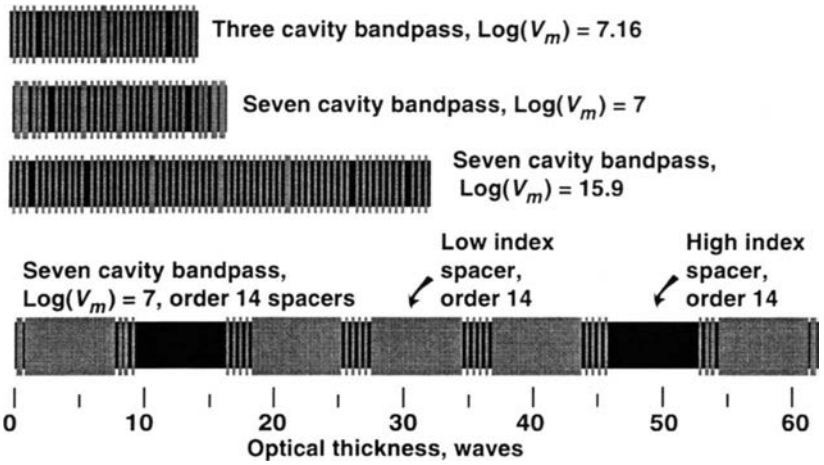


Fig. 7-46 Versus optical thickness (in waves, at 633 nm), designs of the three-cavity and seven-cavity bandpasses whose optical properties are summarized in Table 7-13. The low-index films are lightly shaded and the high-index films are black.

Table 7-14 Comparison of bandpasses containing five cavities

Lyrs	BW, nm	BW %	Ripple %	D_{max}	η	V_0	V_1	V_2
44	41.8	6.6	0.218	4.05	6.41	3.14	16.0	31.9
55	29	4.6	0.28	5.1	7.91	4.42	31.9	63.8
64	20	3.2	0.36	6.25	9.41	6.28	63.8	127
75	14	2.2	0.348	7.4	10.91	8.81	127	255

Bandwidths BW of the bandpasses whose transmittance curves appear in Fig. 7-47 *Lyrs* is the number of layers. *Ripple* is the ripple in the transmittance in the passband, in percent. η appears in Eq. (7-36a) and D_{max} is the largest offband absorbance [see Eq. (7- 27b)]. V_0, V_1 and V_2 are the standing wave ratios of the outer, intermediate and center reflectors, respectively.

The immense optical thickness of the order 14 spacers makes it evident that this is not a practical solution to the problem.

In §7.6.4.7 a series of five-cavity bandpasses is designed in which the layer count increases from 44 to 75. With increasing layer count, T_{min} decreases from 4 absorbance units to 7.40.

7.6.4.7 Control of the spectral bandwidth by the adjustment of V_m

Figure 7-47 depicts four bandpasses in which the η [see Eq. (7-36a)] progresses from 6.4 to 10.91. Table 7-14 shows that the bandwidth shrinks from 6.6% to 2.2%. The offband rejection, depicted as the arrowed T_{min} in Fig. 7-47, decreases by a factor of one thousand as the V_m progresses from smaller to larger values.

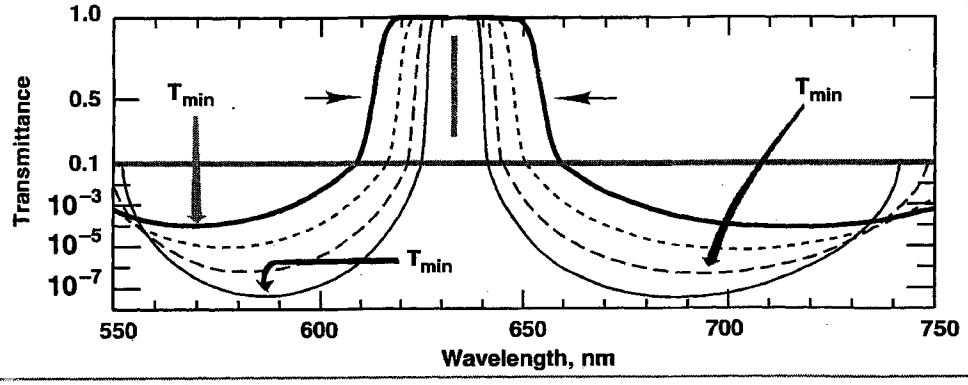


Fig. 7-47 $T(\lambda)$ of five-cavity bandpasses with small amounts of ripple in their passbands. The designs appear in §7.10.3.5. T_{min} is the minimum offband transmittance. The scale of the ordinate changes from linear to log at 0.10.

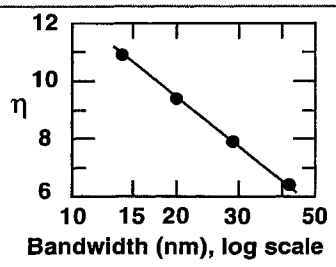


Fig. 7-48 For the five-cavity bandpass cited in Table 7-14, sine(fractional bandwidth) is plotted vs η on the ordinate.

According to Eq. (7-48b) that is developed in an appendix (§7.9.8.1.1), a plot of η on the ordinate plotted vs "log sine (bandwidth)" is linear. This is evident in Fig. 7-48, which shows such a plot, based upon the data in Table 7-14. The sine may be approximated by its argument when much less than one.

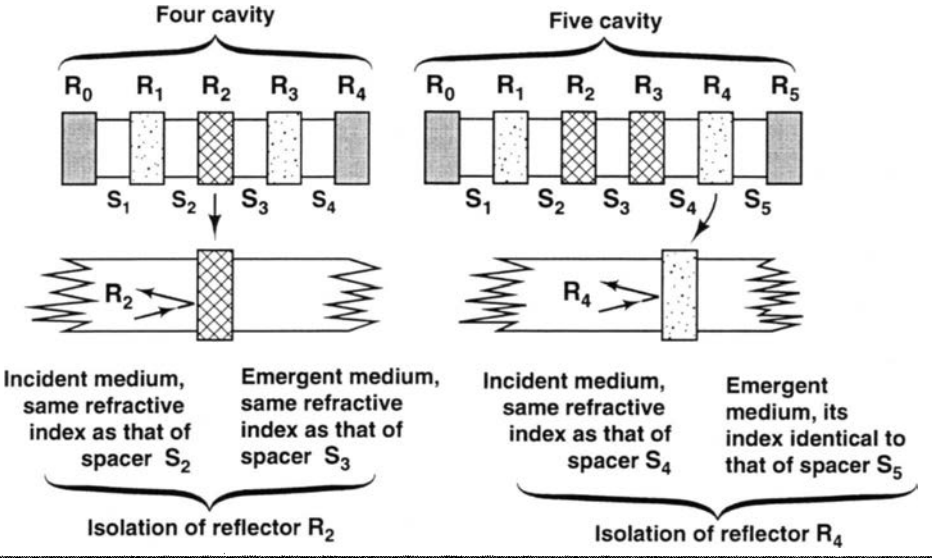


Fig. 7-49 (*Left upper*) Configuration of the reflectors R and spacers S in a four-cavity bandpass. (*Left lower*) In the four cavity, reflector R_2 is *isolated.* (*Right upper*) The bandpass contains five cavities. (*Right lower*) In the five cavity, reflector R_4 is *isolated.* Reflectors with identical shading manifest the same standing wave ratios.

7.6.5 Design procedure

§7.6.5.1 introduces the concept of an isolated reflector and the following section discusses the symmetry of the reflectors. §7.6.5.3 outlines some constrains that may be imposed upon the designs of the reflectors, as manifested in the choice of the layer thicknesses and refractive indices. §7.6.5.6 outlines a procedure for the design of a bandpass.

7.6.5.1 Isolation of a reflector

In §7.6.5.2, reference is made to the standing wave ratio of an individual reflector. How does one compute this standing wave ratio of a reflector? An example is the reflector R_2 in Fig. 7-49. The first step is to *isolate* that reflector. The isolation is accomplished by

discarding the remainder of the bandpass and retaining only the spacers on the incident and emergent sides of R_2. The spacer on its emergent side is transformed into a *pseudo-substrate* that is assumed to extend to infinity. The spacer on its incident side functions as a *pseudo-incident-medium.* The radiant reflectance R_2 of that reflector is measured inside of that incident medium, as depicted in Fig. 7-49. Equation (2-19c) converts this reflectance into a standing wave ratio.

When the actual incident or actual substrate is adjacent to the reflector, then it is used for the incident medium or substrate.

7.6.5.2 Symmetry of the standing wave ratios

There is a symmetry of the standing wave ratio V_i of the ith reflector of the Chebyshev and maximally flat designs. For a bandpass in which the number of cavities q is odd

$$V_0 = V_q \ , \qquad V_1 = V_{q-1} \ , \qquad V_2 = V_{q-2} \ , \qquad V_3 = V_{q-3} \tag{7-39}$$

and so on. The same relationships hold for an even number of cavities, with the exception that the center reflector (whose standing wave ratio is $V_{q/2}$) is unique.

An example is a five-cavity bandpass (see Fig. 7-50) containing six reflectors. The six standing wave ratios of the reflectors are constrained as follows:

$$V_0 = V_5 \qquad V_1 = V_4 \qquad V_2 = V_3 \, . \tag{7-40}$$

The right panel of Fig. 7-50 depicts a four-cavity bandpass containing five reflectors. Symmetry requires that the standing wave ratios are constrained as

$$V_0 = V_4 \qquad V_1 = V_3 \tag{7-41}$$

and V_2 is unique.

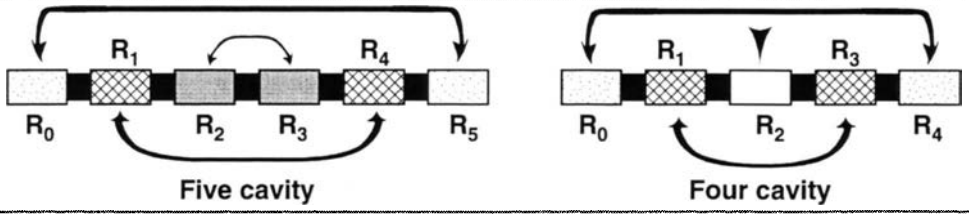


Fig. 7-50 (*Left*) Symmetry (shown by the arrows) of the reflectors R and spacers (*solid black*) in a five-cavity bandpass. (*Right*) The symmetry is shown for a four-cavity bandpass, with the exception that reflector R_2 (*vertical arrow*) is unique.

7.6.5.3 Constraints imposed upon bandpass design

The important question is: What are the constraints that may be placed upon the refractive indices and layer thicknesses? Table 7-15 lists some possibilities.

7.6.5.4 Constraints on the bandpass design—spacer layers

Suppose that a five-cavity bandpass is fabricated using two coating materials, as represented by **L** and **H**. Each layer has a quarterwave optical thickness. If $\mathbf{R_0}$, $\mathbf{R_1}$ and $\mathbf{R_2}$ denote the reflectors, possible configurations are:

(1) $\mathbf{R_0}$ LL $\mathbf{R_1}$ LL $\mathbf{R_2}$ LL $\mathbf{R_2}$ LL $\mathbf{R_1}$ LL $\mathbf{R_0}$ Homologous cavity;

(2) $\mathbf{R_0}$ HH $\mathbf{R_1}$ HH $\mathbf{R_2}$ HH $\mathbf{R_2}$ HH $\mathbf{R_1}$ HH $\mathbf{R_0}$ Homologous cavity.

That is, the refractive indices of the spacers are identical. If this constraint were relaxed, then other possibilities include:

(3)	R_0 LL R_1 HH R_2 LL R_2 HH R_1 LL R_0	Alternate cavity;
(4)	R_0 HH R_1 LL R_2 HH R_2 LL R_1 HH R_0	Alternate cavity;
(5)	R_0 HH R_1 LL R_2 LL R_2 LL R_1 HH R_0	Mixed cavity;
(6)	R_0 LL R_1 LL R_2 HH R_2 LL R_1 LL R_0	Mixed cavity.

Table 7-15 Design constraints imposed upon a bandpass

Design constraint	Where is an example?
(1) Two (and only two) coating materials are used.	Two cavity, §7.7.2.2 Six cavity, §7.7.7 Five-cavity WDM, §7.8.4.1
(2) Layers whose optical thicknesses are quarterwaves (or multiples thereof) are used exclusively.	Caption, Fig. 7-66 Five cavity, §7.8.4.3
(3) Equivalent layers may be used on the exterior of the bandpass.	Three cavity, §7.7.4.1 Four cavity, §7.8.3.2
(4) Equivalent layers may be used anywhere.	Seven cavity, §7.7.8
(5) More than two refractive indices may be used in the bandpass.	Caption, Fig. 7-80 Bandpass, §7.5.2 Caption, Fig. 7-80
(6) All spacers are of the same refractive index.	All bandpasses in Table 7-20
(7) The refractive indices of the spacer layers need not be the same.	Four cavity, §7.7.5.1 Three cavity, §7.7.4.2

Constraints that may be imposed upon the layer thicknesses and refractive indices of a bandpass filter.

The advantage of using all H spacers is that the metric thicknesses of the layers are smaller (as compared to using L spacers) and that the angle shift of the passband central wavelength (see §1.6.4.1) is smaller. Low-index layers, on the other hand, usually have a smaller amount of absorption, as compared to H spacers. A disadvantage of using the alternate-cavity or mixed-cavity configuration is that the temperature coefficient of change of the refractive index is different for the L and H cavities, the next effect of which is to detune the bandpass. §7.7.5.4 discusses the performance of the alternate cavity when the angle of incidence of the light is as large as 20°.

7.6.5.5 Design of the prototype

Given all of the bandpass specifications cited in §7.6.1 (spectral bandwidth, shape factor, offband rejection) the first step is to design a prototype. The software mentioned in §7.10.1.7 or the maximally flat equations cited in §7.10.1.8 may be used.

7.6.5.6 Conversion of the prototype into an actual design

The following procedure is used to translate a prototype bandpass into an optical multilayer bandpass:

(1) Replace each *halfwave thickness section* of the prototype with a spacer layer of halfwave optical thickness.

(2) Determine the standing wave ratio at each interface of the prototype.
(3) Design reflectors (usually composed of quarterwave optical thickness layers) that manifest the standing wave ratios cited in step (2) above.
(4) Sandwich the spacer layers [step (1) above] between the reflectors [step (3) above].

Table 7-16 Standing wave ratios of inner reflectors

q	V_a	V_b	V_c	q	V_a	V_b	V_c	q	V_b
1	1.41	1.40	1.48	16	255	230	546	30	26797
2	2.00	1.97	2.20	17	360	323	809	31	37643
3	2.83	2.77	3.26	18	509	454	2639	32	52879
4	4.00	3.89	4.83	19	720	637	3913	33	74283
5	5.65	5.47	7.17	20	1018	895	5802	34	104350
6	7.99	7.68	10.6	21	1439	1258	8602	35	146587
7	11.3	10.8	15.8	22	2035	1767		36	205920
8	16.0	15.2	23.4	23	2877	2482		37	289268
9	22.6	21.3	34.6	24	4067	3487		38	406353
10	31.9	29.9	51.4	25	5750	4899		39	570829
11	45.1	42.0	76.2	26	8129	6881		40	801879
12	63.8	59.1	113	27		9667		41	1.12645e+06
13	90.2	83.0	167	28		13579		42	1.58239e+06
14	127	117	248	29		19076		43	2.22289e+06
15	180	164	368						

Column V_a contains $(2.05/1.45)^q$; V_b contains $(2.065/1.47)^q$; V_c contains $(2.15/1.45)^q$.

7.6.5.6.1 The bandpass contains two materials

The refractive indices alternate between two materials, with known indices of n_H and n_L.

An *outer reflector* is contiguous to either the substrate or incident medium. The *inner reflectors* are buried in the stack. The first layer and the final layer of an inner reflector abut a spacer layer. Assume that each layer manifests an optical thickness of $\lambda_0/4$. The V (at $\lambda = \lambda_0$) of any inner reflector containing q layers is

$$V = \left[\frac{n_H}{n_L}\right]^{q+1} . \tag{7-42}$$

For an inner reflector, Table 7-16 is useful in determining the number of layers needed to achieve a given standing wave ratio.

Example: An inner reflector consists of eleven layers [$q = 11$] of indices of 1.45 and 2.30. From Eq. (7-42) its standing wave ratio is

$$V = n_H^{12}\, n_L^{-12} = (2.30)^{12}\ (1.45)^{-12} = 254 . \tag{7-43a}$$

Table 7-16 lists the standing wave ratios produced by multilayers of quarterwave optical thickness layers alternating between silica (index 1.45) and films of higher refractive index.

7.6.5.6.2 Inclusion of equivalent layers

The incorporation of equivalent layers allows the designer to achieve virtually any standing wave ratio. An example is an outer reflector (adjacent to an incident medium of air) that should achieve a V of 11.2. The **L** and **H** layers indices are 1.45 and 2.30, respectively.

The index of the *low-index-pseudo-substrate* is identical to that of the spacer layer of index 1.45. After the principle of isolation (§7.6.5.1) is applied, a trial reflector design is

air **H L H** *low-index-pseudo-substrate* .

Equation (2-48c) provides the reduced admittance, which, according to Eq. (2-20b), is identical to the standing wave ratio

$$V = n_H^4 n_L^{-3} (1.0003)^{-1} = (2.30)^4 (1.45)^{-3} (1.0003)^{-1} = 9.2 \ , \qquad (7\text{-}43b)$$

which falls short of the *desideratum* of 11.2. The 1.0003 index of refraction of air is included for pedagogical reasons. The standing wave ratio of the design

air **H** $(\mathbf{L\,H})^2$ *low-index-pseudo-substrate*

is

$$V = n_H^6 n_L^{-5} (1.0003)^{-1} = (2.30)^6 (1.45)^{-5} (1.0003)^{-1} = 23.1 \ , \qquad (7\text{-}43c)$$

which is too large. A solution to the problem is to use a quarterwave optical thickness layer **E** of intermediate refractive index, as in the design

air **E** $(\mathbf{L\,H})^2$ *low-index-pseudo-substrate* .

The standing wave ratio is found from Eq. (2-48c)

$$11.2 = n_E^2 n_L^{-5} n_H^4 (1.0003)^{-1} \ . \qquad (7\text{-}43d)$$

Given that n_H and n_L are known, solving the foregoing for n_E provides 1.60. If **E** is to be replaced by an equivalent layer, the procedure in §2.11.3 is used to synthesize it.

7.7 Examples of conventional bandpass design

Conventional in the title above intends that the bandpasses are not used in wavelength division multiplexing WDM systems. The spectral bandwidth refers to the full spectral width at the 50% transmission level. The maximum offband rejection is typically in the range of 4 to 6 absorbance units.

The microwave system is used to design bandpasses with varying numbers of cavities. When the cavity count is two or three, it makes little difference whether the microwave or periodic structure (see §7.3) method is used. The results are quite similar.

7.7.1 Offband rejection

This parameter, as illustrated by T_{min} in Fig. 7-32, is often important in the design of a conventional bandpass. Determine the minimum allowed maximum standing wave ratio V_m [defined in Eq. (7-35)] from

$$V_m \approx 200\, T_{min}^{-1} \ . \qquad (7\text{-}43e)$$

That is, the V_m used in the design should exceed the foregoing value.

Example: If T_{min} in the offband region is to be less than 10^{-4}, then the V_m should exceed $2{\times}10^6$.

7.7.2 Two-cavity designs

This coating is used principally when a narrow spectral bandwidth is needed. The bandpasses in this section have a maximally flat passband shape. Coating materials of indices 1.45 and 2.30 are used in §7.7.2.

7.7.2.1 Homologous-cavity designs with two materials

The standing wave ratios of the inner and outer reflectors are provided by Eqs. (7-57e) and (7-57f). The design procedure is simplified by noting that the standing wave ratio of the inner reflector is the *square* of the outer reflector. As explained below, there are two basic designs if a *homologous-cavity* bandpass is used. These designs are written for a bandpass containing low-index spacers. The substrate index is *glass* of refractive index 1.52 and *air* is the incident medium.

In the remainder of §7.7.2, the massive media *low-index-pseudo-incident-medium* and *low-index-pseudo-substrate* manifest a refractive index identical to that of a low-index spacer, namely 1.45.

Table 7-17 Optical properties of two-cavity filters

Design					Lyrs	BW	η
F	LL	HLH	LL	H	7	22.0	1.6
H	LL	HLHLH	LL	HLE	11	11.4	2.4
FLH	LL	HLHLHLH	LL	HLH	15	6.2	3.2
HLH	LL	HLHLHLHLH	LL	HLHLE	19	3.7	4.0
FLHLH	LL	HLHLHLHLHLH	LL	HLHLH	23	2.2	4.8
HLHLH	LL	HLHLHLHLHLHLH	LL	HLHLHLE	27	1.4	5.6
FLHLHLH	LL	HLHLHLHLHLHLHLH	LL	HLHLHLH	31	0.85	6.4

Designs and bandwidths BW (in %) of two-cavity filters in which the incident medium is air and the substrate index is 1.52. **L**, **E**, **F** and **H** represent layers of optical thickness $\lambda_0/4$ and refractive indices of 1.45, 1.87, 1.91 and 2.30, respectively. η is in Eq. (7-36a).

7.7.2.1.1 Inner reflector contains three layers

Consider a bandpass with three layers in its center reflector. After that reflector is isolated (see §7.6.5.1), its design is

low-index-pseudo-incident-medium **H L H** *low-index-pseudo-substrate*

which, according to Eq. (7-42), manifests a V_1 of 6.33.

The next task is to design the two outer reflectors. From Eq. (7-57e) it is found that the standing wave ratio of each should be $\sqrt{V_1} = \sqrt{6.33} = 2.52$. For the reflector next to the substrate, a single **H** layer produces a V_2 of 2.40, which is close enough to the *desideratum* of 2.52. That is, after isolation (see §7.6.5.1), the outer reflector adjacent to the substrate is

low-index-pseudo-incident-medium **H** *glass* .

The final task is to design the outer reflector on the *air* side. The addition of a single equivalent layer **F** of index 1.91 (and 90° phase thickness) on the exterior of the bandpass produces the requisite standing wave ratio.

The completed bandpass design is

air **F LL H L H LL H** *glass*

and contains either seven layers (if an actual material is used for **F**) or eight layers (if an equivalent layer is used for **F**). Table 7-17 (first entry) lists its spectral bandwidth* as 22%.

7.7.2.1.2 Inner reflector contains five layers

Using the principle of isolation (see §7.6.5.1), the design of the inner reflector is

low-index-pseudo-incident-medium **H L H L H** *low-index-pseudo-substrate* ,

which, according to Eq. (7-42), manifests a standing wave ratio V_1 of 15.9.

From Eq. (7-57e), the standing wave ratios V_2 and V_0 of the outer reflectors should attain a value of

$$V_2 = V_0 = \sqrt{V_1} = \sqrt{15.9} = 3.99 \ . \qquad (7\text{-}44a)$$

The next task is to design the reflector that is adjacent to the substrate. A calculation similar to that of §7.7.6.1 shows that a reflector of the design

low-index-pseudo-incident-medium **H L E** *glass*

(where **E** is an equivalent layer of index 1.87 and 90° phase thickness) meets the requirement of Eq. (7-44a). For the outer reflector adjacent to air, a single **H** layer produces a V_0 of 3.73, which is close enough to the target value of 3.99.

The design of the final bandpass is

air **H L L H L H L H L L H L E** *glass* ,

which contains eleven layers if an actual **E** layer is used and twelve layers if **E** is simulated with an equivalent layer. The second entry in Table 7-17 lists its spectral bandwidth as in excess of 11%.

7.7.2.1.3 Narrowing of the spectral bandwidth

Both the spectral bandwidth and offband rejection of the bandpasses in §7.7.2.1.1 or §7.7.2.1.2 may be altered by the following procedure: (1) Add an **L H** pair to each of the outer reflectors; and (2) augment the inner reflector with four layers **L H L H**. The result is that the standing wave ratio of each outer reflector is multiplied by

$$V = n_H^2 n_L^{-2} = 2.42 \ . \qquad (7\text{-}44b)$$

Because twice as many layers are added to the inner reflector, its standing wave ratio is increased by a factor of $(2.42)^2$, thus preserving the relationships in Eqs. (7-57e) and (7-57f). Table 7-17 illustrates the passband narrowing that occurs each time eight layers (four **L H** pairs) are added to these basic designs. Each increment of eight layers increases η by 1.6. The bandpass containing 31 layers manifests a spectral bandwidth of less than 1%.

7.7.2.2 Alternate-cavity design with two materials

The goal is to produce a bandpass with a maximally flat passband shape and spectral bandwidth of 1%. A possible solution is to use an alternate-cavity bandpass of the design

air **E L H H L H L H L L H L F** *glass*

where the indices of the equivalent layers **E** and **F** (of 90° phase thickness at λ_0) are yet to be determined. The isolation of the center reflector is shown in §7.6.5.1.

From Eq. (7-42), the standing wave ratio of the center reflector is $n_H^5 n_L^{-5}$.

Equation (7-57e) requires that

---

* The bandwidth is the full spectral width at the 50% transmission level.

$$n_H^5 n_L^{-5} = \left[n_E^2 n_L^{-2} n_H \right]^2 \qquad (7\text{-}44c)$$

for the reflector adjacent to air. This leads to an index for the **E** layer of

$$n_E = n_H^{3/4} n_L^{-1/4} = (2.30)^{3/4} \times (1.45)^{-1/4} = 1.70 \ . \qquad (7\text{-}44d)$$

A similar calculation for the reflector adjacent to the substrate produces

$$n_F = n_H^{1/4} n_L^{1/4} n_s^{1/2} = (2.30)^{1/4} \times (1.45)^{1/4} \times (1.52)^{-1/2}) = 1.67 \ , \qquad (7\text{-}44e)$$

which is close enough to 1.70 that an index of 1.70 is used.

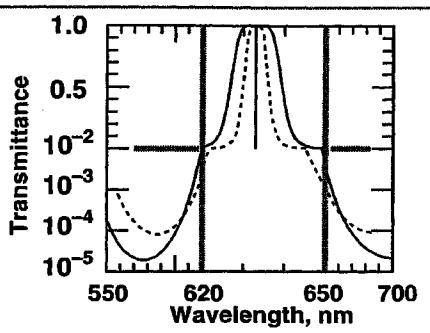


Fig. 7-51 $T(\lambda)$ of a three-cavity bandpass with alternate refractive index spacers (*solid*)

air **H L H** 4**L** (**H L**)^5 4**H** (**L H**)^5 4**L** (**H L**)^2 0.35**L** 0.276**H** 0.35**L** *glass*

and two-cavity bandpass(*dashed*)

air 0.38**L** 0.22**H** 1.38**L** (**H L**)^2 6**H** (**L H**)^6 6**L** (**H L**)^3 0.38**L** 0.22**H** 1.38**L** *glass*

where the refractive indices of *glass*, **L** and **H** are 1.52, 1.45 and 2.30, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 633 nm. The ordinate scale changes from linear to logarithmic at 0.01. The abscissa changes scale at 620 nm and 650 nm.

The spectral bandwidth of the foregoing bandpass is shrunk by the procedure of §7.7.2.1. A single **L H** pair is added to each of the outer reflectors and a double **L H** pair is appended to the inner reflector. The resultant bandpass contains 27 layers

air **E L** (**H L**)^2 **HH** (**L H**)^6 **LL** (**H L**)^3 **E** *glass*

and manifests a spectral bandwidth of 1.7%. Its spectral bandwidth is decreased to 1% by using third-order spacers. When the **E** and **F** layers are replaced by equivalent layers, the resulting design is

air 0.38**L** 0.22**H** 1.38**L** (**H L**)^2 6**H** (**L H**)^6 6**L** (**H L**)^3 0.38**L** 0.22**H** 1.38**L** *glass* .

Its transmittance is the dashed curve in Fig. 7-51.

7.7.3 Discarding of halfwaves

A layer with halfwave optical thickness is absentee and may be removed from the bandpass. §2.6.3.1 discusses the mathematical basis for this process. At a wavelength λ_0, the bandpass in the final paragraph of §7.7.2.1.2 is analyzed in this manner. Because both **LL** halfwave thickness spacers may be removed, the bandpass reduces to

air **H** ... **H L H L H H L E** *glass* .

The two quarterwave thickness underlined **H** layers coalesce into a halfwave thickness layer. It is removed and the bandpass design is

air **H H L H L L E** *glass* .

The contiguous **L** layers of quarterwave optical thickness and adjacent underlined **H** layers above fuse into halfwave thickness layers—they are removed from the bandpass. There are three remaining of quarterwave thickness layers

air **L H E** *glass* .

The coatings manifests a reduced admittance of

$$y = n_L^2 n_H^{-2} n_E^2 n_s^{-1} = (1.45)^2 \ (2.30)^{-2} \ (1.87)^2 \ (1.52)^{-1} = 0.914 \ , \qquad (7\text{-}45a)$$

which produces a radiant reflectance R of

$$R = (1 - 0.914)^2 \ (1 + 0.914)^{-2} = 0.2\% \qquad (7\text{-}45b)$$

which could be reduced to zero by increasing slightly the refractive index of the **E** layer.

7.7.4 Three-cavity designs

The objective is a bandpass with a 2% BW.* The passband spectral shape is maximally flat. Refractive indices of 1.45 and 2.30 are used. The incident medium is *air* and the *glass* substrate index is 1.52.

7.7.4.1 A bandpass containing homologous cavities

As mentioned in §7.6.5.4, homologous means that all spacer layers have the same refractive index. The first step is to estimate the spectral bandwidth produced by a given V_m of 10^7 by assuming a maximally flat spectral shape. Remembering that the F_{max} in Eq. (7-58a) is $V_m/4$ and that the division of the left of Eq. (7-58a) by 1.571 produces the fractional bandwidth, the substitution of $q = 3$ shows that the prototype spectral bandwidth is 5.5%. There is hope that phase-dispersion narrowing will shrink this 5.5% to the requisite 2%. The final entry in Table 7-9 illustrates that the phase dispersion of the multilayer bandpass narrows its spectral bandwidth (as compared to the prototype bandpass) by a factor of three. Low-index spacers are used.

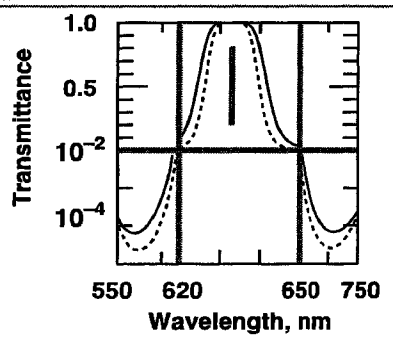


Fig. 7-52 $T(\lambda)$ of a three-cavity bandpass with homologous spacers (*dashed*)

air 0.427**L** 0.132**H** 0.427**L** (**L H**)^2 **LL** (**H L**)^5 **H LL** (**H L**)^5 **H LL** (**H L**)^2 0.258**L** 0.452**H** 0.258**L** *glass*

and a bandpass with spacers of alternating refractive index (*solid*)

air **H L H LL** (**H L**)^5 **HH** (**L H**)^5 **LL** (**H L**)^2 0.35**L** 0.276**H** 0.35**L** *glass* .

The meaning of *glass*, **L** and **H** appears in the caption to Fig. 7-51. The ordinate scale changes from linear to logarithmic at 0.01. The abscissa changes scale at 620 nm and 650 nm.

Equation (7-43a) shows that an eleven-layer reflector has a V of 254. The foregoing standing wave ratio is obtained when $\eta = 6.91$ is substituted into Eq. (7-57d), using the data of Table 7-25. After the principle of isolation is used (see §7.6.5.1), each of the two center reflectors has the design

low index **HL HLHLHL HL H** *low index*

where the *low index* represents either the massive *pseudo-incident medium* or *pseudo-substrate* of index 1.45 — namely, that of the low-index spacer.

The next step is to design the outer reflectors. Using data from Table 7-25 and Eq. (7-57d), the required standing wave ratio V_0 is

$$V_0 = 0.803 \ 10^{0.1661 \times 6.91} = 11.2 \ . \qquad (7\text{-}46a)$$

An analysis similar to that of §7.6.5.6.2 indicates that an equivalent layer should be used for the reflectors on both the incident and emergent sides of the bandpass. The bandpass design is

air **E** (**LH**)^2 **LL** (**HL**)^5 **HLL** (**HL**)^5 **HLL** (**HL**)^2 **F** *glass*

* The BW is the full spectral width at the 50% transmission level.

where **E** and **F** are equivalent layers of 90° phase thickness at λ_0. §7.6.5.6.2 [i.e., Eq. (7-43d)] shows that the index of **E** is 1.60. A similar calculation shows that the **F** layer index is 1.97. The fractional bandwidth of the foregoing design is quite close to 2%. Its transmittance appears in Fig. 7-52.

7.7.4.2 A bandpass containing alternate cavites

It appears that the attainment of the requisite spectral bandwidth of the foregoing coating is a stroke of luck. What happens when this is not the case? An alternate-cavity bandpass serves as another example. "Alternate cavity" means that the refractive index of the spacers alternates between a high index and low index.

The first step is to isolate (see §7.6.5.1) the ten layer inner reflector

high-index-pseudo-incident-medium $(\mathbf{L}\,\mathbf{H})^5$ *low-index-pseudo-substrate*

where the refractive indices of the *high-index-pseudo-incident-medium* and *low-index-pseudo-substrate* are equal to the index of an **H** layer and **L** layer, respectively. From Eq. (7-42) its standing wave ratio is $V_1 = 160$.

The design of the outer reflectors is the next task. Equation (7-57d) shows that V_1 is 160 when η [see Eq. (7-36a)] is 6.31. Given η, Eq. (7-57d) produces a V_0 of 9.0. After isolation of the outer reflector adjacent to air, its design is

air **H L H** *low-index-pseudo-substrate*

and its standing wave ratio is

$$V_0 = n_H^4\, n_L^{-3}\, n_0^{-1} = (2.30)^4\ (1.45)^{-3}\ (1.0003)^{-1} = 9.2\ , \qquad (7\text{-}46b)$$

which is close enough to the *desideratum* of 9.0. The 1.0003 (the refractive index of air) is retained for pedagogical reasons only.

A procedure similar to that of §7.6.5.6.2 is followed to design the reflector contiguous to the substrate. After isolation (see §7.6.5.1) of that reflector, the design is

low-index-pseudo-incident-medium $(\mathbf{H}\,\mathbf{L})^2$ **E** *glass*

where **E** is an equivalent layer of quarterwave optical thickness. The standing wave ratio V_2 of the reflector is

$$V_2 = 9.0 = n_H^4\, n_E^2\, n_L^{-5}\, n_s^{-1} = (2.30)^4\ (1.45)^{-5}\ (1.52)^{-1}\ n_E^2\ . \qquad (7\text{-}46c)$$

Solution of the foregoing equation results in an index of 1.77 for the **E** layer.

After the final bandpass is assembled, its design is

air **H L HLL** $(\mathbf{H}\,\mathbf{L})^5$ **HH** $(\mathbf{L}\,\mathbf{H})^5$ **LL** $(\mathbf{H}\,\mathbf{L})^2$ **0.35L 0.276H 0.35L** *glass* .

The three-layer sequence **0.35L 0.276H 0.35L** above represents an equivalent layer of index 1.77, which was obtained by following the procedure in §2.11.3. A thin film design software package such as TFCalc [Software Spectra 02] may also be used. As shown in Fig. 7-52, this bandpass manifests a spectral bandwidth of 2.7%, which is in excess of our target of 2%. The conversion of the spacers to second order (see §7.2.1) narrows the spectral bandwidth. That is, the halfwave optical thickness layer **LL** is replaced with the fullwave layer **LLLL**. In a like manner, the fullwave **HHHH** is substituted for halfwave **HH**. Figure 7-51 illustrates that its spectral bandwidth is quite close to 2%.

A three-cavity bandpass for an optical fiber communication system is developed in §7.8.2. As compared with the fractional bandwidth of 2×10^{-2} in the bandpass developed above, the bandpass in the later section manifests a fractional bandwidth of 7×10^{-4}.

7.7.5 Four-cavity bandpasses

The bandpasses designed in this section are:

(1) Homologous-cavity bandpass with high-index spacers; and
(2) Bandpass with alternating refractive index spacers.

The performance at nonnormal incidence is examined. In addition the performance of the bandpass is examined when it transmits a cone of light—the biconical transmittance is computed. In each instance, the bandpass is to transmit 633 nm. Materials of 1.45 (for the low-index layers) and 2.05 (for the high-index layers) are used. The substrate is soda lime glass of index 1.52.

7.7.5.1 Four-cavity bandpasses—the refractive indices of the spacers alternate

The code described in §7.10.1.7 was executed and it showed that an η of 6.6345 and passband ripple of 0.03% required standing wave ratios for the reflectors of 4.85, 45.12 and 90.2 for V_0, V_1 and V_2, respectively. Table 7-16 shows that the center interior reflector and intermediate interior reflector should contain 12 layers and 10 layers, respectively. These are cited in Table 7-18 as entries 1 and 2.

The next task is to design the outer reflector adjacent to air. The spacer layer closest to air is of low refractive index. The three-layer sequence

air **E L H** *low index*

(where **E** is an equivalent layer of 90° phase thickness) should manifest a standing wave ratio of 4.85. According to Eq. (2-48c), this requires that

$$V_0 = n_H^2 n_L^{-3} n_E^2 = (2.05)^2 (1.45)^{-3} n_E^2 = 4.85 \ . \qquad (7\text{-}46d)$$

The result is that n_E is 1.867. This appears as the fourth entry in Table 7-19.

Table 7-18 Designs of interior reflectors

Entry	Cavities	Incident	Design of reflector	Emergent	SWR
1	4	*low*	$(\mathbf{H\,L})^5$	*high*	45.1
2	4	*high*	$(\mathbf{L\,H})^6$	*low*	90.2
3	4, 5	*low*	$(\mathbf{H\,L})^5\,\mathbf{H}$	*low*	63.8
4	4, 5	*low*	$(\mathbf{H\,L})^6\,\mathbf{H}$	*low*	127
5	4	*high*	$(\mathbf{L\,H})^5\,\mathbf{L}$	*high*	63.8
6	4	*high*	$(\mathbf{L\,H})^6\,\mathbf{L}$	*high*	127
7	5	*low*	$(\mathbf{H\,L})^3\,\mathbf{H}$	*low*	16
8	5	*low*	$(\mathbf{H\,L})^4\,\mathbf{H}$	*low*	31.9
9	5	*low*	$(\mathbf{H\,L})^7\,\mathbf{H}$	*low*	255

After isolation (see §7.6.5.1), standing wave ratios SWR of inner reflectors for bandpasses specified in the *Cavities* column. The optical thickness of layers **H** and **L** is $\lambda_0/4$. **H**, *high*, **L** and *low* represent refractive indices of 2.05, 2.05, 1.45 and 1.45, respectively. *Incident* is the incident medium and *Emergent* is the emergent medium.

Since the refractive index of the spacers alternates between low and high, the spacer layer nearest the substrate is a high refractive index material. The outer reflector adjacent to the substrate is designed using a procedure similar to the above. The result appears as entry 5 in Table 7-19.

7.7.5.1.1 Wavelength centering of the bandpass — initial guess

With the objective that a cone of light (at 633 nm) will be transmitted through the bandpass, the central wavelength is not 633 nm but rather a longer wavelength of 640 nm. This shift of the passband center to longer wavelengths is discussed in more detail in §1.6.6.

Table 7-19 Outer reflectors of bandpasses

Section	Cavities	Design of reflector	Simulation	Phase, E	Index, E	SWR
7.7.5.2	4	*air* **E L H L** *high*	0.327**L** 0.329**H** 0.327**L**	90°	1.725	5.8
7.7.5.2	4	*high* **L H L E** *glass*	0.535**H** 1.933**L** 0.535**H**	270°	2.127	5.8
	4	*low* **H L H E H** *glass*	0.294**L** 0.394**H** 0.294**L**	90°	1.777	5.8
7.7.5.1	4	*air* **E L H** *low*	0.233**L** 0.515**H** 0.233**L**	90°	1.867	4.85
7.7.5.1	4	*high* **L H L E** *glass*	0.172**L** 0.639**H** 0.172**L**	90°	1.945	4.85
7.7.6.1	5	*air* **E** *low*	0.542**H** 1.92**L** 0.542**H**	270°	2.143	3.14
7.7.6.1	5	*low* **H L E** *glass*	0.238**L** 0.505**H** 0.238**L**	90°	1.86	3.14
7.7.6.2	5	*air* **E L H** *low*	0.285**L** 0.412**H** 0.285**L**	90°	1.791	4.42
7.7.6.2	5	*low* **H L E** *glass*	0.526**H** 1.951**L** 0.526**H**	270°	2.106	4.42
7.7.6.2	5	*air* **E L H** *low*	0.538**H** 1.927**L** 0.538**H**	270°	2.134	6.28
7.7.6.2	5	*low* **H L H L E** *glass*	0.237**L** 0.508**H** 0.237**L**	90°	1.862	6.28
7.7.6.2	5	*air* **E L H L H** *low*	0.287**L** 0.408**H** 0.287**L**	90°	1.788	8.81
7.7.6.2	5	*low* **H L H L E L** *glass*	0.551**H** 1.903**L** 0.551**H**	270°	2.163	8.81

For the four-cavity and five-cavity bandpasses developed in §7.7.5.1, §7.7.5.2 and §7.7.6.1, designs and optical properties of the outer reflectors after each of those reflectors is isolated (see §7.6.5.1). **H**, **L** and **E** represent layers of optical thickness $\lambda_0/4$. The refractive indices of **H**, **L** and *glass* are 2.05, 1.45 and 1.52, respectively. *Simulation* is the equivalent layer representation (see §2.6.7) of the **E** layer, given its index and phase thickness. SWR is the standing wave ratio (at a wavelength λ_0) of the reflector. Depending upon the context, *low* may refer to either *low-index-pseudo-incident-medium* (index equals 1.45) or *low-index-pseudo-substrate* (its index is 1.45). Depending upon the design, *high* may refer to *high-index-pseudo-substrate* (index equals 2.05) or *high-index-pseudo-incident-medium* (of index 2.05).

7.7.5.2 A bandpass with high-index spacers

The equations embedded in the computer code cited in §7.10.1.7 were solved. This shows that an η of 7.2404 and passband ripple of 0.037% require for the reflectors standing wave ratios of 5.8, 63.9 and 127 for V_0, V_1 and V_2, respectively. Table 7-16 shows that the intermediate interior reflector and center interior reflector should contain 11 layers and 13 layers, respectively. These appear as entries 5 and 6 in Table 7-18.

The designs of the outer reflectors are similar to those in §7.7.5.1, with one exception. The equivalent layer adjacent to the glass substrate should attain a value of 2.127. Since this exceeds 2.05 of the high-index layer, §2.6.7.5.6 points out that a layer with 90° phase thickness may not be used. It is for that reason that this layer has a phase

thickness of 270°. The designs of each of the outer reflectors appear in Table 7-19. The transmittance of the final design appears in Fig. 7-54.

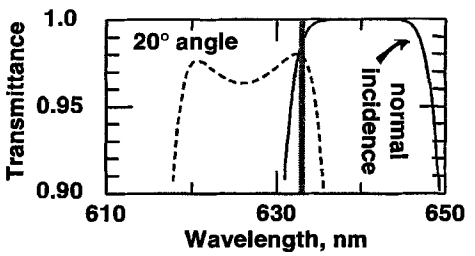


Fig. 7-53 $T(\lambda)$ (for randomly polarized light) of a four-cavity Chebyshev bandpass at normal incidence (*solid*) and at 20° incidence angle in air (*dashed*). The design appears in §7.10.3.4.1. The shaded vertical line is 633 nm.

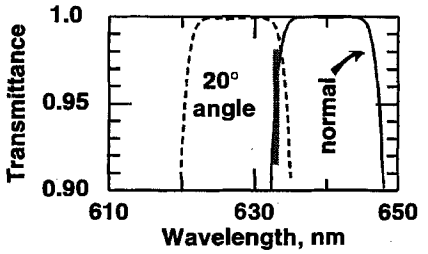


Fig. 7-54 $T(\lambda)$ (for randomly polarized light) of a four-cavity Chebyshev bandpass at normal incidence (*solid*) and at 20° incidence angle in air (*dashed*). The design appears in §7.10.3.4.2. The shaded vertical line is 633 nm.

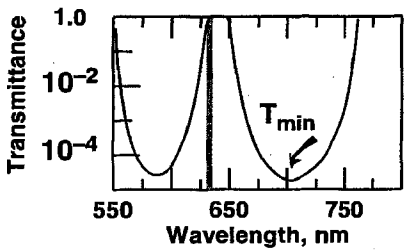


Fig. 7-55 $T(\lambda)$ (on an absorbance scale) of a four-cavity Chebyshev bandpass at normal incidence. The design appears in §7.10.3.4.2. The shaded vertical line is 633 nm. T_{min} is the maximum offband rejection.

7.7.5.3 Offband rejection

Figure 7-55 graphs the transmittance on an absorbance scale. The maximum offband rejection T_{min} is below 10^{-4}.

7.7.5.4 Performance at nonnormal incidence

Figures 7-53 and 7-54 present, respectively, the performance at nonnormal incidence of the alternate-cavity bandpass (§7.7.5.1) and homologous-cavity bandpass (§7.7.5.2). In essence, at 20° incidence angle the effective thicknesses (see §2.5.2) of the low-index cavities and high-index cavities are different. This, in turn, produces a detuning of the cavities that degrades the passband shape, as is evident in Fig. 7-53. At an incidence angle of 20°, there is a bothersome dip in the passband transmittance of the *alternate-cavity* bandpass that may be attributed to the detuning mentioned above.

7.7.5.5 Biconical transmittance

Imagine that a cone of light is incident upon the bandpass, as shown in Fig. 1-17. Equation (10-11b) is used to find the transmittance for a 20° (semi-angle) cone. These results appear in Fig. 7-56.

7.7.5.5.1 Wavelength centering

It is evident from Fig. 7-56 that the initial guess of 640 nm for the wavelength centering of the bandpass should be adjusted a few nm to longer wavelengths, in order that the peak biconical transmittance be centered at 633 nm. It is also evident that if the cone angle were increased above the present 20°, the peak biconical transmittance would begin to decrease rapidly.

7.7.5.6 Additional four-cavity bandpass designs

A four-cavity bandpass for an optical fiber communication system is developed in §7.8.3. Its fractional bandwidth is at least 100 times smaller than the relatively broad bandpasses developed earlier in §7.7.5.

7.7.6 Five-cavity bandpasses — specifications

Indices of 1.45 and 2.05 are to be used for the layers. The substrate is soda-lime glass. The goal is to include layers of only quarterwave optical thickness (or multiples of quarterwaves) in the interior of the bandpass. Low-index spacers are used.

Figure 7-47 plots the transmittances of a family of five-cavity bandpasses. For each filter, Table 7-14 lists the η [see Eq. (7-36a)], passband ripple and the standing wave ratio of each isolated reflector. The η and passband ripple were obtained by using the code in §7.10.1.7.

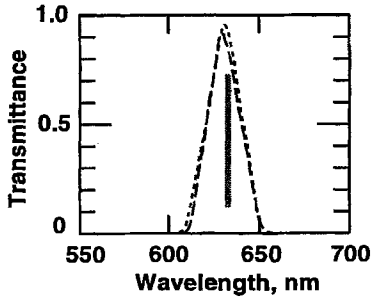


Fig. 7-56 Biconical transmittance of a four-cavity bandpass whose design appears in §7.10.3.4.2. The illumination is Lambertian and the semi-angle of the cone is 20°. The shaded vertical line is 633 nm.

7.7.6.1 Five-cavity bandpasses — the bandpass with the smallest V_m

The target standing wave ratios of 16.0 and 31.9 for the intermediate and center reflectors, respectively, appear as the first entry in Table 7-14. Entries 7 and 8 in Table 7-18 show that quarterwave stack reflectors with seven and nine layers manifest the requisite standing wave ratios.

Following the procedure of §7.6.5.1, the next step is to isolate the outer reflector next to air. The standing wave ratio V_0 of this reflector is 3.14. Since a single **H** layer, as in the design

air **H** *low index* ,

manifests a standing wave ratio of 2.9, this hints that a slight increase in the layer refractive index will boost the standing wave ratio to the required 3.14. The sixth line in Table 7-19 lists the index of the layer and the equivalent layer (of 270° phase thickness) that simulates it.

The design of the outer reflector next to the substrate is

low index **H L E** *glass*

where the pseudo-incident-medium *low index* has the same refractive index as a low-index spacer. The refractive index of the **E** layer is obtained by solving the following equation

$$3.14 = n_H^2 n_E^2 n_s^{-1} n_L^{-3} . \qquad (7\text{-}46e)$$

Table 7-19 shows that the index of the **E** layer is 1.86. After assembly of the final filter, its design appears in §7.10.3.5.1 and the heavy curve in Fig. 7-47 graphs its transmittance.

7.7.6.2 Three additional designs of five-cavity bandpasses

Table 7-18 lists the standing wave ratios of the reflectors in five-cavity bandpasses in which the η is larger than the one in §7.7.6.1. The η and passband ripple are selected so that the inner reflectors contain layers of only quarterwave optical thickness. Table 7-19

lists the designs of the outer reflectors. The transmittances of the three remaining five-cavity bandpasses appear in Fig. 7-47.

7.7.7 A six-cavity bandpass—mixed-cavity design

The passband of the bandpass in the next example manifests a maximally flat spectral shape. Refractive indices of 1.45 and 2.15 are used. The incident medium is air and the substrate is BK-7 glass.

Equation (7-57d) shows that for an η [see Eq. (7-36a)] of 8.44 the standing wave ratios of the reflectors V_3, V_2, V_1 and V_0 are 72, 51.9, 15 and 2.5, respectively. Inspection of Table 7-16 shows that a nine-layer reflector produces V_2.

High-index spacers are used in the center of the bandpass. With a single equivalent layer **E** (of quarterwave optical thickness) and 22 other layers in the center reflector, the required standing wave ratio is obtained if n_E satisfies the following equation

$$V_3 = 72 = n_H^{12}\, n_E^{-2}\, n_L^{-10} \quad . \tag{7-46f}$$

This provides for an n_E of 1.816. The principle of isolation is used (see §7.6.5.1) to produce a center reflector of the following design

high index (**L H**)^2 **E** (**H L**)^3 *high index* .

The next step is to design a reflector with the standing wave ratio V_1 of 15. Table 7-16 shows that a reflector containing six layers produces a standing wave ratio of 15.8, which is close enough to our *desideratum*. This reflector contains an even number of layers and hence has a low-index spacer on one side and a high-index spacer on the opposite side

high index (**L H**)^3 *low index* .

The final task is to design the outer reflector on the air side, which consists of a single equivalent layer **F**. Its refractive index should satisfy

$$V_0 = 2.5 = n_F^2\, n_L^{-1} \tag{7-46g}$$

which results in an **F** index of 1.90. On the glass side, a single equivalent layer **G** should have an index of

$$n_G = \sqrt{2.5 \times 1.45 \times 1.52} = 2.35 \, . \tag{7-46h}$$

For reasons mentioned in §2.6.7.5.6, this equivalent layer must be of three-quarterwave optical thickness rather than one quarterwave optical thickness.

The final design incorporates four equivalent layers

air **F 2L** (**H L**)^3 **2H L** (**H L**)^4 **2H** (**L H**)^2 **E** (**H L**)^3 **2H** (**L H**)^3 **E** (**H L**)^2 **2H**
L (**H L**)^4 **2H** (**L H**)^3 **2L G** *glass*

and its $T(\lambda)$ appears as the dashed curve in Fig. 7-57. Because it contains four contiguous **HH** cavities and two **LL** cavities, it cannot be classified as an alternate-cavity bandpass (see §7.6.5.4) but rather as a *mixed-cavity* bandpass.

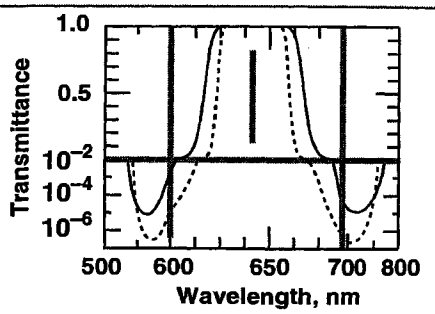


Fig. 7-57 $T(\lambda)$ of a seven-cavity bandpass (*solid*). Its design appears in §7.10.3.3.2. The *dashed curve* is for a six-cavity bandpass whose design appears in §7.10.3.3.1. The ordinate scale changes from linear to log at 0.01. The abscissa changes scale at 600 nm and 680 nm.

7.7.8 Seven-cavity bandpass — homologous cavity

The same refractive indices for the layers and the substrate are used, as in §7.7.7. Low-index spacers are used throughout the coating and *low index* refers to a massive medium of refractive index 1.45. Given an η [see Eq. (7-36a)] of 8.1, Eq. (7-57d) is used to find the standing wave ratios of the reflectors. The result is that V_3, V_2, V_1 and V_0 are 39.6, 23.4, 6.9 and 1.77, respectively.

The first step is to isolate (see §7.6.5.1) the center reflector. Its design is

low index **H L H L E L H L H** *low index*

where the index n_E of the **E** layer is obtained from

$$V_3 = 39.6 = n_H^8 n_E^2 n_L^{-10} \quad . \qquad (7\text{-}46i)$$

The **E** layer index is 1.89 and its phase thickness is 90° at λ_0, which is the central wavelength of the passband.

Inspection of Table 7-16 shows that a seven-layer sequence produces the standing wave ratio of 23.4 for the next intermediate reflector.

The next step is to isolate the adjacent intermediate reflector, which is a five layer reflector

low index **H L F L H** *low index* .

It contains a single equivalent layer **F**, whose refractive index is obtained from

$$V_1 = 6.9 = n_H^4 n_F^2 n_L^{-6} \quad . \qquad (7\text{-}46j)$$

Solution of the foregoing equation provides an n_F of 1.732.

A single equivalent layer is used on the exterior sides of the bandpass as the outer reflector. The analysis is essentially identical to that in §7.7.7. The refractive indices of the equivalent layers are 1.60 and 1.89 on the air and glass sides of the bandpass, respectively. The final design appears in the caption to Fig. 7-57.

7.7.9 Bandpasses containing more than seven cavities

An eight-cavity design appears in the caption to Fig. 7-75. §7.9.10 discusses a bandpass design with seventeen cavities.

7.8 Bandpasses for optical fiber communication

This section focuses upon bandpasses for Wavelength Division Multiplexing WDM. There is a general discussion of this topic in §1.6.3.2. The goal in §7.8 is to produce a bandpass at a central wavelength λ_0 of 1552.5 nm. The filter should manifest a spectral bandwidth of 1.2 nm or less at the –25 dB transmittance level, which permits the spacing of adjacent frequency channels to be at intervals of 100 GHz, as depicted in Fig. 1-85.

At the time of this writing (in the year 2001), a bandwidth of 0.4 nm (at the –0.5 dB transmittance level) is usually required when the following phenomena are

considered: (1) change in the passband wavelength due to temperature drift; and (2) frequency jitter of the communications laser.

Nonabsorbing layers of refractive indices $n_L = 1.47$ (symbol: **L**) and $n_H = 2.065$ (symbol: **H**) are deposited upon a glass substrate of refractive index 1.50. Low-index spacers are used. In the interior of these bandpasses, the goal is to use only layers whose optical thicknesses are integral multiples of quarterwaves. An equivalent layer is permitted as the first layer and/or the final layer of the bandpass. A consequence is that the layers adjacent to either the substrate and/or the incident medium may be layers of nonquarterwave optical thickness.

7.8.1 Choice of the number of cavities

The goal is a bandpass with a passband whose spectral transmittance approximates a rectangle. This is equivalent to a bandpass with a small shape factor [see §1.6.4 (f)]. Figure 7-95 (in §7.10.2) illustrates that the shape factor decreases as the number of cavities increases. As compared to a three-cavity or four-cavity bandpass, a five-cavity bandpass manifests a more rapid drop off of the transmittance curve at the edges of the passband.

Bandpasses are designed containing three, four and five cavities in §7.8.2, §7.8.3 and §7.8.4, respectively. Their attributes appear in Table 7-20 and are summarized in §7.8.5.

7.8.1.1 Control of the bandwidth

Both §7.2 and §7.6.4 illustrate that the bandwidth of a bandpass may be controlled by:

- adjusting the order number of one or more of the spacers;
- controlling the "reflectance of the reflectors" contiguous to those spacers;
- altering the phase dispersion of the reflectors.

These concepts are applied to control the bandwidths of the bandpasses designed in §7.8. The "reflectance of the reflectors" cited above is adjusted by changing V_m [defined in Eq. 7-35)].

7.8.1.2 Standing wave ratios of inner reflectors

Given the refractive indices of the layer materials cited in §7.8, Eq. (7-42) is used to generate a table of the standing wave ratios of inner reflectors, as listed in Table 7-16.

7.8.2 Three-cavity bandpass

At least two design methods may be used to produce a three-cavity bandpass: (1) the periodic-structure design method (§7.3.2); and (2) equations for the standing wave ratios of a Chebyshev bandpass. The latter method is used here.

The code in §7.10.1.7 is used to design a bandpass in which the passband spectral transmittance resembles a Chebyshev polynomial. Two parameters are varied: the log of the maximum standing wave ratio η [see Eq. (7-36a)] and the ripple in the transmittance in the passband. The latter parameter is set at 0.5%. An η of 16.672 results in a bandpass with standing wave ratios of 533 and 406358 for the outer and center reflectors, respectively. Table 7-16 shows that at λ_0, a standing wave ratio of 406353 is produced by a reflector with 37 layers.

The design of the reflector adjacent to air is

air **E** (**L H**)^8 *low index*

and the index for the **E** layer is obtained by solving the equation

$$533 = n_H^{16}\, n_E^2\, n_L^{-17} = n_E^2 (2.065)^{16}\, (1.47)^{-17} \quad . \qquad (7\text{-}46k)$$

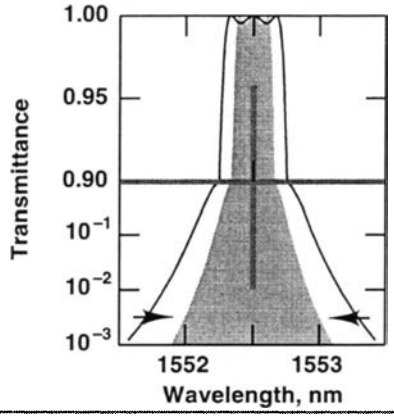


Fig. 7-58 $T(\lambda)$ of a bandpass with three cavities containing first-order spacers (*solid*) and second-order spacers (*shaded*). The ordinate scale changes from linear to log at 0.90. The designs appear in §7.10.3.2.1. The –25 dB spectral width is arrowed.

The result is that n_E is 1.846. The *low index* above is a pseudo-substrate that manifests a refractive index identical to that of an **L** layer. The **E** layer has a quarterwave optical thickness and is replaced by an equivalent layer in the final design.

The design of the reflector adjacent to the glass substrate is

low-index $(\mathbf{H}\,\mathbf{L})^9\,\mathbf{F}$ *glass*

and the index for the **F** layer is obtained from

$$533 = n_H^{18}\, n_F^2\, n_L^{-19}\, n_s^{-1} \quad ,(7\text{-}47a)$$

which results in an n_F of 1.61. As shown in the caption to Fig. 7-58 (see §7.10.3.2.1), the **E** and **F** layers are replaced by equivalent layers of 90° phase thickness.

7.8.2.1 Three-cavity bandpass with third-order spacers

The problem is that the coating developed in §7.8.2 manifests a spectral width at the –25 dB level of 1.58 nm, which exceeds our target of 1.2 nm. A bandwidth of 1.23 nm at the –25 dB level is obtained by using second-order spacers. This, again, exceeds our specifications. The transmittance of the final filter with third-order spacers is depicted as the shaded curve in Fig. 7-58 and its design appears in §7.10.3.2.1. Its bandwidth at the –25 dB transmittance level is slightly in excess of 1 nm.

7.8.3 Four-cavity bandpasses

As stated earlier, a design constraint is that all inner reflectors must be composed of layers of only quarterwave optical thickness layers, or layers that are integral multiples of quarterwaves. The equations in §7.10.1.7.2 are used to design a bandpass in which the passband spectral transmittance resembles a Chebyshev polynomial. As in §7.8.2, the two parameters that are varied include:

- η, which is the log of the maximum standing wave ratio [see Eq. (7-36a)]; and
- the ripple in the transmission in the passband spectral region.

Three solutions are produced, each with a progressively larger value of η.

7.8.3.1 Four-cavity bandpass — first solution

By use of the code in §7.10.1.7, it is found that a passband ripple of 0.0669% and η of 21.2782 produce a bandpass with standing wave ratios of 332, 205916 and 406355 for the outer, intermediate and center reflectors, respectively. Table 7-16 shows that at λ_0 the standing wave ratios are 205920 and 406353 for reflectors with 35 and 37 layers, respectively.

The remaining task is to design the outer reflectors. On the air side, a fifteen layer sequence of the design **H** (**L H**)^7 attains a standing wave ratio of

$$V_0 = n_H^{16}\, n_L^{-15}\, n_0^{-1} = (2.065)^{16}\,(1.47)^{-15}\;\;(1.0003)^{-1} = 338 \qquad (7\text{-}47b)$$

where the refractive index n_0 of air "1.0003" is retained for pedagogical reasons. The foregoing equation is derived from Eq. (2-48a). This result of 338 is quite close to the target of 332.

The reflector on the glass side contains an equivalent layer **E**, as in the design

low index (**H L**)^8 **E** *glass*

where the *low index* is the pseudo-incident-medium and manifests the same refractive index as that of an **L** layer. Equating the standing wave ratio of the foregoing reflector to the *desideratum* of 332 yields

$$332 = V_0 = n_H^{16}\, n_L^{-17}\, n_E^{\,2}\, n_s^{-1} = n_E^{\,2}\,(2.065)^{16}\,(1.47)^{-17}\;\;(1.50)^{-1}\;. \qquad (7\text{-}47c)$$

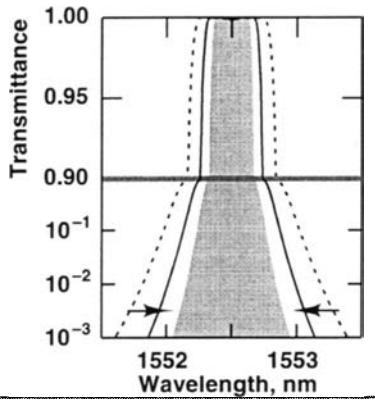


Fig. 7-59 $T(\lambda)$ of bandpasses with four cavities. The ordinate scale changes from linear to log at 0.90. The designs are identified in §7.10.3.2.2 and Table 7-20.

Solution of this equation for n_E yields 1.784. The procedure of §2.11.3 is used to synthesize an equivalent layer of 90° phase thickness and the desired refractive index. As depicted in Fig. 7-59, the passband transmittance (*dashed*) that approaches 100% is evidence that the calculations have been done correctly.

Table 7-20 lists the full spectral width at the –25 dB level as 1.56 nm, which exceeds our target width of 1.2 nm.

7.8.3.2 Four-cavity bandpass — second solution

Following the philosophy of §7.2.2, the passband spectral width is decreased by adding layers to the reflectors, which is equivalent to increasing V_m. This increase in V_m is achieved by adding an **H L** pair to each of the inner reflectors. For example, using a "ripple in percent" of 0.0651 and an η of 22.4588, standing wave ratios of 466, 406349 and 801881 are required for V_0, V_1 and V_2, respectively. The foregoing numbers are obtained from the code in §7.10.1.7. Table 7-16 shows that the standing wave ratios of the inner reflectors are produced by reflectors with 37 and 39 layers.

A standing wave ratio of 444 is produced by the seventeen-layer sequence

low index (**H L**)^8 **H** *glass* ,

which is reasonably close to the target V_0 of 466.

The design of the reflector adjacent to air is

air **E** (**L H**)^8 *low index*

where *low index* is a massive material with a refractive index of 1.47. The **E** layer has a refractive index of 1.726 and quarterwave optical thickness. Its transmittance curve in Figure 7-59 (*solid*) shows that the bandwidth at the –25 dB level is 1.12 nm, which is acceptable.

Table 7-20 Comparison of bandpasses for WDM

Cav	η	Ripl	BW	BW25	Ratio	Lyrs	Tkns	Order	Fig.	Curve	Comment
3	16.672	0.5	0.51	1.584	3.1	115	26.9	1	7-58	solid	
3	16.672	0.5	0.324	1.00	3.1	115	30.1	3	7-58	shaded	
4	21.28	0.07	0.682	1.564	2.5	144	33.9	1	7-59	dash	
4	21.28	0.07	0.53	1.21	2.3	144	36.0	2	7-60	solid	
4	22.46	0.07	0.478	1.12	2.3	152	35.2	1	7-59	solid	
4	23.64	na	0.35	0.79	2.3	160	37.5	1	7-59	shaded	
5	26.96	0.48	0.744	1.27	1.7	185	43.2	1	7-61	solid	
5	28.44	0.48	0.53	0.90	1.7	195	45.5	1	7-61	shaded	
5	27.14	6.0	0.78	1.26	1.62	184	42.7	1, 2	7-62	solid	**
5	27.14	0.05	0.62	1.13	1.8	184	44.5	1, 2	7-62	shaded	PDN **

Comparison of bandpasses: η appears in Eq. (7-36a). *Ripl* (if used) is the ripple in percent used in the design of the bandpass. *BW* is the passband full spectral width (in nm) at a transmittance level of 0.891 and *BW25* is the width (in nm) at an absorbance level of 2.5. *Ratio* is the quotient BW25/BW. *Lyrs* is the layer count. *Tkns* is the metric thickness of the coating in μm. *Order* is the optical thickness of the spacer layers, as multiples of halfwaves at 1552.5 nm. *Fig.* is the figure in which $T(\lambda)$ appears.
Comment: *PDN* Phase-dispersion narrowing used. **Not a Chebyshev

7.8.3.3 Four-cavity bandpass —third solution

Having achieved an adequate 100 GHz bandpass in §7.8.3.2, it is tempting to retire from the field, *hors de combat.* A final four-cavity bandpass is presented, which, with higher order spacers, could evolve into a 50 GHz bandwidth filter.

A "ripple in percent" of 0.0627 and an η of 23.6393 are substituted in the computer code in §7.10.1.7. The result is that the interior reflectors should contain 39 and 41 layers. The design of the outer reflectors follows, quite closely, the procedure of §7.8.3.2. An equivalent layer is inserted into the reflector that is adjacent to the glass and a seventeen-layer quarterwave stack is interposed between the outer spacer and air.

Figure 7-59 illustrates the narrowing of the passband that is produced by increasing the reflectance of the reflectors. Table 7-20 shows that for "solution three," the spectral width at the –25 dB level has shrunk to 0.79 nm.

7.8.3.4 Four-cavity bandpass with second-order spacers

The bandwidth of the coating designed in §7.8.3.1 is too wide. The use of second-order spacers reduces the bandwidth at the –25 dB level to 1.2 nm. Its transmittance is depicted in Fig. 7-60.

7.8.4 Design of five-cavity bandpasses

Five-cavity bandpasses are designed in this section. A design goal is that all inner reflectors are composed of layers of only quarterwave optical thickness layers.

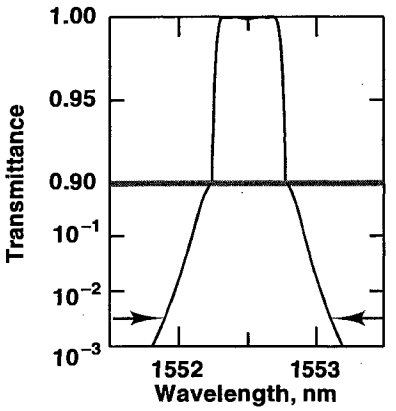


Fig. 7-60 $T(\lambda)$ of a bandpass with four cavities and second-order spacers. The ordinate changes scale from linear to log at 0.90. The design appears in §7.10.3.2.3. The –25 dB spectral width is arrowed.

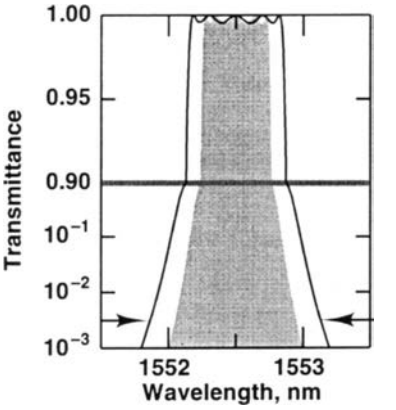


Fig. 7-61 $T(\lambda)$ of bandpasses with five cavities containing 185 layers (*solid*) and 195 layers (*shaded*). The ordinate scale changes from linear to log at 0.90. The designs appear in §7.10.3.2.4.

7.8.4.1 Five-cavity bandpass— first solution

The code in §7.10.1.7 is used to design a bandpass in which the passband spectral transmittance approximates that of a Chebyshev polynomial (see §7.10.1). Two parameters are varied: the passband ripple and η, which is the log of the maximum standing wave ratio [see Eq. (7-36a)]. By permutation of the equations cited in §7.10.1.7, it was found that a passband ripple of 0.4771% and η of 26.9593 produce a bandpass in which the standing wave ratios of the reflectors are 361, 205913 and 406356 for the outer, intermediate and center reflectors, respectively. Table 7-16 shows that standing wave ratios of 205920 and 406353 are produced by reflectors with 35 and 37 layers, respectively. The foregoing values obtain only at the "tuned" wavelength λ_0, which is the passband center.

The remaining task is to design the outer reflectors. The reflector adjacent to the substrate contains 16 quarterwave optical thickness layers and an equivalent layer **F**

low index (**H L**)^8 **F** *glass* .

It manifests a standing wave ratio of

$$V_0 = n_H^{16}\, n_L^{-17}\, n_F^2\, n_s^{-1} = (2.065)^{16}\,(1.47)^{-17}\, n_F^2\,(1.50)^{-1} = 361 \ . \qquad (7\text{-}47d)$$

Manipulation of the foregoing equation yields a refractive index of 1.86 for the **F** layer.

On the air side, the outer reflector design is more complex. The fifteen layer sequence

air **E** (**L H**)^7 *low index*

attains a standing wave ratio V_0 of 361 when the refractive index of the **E** layer is 2.134. As discussed in §2.6.7.5.6, the phase thickness of this equivalent layer cannot be 90°. The reason relates to the refractive index of the layer that is to be synthesized. Its index is 2.134, which exceeds 2.065, i.e., that of the high-index layer.

A consequence is that this layer is simulated by a layer of 270° phase thickness. Two layers of phase thickness 135° are concatenated. The spectral transmittance is

depicted as the solid curve in Fig. 7-61. The problem is that the bandwidth at the –25 dB level of 1.27 nm is slightly in excess of our *desideratum* of 1.2 nm.

7.8.4.2 Five-cavity bandpass— second solution

Similar to the procedure of §7.8.4.1, the design parameters are η (log of the maximum standing wave ratio) and the passband ripple. A passband ripple of 0.4809% and an η of 28.4363 produce a bandpass in which the standing wave ratios of the reflectors are 507, 406348 and 801882 for the outer, intermediate and center reflectors, respectively. According to Table 7-16, standing wave ratios of 406353 and 801879 are produced by quarterwave reflectors with 37 and 39 layers, respectively.

The design of the outer reflectors follows closely the procedure of §7.8.4.1. The design of the reflector adjacent to air is

air **E** (**L H**)^8 *low index*

and the reflector contiguous to the glass substrate manifests the design

low index (**H L**)^9 **F** *glass*

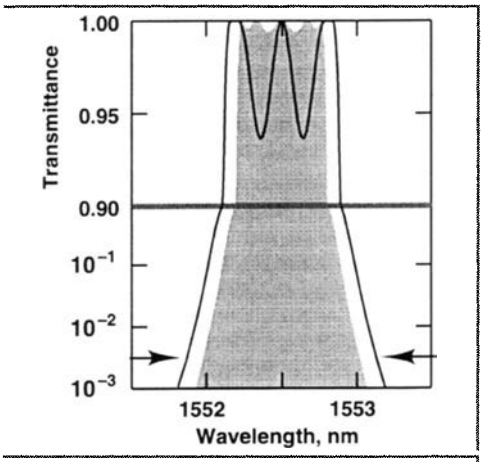


Fig. 7-62 $T(\lambda)$ of a five-cavity bandpass with excessive passband ripple (*solid*) and with phase-dispersion narrowing of the passband spectral width (*shaded*). The ordinate scale changes from linear to log at 0.90. The designs appear in §7.10.3.2.5.

where the **E** and **F** layers are quarterwave optical thickness and manifest refractive indices of 1.80 and 1.57, respectively. As shown in Table 7-20 and graphed in Fig. 7-62 (*shaded curve*), the width at the –25 dB level is 0.9 nm, which, although it meets specifications, is "overkill" because it is too small.

7.8.4.3 Five-cavity bandpass— symmetrical "glass-to-glass" design

The solutions in §7.8.4.1 and §7.8.4.2 present a quandary. On one hand, the bandwidth at the –25 dB level of the first solution is a bit too large and the second solution is decidedly too small.

The bandwidth of the first solution (§7.8.4.1) could be narrowed by using second-order spacers throughout the bandpass. This is objectionable for several reasons. First, since there are five spacers, the total thickness of the bandpass is increased considerably. Second, the bandwidth would be reduced to considerably smaller values than the target bandwidth of 1.2 nm.

The solution is to use phase dispersion to narrow the bandwidth. The starting design is completely symmetrical about the center layer and is "glass to glass"

glass **H** (**L H**)^8 **2L** (**H L**)^{17} **H 2L** (**H L**)^{18} **H 2L** (**H L**)^{18} **H 2L** (**H L**)^{17} **H 2L**
(**H L**)^8 **H** *glass* .

Its passband transmittance, depicted as the solid curve in Fig. 7-62, dips to a value below 0.95. It is not difficult to see the reason that this dip occurs. Compare the foregoing design with the design in §7.8.4.1. The design of the outer reflector in §7.8.4.1 is

low index (**H L**)^8 **E** *glass*

where the E layer refractive index is 1.86. The outer reflector in §7.8.4.3 substitutes an H layer for the E layer. A consequence is the excessive passband ripple shown in Fig. 7-62.

7.8.4.4 Five-cavity bandpass—phase-dispersion narrowing of the passband spectral bandwidth

The objective is to use both higher order spacers and phase-dispersion narrowing to adjust the spectral bandwidth. Two decades ago Minowa [84] showed that the inclusion of higher order spacers in the interior of the bandpass removed the passband ripple. The phase-dispersion narrowing is achieved by adding a layer of three quarterwave optical thickness next to the spacer. Figure 7-62 shows that the bandwidth at the –25 dB level is quite close to 1.0 nm.

A final comment relates to the conversion of the "glass-to-glass" design in §7.8.4.3 to the "air-to-glass" design in this section. Consider the admittance Y of a single H layer that is contiguous to a glass emergent medium. The admittance is pure real and is

$$Y = n_H^2 \; n_s^{-1} = (2.065)^2 \; (1.50)^{-1} \qquad (7\text{-}47e)$$

where n_H is the refractive index of the H layer and n_s is the index of the glass.

It is required that this admittance be unaltered when the H layer is replaced by an equivalent layer E and air (of refractive index n_0 = 1.0003) replaces the glass. This requires that

$$n_E^2 \; n_0^{-1} = (2.065)^2 \; (1.50)^{-1}. \; , \qquad (7\text{-}47f)$$

which shows that $n_E = n_H \; (1.50)^{-1/2} = 1.686$. This intermediate refractive index layer is replaced by an equivalent layer. The design of the completed bandpass appears as the second entry in §7.10.3.2.5.

7.8.5 A comparison of the results

Table 7-20 also lists the ratio of the bandwidth at the –0.5 dB transmittance level to the bandwidth at the –25 dB level. That ratio decreases from a number near 3.0 for the three cavity to nearly half that value for the five-cavity bandpass. This ratio, which is inversely proportional to the "squareness" of the passband spectral transmittance, shows that this "squareness" is increasing as the cavity count increases.

Figure 7-63 graphs the optical thickness of the layers of the three-cavity, four-cavity and five-cavity designs. Some of the methods that are used to achieve the correct bandwidth are now more evident. For example, the second-order low-index spacers in the designs are now more visible. Also evident is the steady increase of the coating thickness as the cavity count progresses from three to five.

7.8.6 Use of a composite spacer layer

Thelen [89] (p. 205) suggests that the sensitivity of a thick spacer layer to thickness errors is mitigated by splitting it up into a succession of halfwave optical thickness layers, alternating between a high index and a low index.

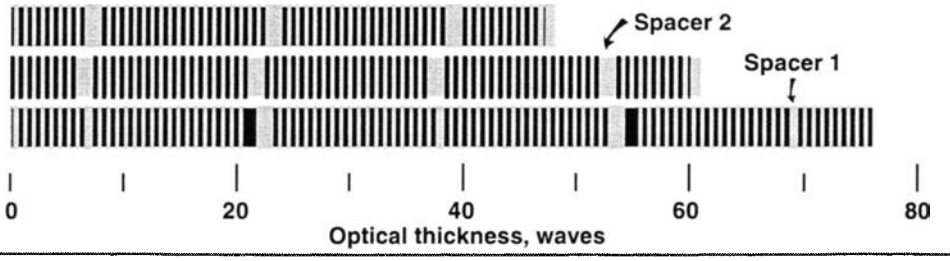


Fig. 7-63 Versus optical thickness (in waves at 1552.5 nm), designs of a three-cavity (*upper*), four-cavity (*middle*) and five-cavity (*lower*) bandpass. The designs appear in the captions: §7.10.3.2.1 (*shaded*), §7.10.3.2.3 and §7.10.3.2.5 (*shaded*), respectively. The dark bands represent high refractive index layers and the lightly shaded bands, low index. Air is on the left and the substrate is on the right. *Spacer 1* and *Spacer 2* represent spacer layers of halfwave optical thickness and fullwave optical thickness, respectively.

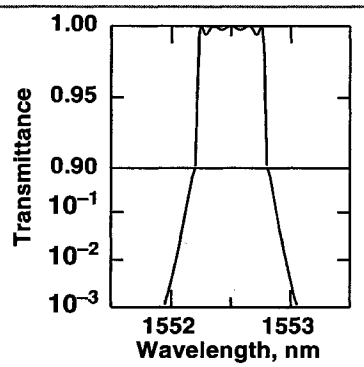


Fig. 7-64 $T(\lambda)$ of a five-cavity bandpass. The refractive indices and meaning of the symbols are identical to those of Fig. 7-62. The design appears in the text in §7.8.6. The ordinate scale changes from linear to log at 0.90.

As an example, suppose that the five-cavity bandpass captioned in §7.10.3.2.5 is modified so that the two intermediate spacer layers are third order. Its design is

air 0.363L 0.260H 0.363L (L H)^8 2L (H L)^{17} H 6L
(H L)^{18} H 2L (H L)^{18} H 6L (H L)^{17} H 2L (H L)^8 H
glass .

Its transmittance appears in Fig. 7-64. When composite spacer layers are used, the third-order spacers are replaced with three halfwave optical thickness layers, alternating between a high and low index

air 0.363L 0.260H 0.363L (L H)^8 2L (H L)^{17} H 2L 2H 2L
(H L)^{18} H 2L (H L)^{18} H 2L 2H 2L
(H L)^{17} H 2L (H L)^8 H *glass* .

The transmittance of the modified design is quite close to that of the original.

7.8.7 Temperature shift of the passband center wavelength

§1.6.4.2 discusses the temperature shift of the passband center wavelength. The focus of that section is on conventional bandpass filters. The corresponding shift of a WDM bandpass is usually required to be a thousand times smaller. Takahashi [95] published a paper of monumental importance. He demonstrated that the temperature shift depended upon the substrate thermal expansion coefficient. It is lessened when the proper glass is used and for that reason, in many instances the substrate index is not the 1.50 cited above, but as large as 1.65.

The reason for this behavior is that the temperature dependence is related to the mechanical stress of the coating after it is removed from the vacuum chamber. The coating is typically deposited at 150 °C. Mechanical stress is induced as the system cools

to room temperature if there is a difference between the thermal expansion coefficient of the glass substrate and the coating. This induced mechanical stress alters the physical thicknesses of the layers by means of the Poisson ratio. The thickness change in turn alters the passband wavelength.

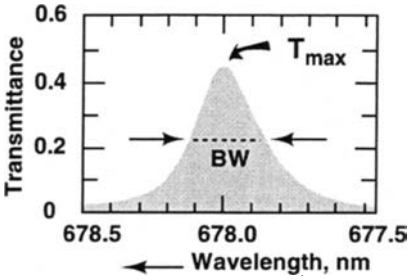


Fig. 7-65 Measured $T(\lambda)$ of a phase conjugate filter deposited on BK-7. Although the patent [Austin 70] is vague about its construction, trials on the computer suggest that its design is

air (0.5L **H** 0.5L)^{10} (0.5M **B** 0.5M)^{10} *glass*

where the refractive indices of **L**, **B**, **M** and **H** are 1.35, 1.35, 2.35 and 2.35, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 570 nm. The optical thickness of **B** and **M** is $\lambda_0/4$ at λ_0 of 788 nm.

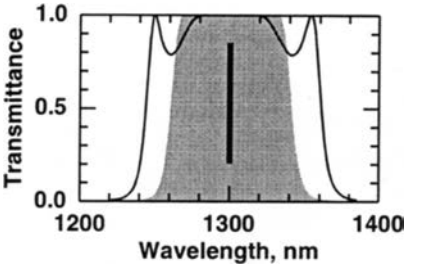


Fig. 7-66 $T(\lambda)$ of a bandpass with equal thickness spacers (*solid*)

glass **H L H 2L H L H L H L H 2L H L H L H L H 2L H L H L H L H 2L H L H L H L H 2L H L H** *glass*

and progressively thicker spacers (*shaded*)

glass **H L H 2L H L H L H L H** 4**L H L H L H L H 6L H L H L H L H** 4**L H L H L H L H 2L H L H** *glass*

where the refractive indices of **L**, **H** and *glass* are 1.45, 2.08 and 1.50, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 1300 nm.

7.9 Additional topics

Additional aspects of bandpass design are discussed.

7.9.1 Phase conjugate bandpass filter

This is a "spacerless" single-cavity bandpass that owes its rather narrow bandwidth to the immense phase-dispersion of a quarterwave reflector (see Fig. 5-93) at the edges of its stop band [Austin 70]. The starting point is Eq. (2-45i), which shows that when two nonabsorbing reflectors are placed in tandem, the transmittance is 100%, provided: (1) the radiant reflectances of the reflectors are identical; and (2) the phase shift upon reflection of one reflector is the complex conjugate of the phase of the other. This is accomplished by adjusting the wavelength centers of the three-layer basic periods $\frac{1}{2}$**L H** $\frac{1}{2}$**L** and $\frac{1}{2}$**H' L'** $\frac{1}{2}$**H'** so that the low frequency edge of the stop band of one group overlaps by about 10% the high frequency stop band edge of the other (see §5.5.1). Figure 7-65 graphs the measured transmittance of such a filter.

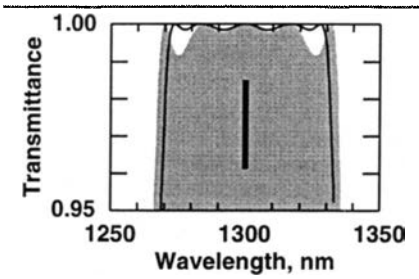


Fig. 7-67 $T(\lambda)$ of a five-cavity bandpass of the design (*solid*) *glass* **H L H 2L H L H L H L H 4L H L H 3L H 3L 3H 2L 3H 3L H 3H L H 4L H L H L H L H 2L H L H** *glass* where the symbols appear in the caption to Fig. 7-66. The shaded curve is from Fig. 7-66.

7.9.2 Miscellaneous design methods

§7.9.2 cites some design methods that do not fit conveniently into the aforementioned categories of *periodic-structure design method* and *microwave design method.*

7.9.2.1 Spacers of progressive thickness

§7.5.8.3.1 alludes to the fact that the "wabbit ears" transmission dips in the passband spectral region of Fig. 7-19 may be attributed to the equal reflectance of the interior reflectors. The same is true of the periodic structure

$$\mathbf{H\,L\,H\ (L\,H\,L\,H\,L\,H\,L\,H\,L)^4\ H\,L\,H}\ .$$

Figure 7-66 shows that even after the addition of the three layer **H L H** antireflection stacks at the ends, there is objectionable ripple in the passband transmittance.

The solution, as suggested by Minowa [84], is to use progressively higher order spacers on the interior of the bandpass. The center spacer is order three and the two inner spacers are second order. Figure 7-66 illustrates that the narrowing of the bandwidth at the center of the bandpass reduces considerably the ripple in the passband transmittance.

7.9.2.2 Phase-dispersion narrowing — bandpass with Chebyshev passband shape

Phase-dispersion narrowing is shown in §7.4.7 to be an alternative to the use of higher-order spacers, as in §7.9.2.1. The phase-dispersion narrowing is achieved by using only a first-order spacer at the center of the bandpass, while using layers of three quarterwave optical thickness adjacent to that spacer. Figure 7-67 illustrates that the passband transmittance approximates that of a Chebyshev, which the original coating does not.

7.9.3 Cavity shifting

The goal is to alter the order number of the cavities in a bandpass. A result of such alteration is that the spectral shape of the passband is distorted. The spectral shape is restored by refractive index optimization. A six-cavity bandpass and a five-cavity bandpass illustrate this procedure.

7.9.3.1 Details of the design procedure

The following procedure is used:

(1) Establish the refractive indices n_H and n_L of the two layers **H** and **L** that will be used in the final design.

(2) Design a bandpass containing spacer layers of equal optical thickness. Design the reflectors (containing an odd number of layers) using n_L for the refractive index of the low-index layers and use refractive indices for the high-index layers that are fictitious, but as close to n_H as possible. An example is in paragraph (3) of §7.7.5.3.2.

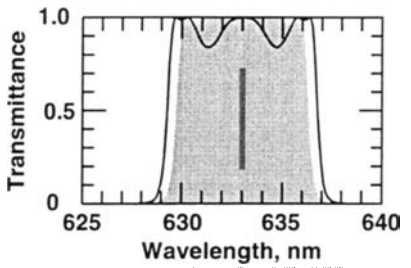


Fig. 7-68 $T(\lambda)$ of the six-cavity bandpass with fictitious indices (*shaded*)

air **A L A** 16**L** (**B L**)^5 **B** 16**L** (**C L**)^6 **C** 16**L** (**D L**)^6 **D** 16**L** (**C L**)^6 **C** 16**L** (**B L**)^5 **B** 16**L** **A L A** *air*

and the shifted-cavity bandpass (*solid*)

air **A L A** 22**L** (**B L**)^5 **B** 10**L** (**C L**)^6 **C** 16**L** (**D L**)^6 **D** 16**L** (**C L**)^6 **C** 10**L** (**B L**)^5 **B** 22**L** **A L A** *air*

where the refractive indices of **L**, **A**, **B**, **C** and **D** are 1.46, 2.10, 2.07, 2.049 and 2.061, respectively. **A**, **B**, **C**, **D** and **L** represent layers of optical thickness $\lambda_0/4$ at λ_0 of 633 nm.

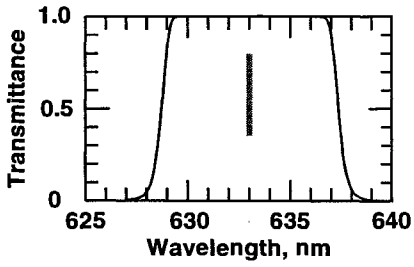


Fig. 7-69 $T(\lambda)$ of the bandpass

air 0.62**H** 1.77**L** 0.62**H** 22**L** **H L H L** 0.57**H** 1.87**L** 0.57**H** **L H L H** 10**L** (**H L**)^3 0.327**L** 0.329**H** 0.327**L** (**L H**)^3 16**L** (**H L**)^3 0.428**L** 0.136**H** 0.428**L** (**L H**)^3 16**L** (**H L**)^3 0.327**L** 0.329**H** 0.327**L** (**L H**)^3 10**L** **H L H L** 0.57**H** 1.87**L** 0.57**H** **L H L H** 22**L** **H L** 0.08**L** 0.827**H** *glass*

where the refractive indices of *glass*, **L** and **H** are 1.52, 1.45 and 2.05, respectively. **H** and **L** represent layers of optical thickness $\lambda_0/4$ at λ_0 of 633 nm.

(3) Perform the "cavity shifting" operation by adding or subtracting thickness from the spacers. In other words, the order number of a few spacers will be increased and the order number of others will be reduced. This operation will, in general, distort the spectral transmission curve of the passband.

(4) Vary those fictitious refractive indices by computer optimization to improve the transmittance in the passband spectral region. The optical thicknesses of all layers should remain constant during this optimization.

(5) Replace the fictitious indices with the "real" indices n_H and n_L, using equivalent layers, if necessary.

7.9.3.2 Shifting of the cavities in a six-cavity bandpass

As illustrated in Fig. 7-68, the spectral transmittance of the shaded design manifests a quasi-Chebyshev response. A procedure similar to that of §7.10.1.2 was used to compute the standing wave ratios of each of the seven reflectors. The high refractive index of each reflector is fictitious.

The order number of each of six spacers is eight and thus the design may be represented symbolically as "8-8-8-8-8-8." Suppose that the outer spacers were thickened and the adjacent inner spacers were made thinner, so that the design reduces to "11-5-8-8-5-11." This operation is termed "cavity shifting." That is, the order number of the outer cavity has been increased to eleven and the order number of the adjacent inner cavity has been decreased to five.

When reduced to layers whose optical thicknesses are multiples of a quarterwave, the design of the bandpass "8-8-8-8-8-8" in Fig. 7-68 may be presented as

A L A 16L **(B L)^5 B** 16L **(C L)^6 C** 16L **(D L)^6 D** 16L **(C L)^6 C** 16L **(B L)^5 B** 16L **A L A**

where the refractive indices appear in the caption to Fig. 7-68.

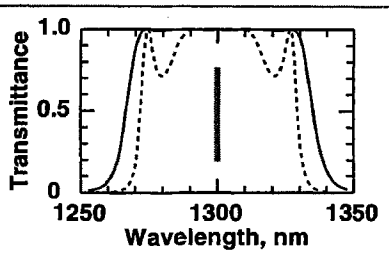


Fig. 7-70 $T(\lambda)$ of a five-cavity bandpass with spacers of unequal order (*solid*) *glass* **H L H 2L (H L)^3 H 4L (H L)^3 H 6L (H L)^3 H 4L (H L)^3 H 2L H L H** *glass* and a five-cavity bandpass with spacers of equal order (*dashed*) *glass* **H L H 6L (H L)^3 H 6L (H L)^3 H 6L (H L)^3 H 6L (H L)^3 H 6L H L H** *glass* where **H** and **L** represent layers of optical thickness $\lambda_0/4$ at λ_0 of 1330 nm. The refractive indices of *glass*, **H**, **L** are 1.51, 2.20 and 1.45, respectively.

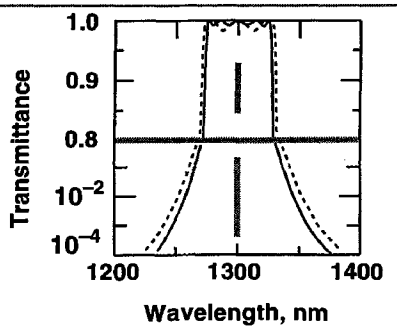


Fig. 7-71 $T(\lambda)$ of a five-cavity bandpass with spacers of equal order (*solid*) *glass* 0.306L 0.364H 0.306L **L H 6L H L H** 0.396H 0.192L 0.396H **H L H 6L H L H** 0.118H 0.746L 0.118H **H L H 6L H L H** 0.118H 0.746L 0.118H **H L H 6L H L H** 0.396H 0.192L 0.396H **H L H 6L H L** 0.306L 0.364H 0.306L *glass* and with spacers of unequal order (*dashed*) *glass* **H L H 2L (H L)^3 H 4L (H L)^3 H 6L (H L)^3 H 4L (H L)^3 H 2L H L H** *glass* where the symbols are captioned in Fig. 7-70.

Shifting of the cavities to the design "11-5-8-8-5-11" distorts the passband transmittance, as is evident by the solid curve in Fig. 7-68. Computer optimization that altered the refractive indices (while maintaining constant optical thickness of each layer) was then used to improve the passband transmittance. The refractive indices of **L**, **A**, **B**, **C** and **D** are 1.46, 2.10, 2.07, 2.049 and 2.061, respectively. Figure 7-69 depicts the transmittance after the fictitious indices have been replaced with n_H and n_L.

In the optimization, no effort was made to alter the transmittance in the spectral region outside of the passband. A result is that the η of the optimized design decreased, which in turn lowers the level of the offband rejection.

7.9.3.3 Shifting of cavities in a five-cavity bandpass

The progression "1-2-3-2-1" represents the relative thicknesses of the spacer layers of a five-cavity bandpass [Minowa 84]. In other words, the thicknesses progress from thin on the outside to thick in the interior. Figure 7-70 shows that the bandpass manifests excellent passband transmittance. As depicted by the dashed curve, however, the incorporation of third-order spacers throughout the bandpass (i.e., a "3-3-3-3-3" design) produces "wabbit ears" transmission dips in the passband.

There are six reflectors, two of which surround each of the spacer layers. Due to symmetry, however, the aforementioned six reduce to three. A fictitious refractive index was assigned to the high index in each of the three reflectors, resulting in the design

glass **A L A 6L (B L)^3 B 6L (C L)^3 C 6L (C L)^3 C 6L (B L)^3 B 6L A L A** *glass*

where **A**, **B**, **C** and **L** represent layers of quarterwave optical thickness at the central wavelength of 1300 nm.

The passband transmittance of the foregoing design was then improved by computer optimization. The refractive indices of layers **A**, **B** and **C** were the variables, while the optical thickness of each layer was maintained constant. The result was that the indices of layers **A**, **B** and **C** were altered to 2.004, 2.046 and 2.183, respectively. The fictitious indices cited above were converted into "real" indices. The transmittance of the final two-component design appears in Fig. 7-71, which illustrates that the passband spectral shape resembles a Chebyshev with 1% ripple.

A comparison between the original Minowa-Fujii design and the shifted-cavity design shows that the coatings contain nearly the same number of layers. Figure 7-71 shows that the spectral width at the 3.0 absorbance level is narrower for the shifted-cavity bandpass. That narrowing is, however, at the expense of a coating that is increased in thickness by 27%, due to the higher order spacers of the "3-3-3-3-3" design.

7.9.4 Effects of absorption upon passband transmittance

Absorption in any of the layers may produce either of two effects: (1) if the absorption is relatively low, the peak transmittance in the passband is reduced; and (2) higher levels of absorption distort the spectral shape of the passband. A computer program is needed to quantify such distortion. The former effect is estimated by a simple procedure outlined below.

7.9.4.1 Equations that calculate approximately the absorption

At the tuned wavelength λ_0 the absorption A_i in the ith layer of a multilayer of quarterwaves is calculated by the procedure of §2.9.6.1. The total absorption A is found by summing the absorption in each layer [see Eq. (2-62a)]. Equation (2-63e) is used to find A_i. Layers of halfwave optical thicknesses are split into two contiguous quarterwave layers. The standing wave ratio V_i at the right interface of the ith layer is found using the procedure:

1. Compute the reflectance R of the entire multilayer at λ_0. This should be close to zero for a carefully designed bandpass. $1 - R$ in Eq. (2-63e) is now known.
2. The substrate is the "ℓ +1" layer for a multilayer of ℓ layers. The reduced admittance y is 1.0 (by definition) in the substrate.
3. The reduced admittance y_i at the emergent interface of the ith layer is found recursively by the formula
$$y_i = n_i \ (n_{i+1} \ y_{i+1})^{-1} \ . \qquad (7\text{-}47g)$$
The foregoing is an approximation (albeit a good one) because the absorption is neglected.
4. Use Eq. (2-20b) to convert y_i into standing wave ratio V_i.
5. Find G_i using Eq. (2-63f).
6. The metric thickness h_i of each layer is found by solving Eq. (2-47i)—since its quarterwave optical thickness and index are known.
7. Estimates are made of the extinction coefficient k_i of each layer. The bulk absorption α_i is found from Eq. (2-14b).
8. Substitute the numerical values from the previous steps into Eq. (2-63e).

Table 7-21 Absorption in a four-cavity bandpass

Layer	*V*	*A* in %	Layer	*V*	*A* in %	Layer	*V*	*A* in %
¶¶1	1.40	0.012	21	12.4	0.081	41	2.58	0.019
2	1.59	0.014	22	12.4	0.081	**42**	1.83	0.055
3	2.25	0.062	**23**	8.74	0.203	43	1.29	0.013
4	3.18	0.023	24	6.18	0.041	**44	1.16	0.012
5	4.50	0.109	**25**	4.37	0.106	45	1.05	0.013
6	6.37	0.042	26	3.09	0.022	**46**	1.35	0.048
7	6.37	0.042	**27**	2.19	0.061	47	1.91	0.016
8	4.50	0.109	28	1.55	0.014	**48**	2.70	0.070
9	3.18	0.023	**29**	1.09	0.046	49	3.81	0.026
10	2.25	0.062	30	1.29	0.013	**50**	5.39	0.128
11	1.59	0.014	**31**	1.83	0.055	51	7.62	0.050
12	1.13	0.046	32	2.58	0.019	52	7.62	0.050
13	1.25	0.013	**33**	3.65	0.090	**53**	5.39	0.128
14	1.39	0.012	34	5.16	0.035	54	3.81	0.026
15	1.55	0.014	**35**	7.30	0.171	**55**	2.70	0.070
**16	2.19	0.061	36	10.3	0.068	56	1.91	0.016
17	3.09	0.022	37	10.3	0.068	**57**	1.35	0.048
18	4.37	0.106	**38**	7.30	0.171	Total absorption		3.29%
19	6.18	0.041	39	5.16	0.035			
20	8.74	0.203	**40**	3.65	0.090			

** Index of this layer is 1.61. ¶¶Index of this layer is 1.62

Versus layer number of a four-cavity bandpass whose design is captioned in Fig. 7-53, standing wave ratio *V* is at the emergent interface of the layer. In computing *V*, the absorption is neglected. *A* is the approximate absorption (in percent) in the layer, assuming an extinction coefficient of 0.0003 for the layers of refractive index 2.05 and an extinction coefficient of 0.00006 for all others. Each layer has an optical thickness of $\lambda_0/4$—each halfwave spacer is subdivided into two QWs. Unless designated otherwise, all **boldface** numbered layers have an index of 2.05. The remaining layers have a refractive index of 1.45. The incident medium is *air* and the substrate refractive index is 1.52.

7.9.4.2 Numerical example

The absorption is found in each layer of the four-cavity bandpass captioned in Fig. 7-53. The equivalent layers in that design have been transformed into actual layers of quarterwave optical thickness. The extinction coefficient of an **H** layer is 0.0003 and the *k* value of all other layers is 0.00006—five times smaller.

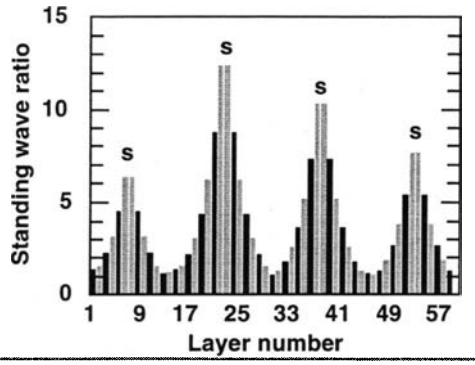


Fig. 7-72 At a wavelength of λ_0, standing wave ratio (at the emergent interface of each layer) vs layer number of the bandpass captioned in Table 7-21. The letter *s* denotes a spacer layer of halfwave optical thickness.

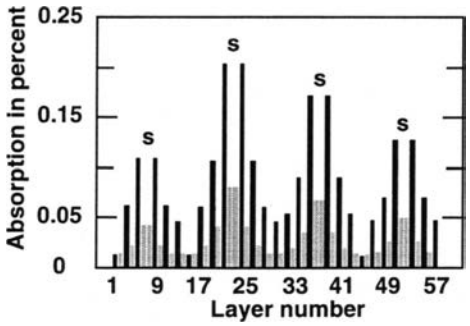


Fig. 7-73 At a wavelength of λ_0, absorption vs layer number of the bandpass whose construction is listed in Table 7-21. The letter *s* designates a spacer layer of halfwave optical thickness.

Table 7-21 summarizes the calculations, which follow the procedure of §7.9.4.1. The final column lists the absorption in each layer. The sum, i.e., the total absorption from Eq. (2-62a), is 3.29%. The actual total absorption is quite close to the approximate — 3.36%. The standing wave ratio V and absorption in each layer are graphed in Figs. 7-72 and 7-73, respectively. Equation (2-63e) predicts that the absorption within any layer scales as the standing wave ratio V, which is evident when Figs. 7-72 and 7-73 are compared. The absorption peaks in the cavities and layers adjacent to the cavities, where the V is the largest.

7.9.4.3 Absorption — computed using exact equations

The absorption may also be computed by the more tedious matrix method (see §2.4.4). As mentioned in the previous paragraph, the approximate value of the absorption is reasonably close to the exact value.

The matrix method is used to find the transmittance at wavelengths other than the tuned wavelength of λ_0. As an example, Fig. 7-74 shows the $T(\lambda)$ of a four-cavity filter with varying amounts of absorption in material of higher refractive index. The conclusion is that the shape of the passband is not appreciably distorted by the absorption, although the peak transmittance is degraded by as much as 30% when the magnitude of the extinction coefficient exceeds one part per thousand.

As another example, Fig. 7-75 shows the transmittance of an eight-cavity bandpass [Goosey 74]. The extinction coefficient of the high-index component of the multilayer is absorbing in several of the designs. The distortion of the spectral shape of the passband at the higher levels of absorption is evident. Compare this bandpass with the four-cavity bandpass captioned in Fig. 7-74. The extinction coefficients of high-index layers in the eight-cavity filter are four times larger than those of the four-cavity bandpass, yet the peak transmittances are nearly the same. The conclusion is that, because of the smaller standing wave ratios of the eight-cavity bandpass, a larger amount of absorption may be tolerated before the passband transmittance is severely degraded.

7.9.5 Production of a bandpass—effects of layer thickness errors

At the time of this writing (in the year 2001), most bandpasses are produced by vacuum evaporation. The thicknesses of the layers are controlled by either: (1) a quartz crystal (see §9.6.1.2); or (2) optical monitoring. §9.6.6.2 shows that there are distinct advantages to composite optical monitoring in transmittance.

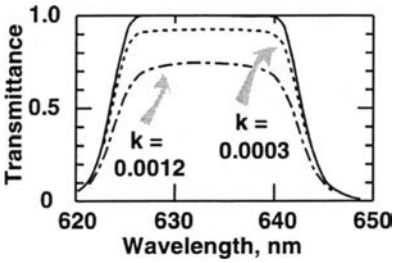


Fig. 7-74 $T(\lambda)$ of a four-cavity bandpass
air 0.392**L** 0.204**H** 1.392**L** **H L H** 2**L**
$(\mathbf{H\,L})^2$ **H** 1.399**L** 0.192**H** 1.399**L H**
$(\mathbf{L\,H})^2$ 2**L H** $(\mathbf{L\,H})^6$ 2**L** $(\mathbf{H\,L})^2$ **H** 1.399**L**
0.192**H** 1.399**L H** $(\mathbf{L\,H})^2$ 2**L H L H L H**
glass
where the refractive index of *glass* is 1.52. The complex optical constants of **H** and **L** are $2.05 - j\,k$ and $1.45 - j\,k$, respectively. k is 0 (*solid*), 0.0003 (*short dash*) and 0.0012 (*long-short dash*). The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 633 nm.

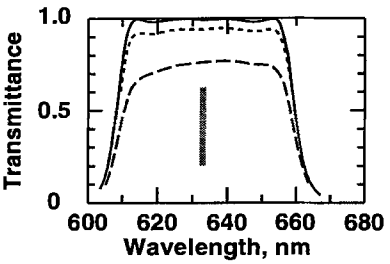


Fig. 7-75 $T(\lambda)$ of an eight-cavity bandpass
air **K H LL** $(\mathbf{H\,L})^2$ **HH** $(\mathbf{L\,H})^2$ **K H LL**
$(\mathbf{H\,K})^3$ **HH** $(\mathbf{K\,H})^3$ **LL H K** $(\mathbf{H\,L})^2$
HH $(\mathbf{L\,H})^2$ **LL H KK** *glass*
where the refractive index of *glass* is 1.52. The complex optical constants of **H**, **L** and **K** are $2.30 - j\,k$, $1.50 - j\,0$ and $1.35 - j\,0$, respectively. k is 0 (*solid*), 0.001 (*dash*) and 0.005 (*shaded*). The optical thickness of **K**, **L** and **H** is $\lambda_0/4$ at λ_0 of 633 nm.

At least two construction parameters of the bandpass may drift during its deposition: (1) refractive index; and (2) either optical thickness or metric thickness. If quartz crystal monitoring is used and the index drifts, then both optical thickness and index will change. If, on the other hand, optical monitoring is used, then the optical thicknesses will be preserved and only the refractive index will vary.

Regardless of which of the thickness control methods mentioned in the first paragraph of §7.9.5 is used, the following generalizations are valid: (1) if the optical thicknesses are preserved and yet there is a drift in the refractive index, the peak transmittance of the passband is reduced; (2) errors in the optical thicknesses of the layers degrade both the peak passband transmittance and the bandwidth; (3) thickness errors in spacer layers and layers near the spacers have a much greater effect than do errors in other layers.

Thickness errors in a two-cavity bandpass illustrate the last point above. It contains 19 layers—its design is captioned in Fig. 7-76. An ensemble of 500 such coatings was constructed on the computer, each layer having a random thickness error (Gaussianly distributed) of ±0.5%. The shaded spectral region in Fig. 7-76 represents the highest and lowest deviation in the transmittance at each wavelength.

The next step is to vary the thicknesses of nine layers of QW thickness at the center and extremities of the bandpass. These are arrowed below:

$\downarrow\downarrow\downarrow \qquad\qquad \downarrow\downarrow\downarrow \qquad\qquad \downarrow\downarrow\downarrow$
air **H L H L H H L H L H L H L H H L H L H** *glass* .

Figure 7-77 shows that the influence of these thickness errors is minimal.

The final "experiment" on the computer is to adjust randomly the thicknesses of just six of the 19 layers in the bandpass — each of the two spacer layers and the QW layer adjacent to that spacer

air **H L H L HH L H L H L H L H L HH L H L H** *glass* .

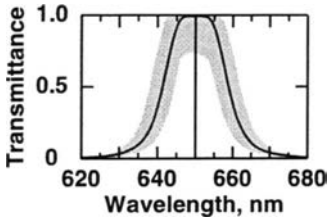


Fig. 7-76 $T(\lambda)$ of a two-cavity bandpass (*solid*) *air* **H L H L HH L HL H L H L HH L H L H** *glass* where the indices of *glass*, **L** and **H** are 1.52, 1.35, and 2.30, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$. The shaded area is the maximum deviation of transmittance produced by a random, Gaussianly distributed ±0.5% thickness variation in each layer.

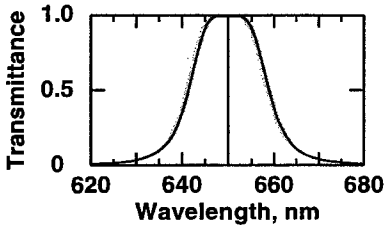


Fig. 7-77 The caption to Fig. 7-76 obtains, with the exception that the thicknesses of only layers 2, 3, 9, 10, 11, 17 and 18 are varied.

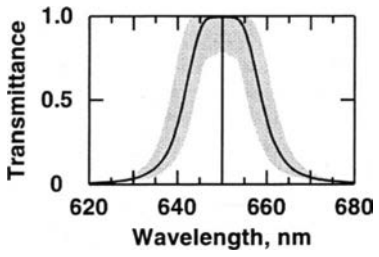


Fig. 7-78 The caption to Fig. 7-76 obtains, with the exception that the thicknesses of only layers 4, 5, 6, 14, 15 and 16 are varied.

Figure 7-78 illustrates that the statistical deviation of the transmittance is virtually the same as that of Fig. 7-76—in which the thicknesses of all layers are varied.

7.9.6 Performance of a bandpass at nonnormal incidence

There are instances in which a bandpass must function at nonnormal incidence. Strong polarization splitting of the passbands usually occurs. An example is the five-cavity bandpass whose $T(\lambda)$ curve appears as the *light solid* curve in Fig. 7-47. That bandpass is modified so that the thickness of each layer is *matched* at 45° (see §2.7.2.2). Figure 7-79 shows that the matched thicknesses cause the spectral centers of the passbands to be centered at 633 nm. However, the *p* bandwidth is nearly double the *s* bandwidth. Although the filter could be redesigned to remove passband ripple for either the *s* or *p* polarization, the strong polarization splitting still remains for reasons enumerated below.

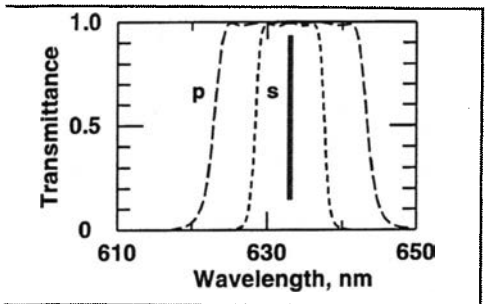


Fig. 7-79 At a 45° angle of incidence in air, $T(\lambda)$ for *p* polarization (*short dash*) and *s* polarization (*long dash*) of a five-cavity bandpass. The design appears in §7.10.3.5.4, with the exception that the layer thicknesses are *matched* at 45°.

The difference in bandwidth for *p* and *s* is caused by the disparity between the standing wave ratios of each reflector. Consider, for example, the center reflector

low index $(\mathbf{H}\,\mathbf{L})^7\,\mathbf{H}$ *low index* .

Effective indices (see §2.5.3) may be substituted for the actual indices into Eq. (7-42) to compute its standing wave ratio at an incidence of 45° in air. The actual indices of **L** and **H** are 1.45 and 2.05, respectfully. The effective indices N_p of **L** and **H** are 1.661 and 2.184, respectively, for *p* polarization. For *s* polarization the corresponding indices are, respectively, 1.266 and 1.924. The standing wave ratio V_p of the foregoing reflector for *p* polarization is

$$V_p = N_H^{16}\,N_L^{-16} = (2.184)^{16}\,(1.661)^{-16} = 80 \qquad \text{(7-47h)}$$

and for *s* polarization it is

$$V_s = N_H^{16}\,N_L^{-16} = (1.924)^{16}\,(1.266)^{-16} = 810\ . \qquad \text{(7-47i)}$$

The *s* standing wave ratio exceeds the *p* standing wave ratio by a factor of ten. A similar disparity exists for the standing wave ratios of the other reflectors. A consequence is that the V_m [see Eq. (7-35)] is smaller for *p* than the V_m for *s*. §7.6.4.7 argues that a smaller V_m produces a wider bandwidth, which is evident in Fig. 7-79 for the *p* polarization.

7.9.6.1 Bandpass with reduced polarization splitting

The foregoing discussion suggests a method of designing a bandpass with nearly equal bandwidth for *s* and *p* [Baumeister 92a]. The design of each interior reflector is

spacer $\mathbf{M}\,(\mathbf{H}\,\mathbf{M}\,\mathbf{L}\,\mathbf{M})^q$ *spacer* .

Such a reflector manifests equal standing wave ratio for *p* polarization and *s* polarization if the refractive indices are chosen using the teachings of §6.4.4. The angle shift of the bandpass is lessened by using spacers of higher index (i.e., $n_{spacer} = 2.30$).

Many of the bandpasses designed previously manifest an η [see Eq. (7-36a)] in excess of 7. The η is reduced to 5.33 for this bandpass in order that the number of layers may be reduced. A five-cavity maximally flat prototype bandpass manifests standing wave ratios of 2.0, 9.0 and 25.6 for the outer, middle and inner reflectors, respectively.

The coating is immersed in a cemented cube in which the angle of incidence is 45° in the *reference medium* index [as it appears in Eq. (2-4a)] of 1.73. To design the inner reflector, the left of Eq. (6-7a) is set equal to $(25.6)^{-1}$ and the left of Eq. (6-7b) is equated to 1.00. Both n_0 and n_s in Eq. (6-7a) are set equal to n_{spacer} . With an index of 1.60 for the $\mathbf{M_c}$ layer, a six-period reflector with indices of 1.386 and 2.277 for the low-index layers and high-index layers, respectively, produces the requisite standing wave ratio. A frightening thought is at this early stage in the design the bandpass (consisting of two reflectors and a spacer layer) contains 51 layers.

The procedure in the foregoing paragraph is repeated for the middle reflector. The U value on the left of Eq. (6-7a) is 1/9. Given the index of 1.60 for the $\mathbf{M_b}$ layer, a reflector with four periods is adequate to produce the standing wave ratio of 9.0.

The final step is to design the outer reflector. The incident medium index is that of the glass prism (i.e., 1.73) and the emergent medium index is that of the spacer — n_{spacer} = 2.30. The left of Eq. (6-7a) is equated to one-half. Given an index of 1.66 for the $\mathbf{M_a}$ layer, a single period produces the proper standing wave ratio.

The center panel of Fig. 7-80 depicts the transmittance of the completed filter at the "tuned" angle of 45°. The *solid* curve is the *s* polarization. The *dashed* or *shaded* curves are the *p*. The polarization splitting of the passband is minimal, both at angles of 44° (in the glass) and at 46°.

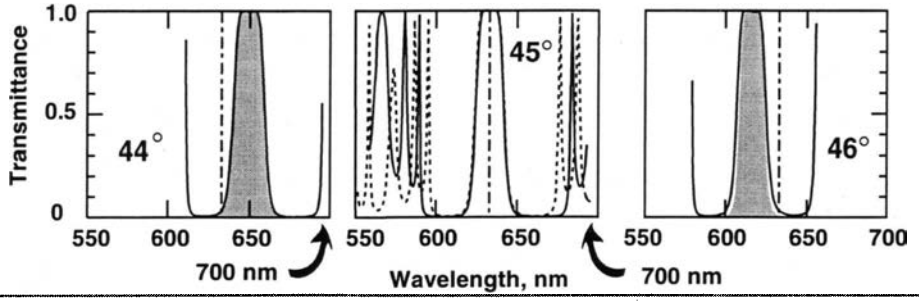


Fig. 7-80 At the indicated angles of incidence in the *glass* of refractive index 1.73, $T_p(\lambda)$ (*dashed* and *shaded*) and $T_s(\lambda)$ (*solid*) of the bandpass

glass $\mathbf{M_a\ H_a\ M_a\ L_a\ M_a\ HH\ M_b\ (H_b\ M_b\ L_b\ M_b)^4\ HH\ M_c\ (H_c\ M_c\ L_c\ M_c)^6\ HH}$

$\mathbf{M_c\ (H_c\ M_c\ L_c\ M_c)^6\ HH\ M_b\ (H_b\ M_b\ L_b\ M_b)^4\ HH\ M_a\ H_a\ M_a\ L_a\ M_a}$ *glass*

where the refractive indices of $\mathbf{L_a}$, $\mathbf{M_a}$, $\mathbf{H_a}$, $\mathbf{L_b}$, $\mathbf{M_b}$, $\mathbf{M_b}$, $\mathbf{L_c}$, $\mathbf{M_c}$, $\mathbf{H_c}$ and **HH** are 1.381, 1.660, 2.343, 1.381, 1.600, 2.228, 1.386, 1.600, 2.277 and 2.30, respectively. Each layer has an optical thickness of $\lambda_0/4$ at λ_0 of 633 nm, *matched at* 45°. The vertical *long-short dashed* line designates λ_0.

Few would argue that this coating (containing 99 layers) is easy to manufacture, because of the many different refractive indices that must be used. The author is hopeful that a computer code (yet to be written) will reduce these to three. As Thelen [89] has shown, a bandpass such as the one in Fig. 7-80 may also be used as an edge filter.

7.9.7 Single-cavity bandpass filter used as a linear polarizer

This system is depicted in Fig. 1-176 and was invented by Ron Ward [96]. It has high *p* transmittance and quite narrow transmission "spikes" for *s* polarization. As long as one of those "spikes" does not coincide with a laser line, it is a superb linear polarizer for that laser. The angle of incidence inside the glass is 41.13°, which is less than 1° from the critical angle (see §2.2.2.2). A system that functions at 1.065 μm is

glass **AR** *gap* **AR** *glass*

where *gap* is an air gap between two prisms. The **AR** is the five layer all-dielectric coating (for *p* polarization) cited in §4.9.5.5. It is coated on each prism.

This system is a single-cavity bandpass. Because the *p* reflectance of each reflector (see Fig. 4-58) is quite low, the *p* transmittance of the assembly is large and relatively independent of the gap spacing.

Figure 4-58 shows that the *s* reflectance of the AR approaches 99%. A consequence is that the *Q*—the reciprocal fractional bandwidth [see Eq. (2-83h)] —is large and hence the transmission "spikes" for *s* polarization are quite narrow. Figure 1-177 shows that T_s is at least 3 or 4 absorbance units at all wavelengths except where these narrow transmission "spikes" occur. During assembly the air gap thickness is adjusted so that those spikes are well separated from the laser wavelength at which the polarizer functions.

7.9.8 Selection of the number of cavities

There are at least three reasons that a particular number of cavities would be selected for a given application:

(1) The achievement of a prescribed spectral width of the passband;
(2) A prespecified shape factor, which is related to the steepness of the decrease of the $T(\lambda)$ curve at the edges of the passband; and
(3) A reduction in the effects of absorption and an increase in the manufacturing tolerances.

7.9.8.1 log(V_m) vs cavity count for a given bandwidth

This subject is approached by posing the following problem:

> Using the minimum number of layers, design multiple-cavity bandpasses that produce a bandwidth of less than 7 nm at the 3.0 absorbance level, given λ_0 of 1552 nm. The lowest refractive index that may be used is 1.45 and the highest is 2.08.

Note that only the –30 dB bandwidth has been specified and that the narrowness of the bandwidth at the –0.5 dB level is of no concern—at the present, at least. Table 7-22 lists the η and number of layers in bandpasses containing 1, 2, 3 and 4 cavities.

7.9.8.1.1 log(V_m) vs cavity count

The starting point of the discussion is the *R/T* of a maximally flat bandpass, as given in Eq. (7-27c), which is an approximation

$$F = \tfrac{1}{4} V_m \ [\sin(\beta)]^{2q} \ . \tag{7-47j}$$

When the absorbance is 3.0, *F* is 999 and $2\Delta\beta$ is (by definition) the full width of the passband (in units of phase)

$$999 \times 4 = V_m \ [\sin(\Delta\beta)]^{2q} \ . \tag{7-47k}$$

Taking the decadic log of both sides of the foregoing equation yields

$$3.60 = \eta - q\,2\,|\{\log[\sin(\Delta\beta)]\}| \tag{7-48a}$$

or

$$\eta = q\,2\,|\{\log[\sin(\Delta\beta)]\}| + \text{constant} \tag{7-48b}$$

where the *constant* depends upon the transmittance level at which the bandwidth is measured. This linear dependence of η upon the cavity count is evident in Fig. 7-81.

Table 7-22 Bandpasses with varying numbers of cavities

q	BW3	Layers	η	Thickness	Abs
1	6.8	43	6.9	9.99	
2	6.4	63	10.0	14.8	2.01%
3	6.0	87	13.8	20.5	
4	6.9	115	16.8	26.4	1.6%

For bandpasses containing q cavities, full spectral passband width BW3 (in nm) at the –30 dB transmittance level, number of layers *Layers*, η [see Eq. (7-36a)] and total thickness *Thickness* in μm. With a k value of 5×10^{-6} assigned to each layer, *Abs* is the total absorption at the center wavelength of the passband. The largest layer refractive index is 2.08 and the smallest is 1.45.

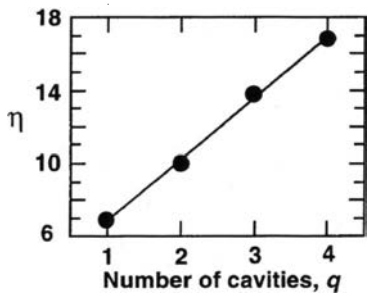


Fig. 7-81 Versus number of cavities q, η [see Eq. (7-36a)] of the bandpasses cited in Table 7-22.

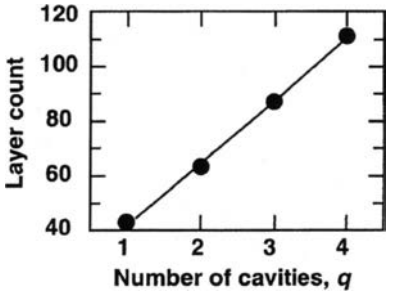


Fig. 7-82 Versus number of cavities q, layer count of the bandpasses cited in Table 7-22.

7.9.8.1.2 Layer count vs cavity count

It is argued that the layer count is directly proportional to η, provided that essentially the same refractive indices are used throughout the bandpass. Equation (7-36a) defines η. An example is a two-cavity bandpass, in which V_0 and V_1 are the standing wave ratios of the outer and inner reflectors (containing, respectively, q_1 and q_2 layers of quarterwave optical thickness), respectively. Then from Eq. (7-35) it follows that V_m (and its log) are

$$V_m = V_0\, V_1\, V_0 \qquad \text{or} \qquad \eta = \log V_0 + \log V_1 + \log V_0\ . \tag{7-48c}$$

Because indices of 1.45 and 2.08 are used in this bandpass, it follows from Eq. (7-42)

$$\eta = 2\,(q_1 + 1)\log(2.08/1.45) + (q_2 + 1)\log(2.08/1.45) \tag{7-48d}$$

and thus η is proportional to the sum—$2\,q_1 + q_2$, which is the layer count. This proportionality is evident in Fig. 7-82, which plots the layer count vs the number of cavities. If a different refractive index ratio were used for the layers in the bandpass, then the slope of the line would change.

7.9.8.1.3 Absorption in a four-cavity bandpass

In §7.9.4.1 the absorption in a four-cavity bandpass is computed. Table 7-21 shows that the absorption in any layer is proportional to the standing wave ratio V in that layer. In the bandpass that is analyzed in §7.9.4.1, the Vs range as large as 12. The net result is that a k value of 3×10^{-4} produces an absorption of about 3% in this 57 layer coating.

Contrast this to the bandpasses cited in Table 7-22. Figure 7-83 plots the standing wave ratios for both the two-cavity and the four- cavity bandpasses. As compared to the bandpass in §7.9.4.1, Fig. 7-83 shows that many of the standing wave ratios are twenty times larger The conclusion is that the extinction coefficient of any of the layers used in bandpasses for optical communication must be at least twenty times smaller in order to maintain an acceptable level of absorption for the entire bandpass.

The next step is to compare the standing wave ratios and the absorption in the two-cavity bandpass and in the four-cavity bandpass. Figure 7-83 shows that the peak standing wave ratios in the four-cavity bandpass are about 60% smaller than those in the two-cavity bandpass. However, the absorption is not 60% less, because there are nearly twice as many layers in the four cavity, as compared to the two cavity. The final column in Table 7-22 lists the total absorption for the two cavity and the four cavity. The smaller standing wave ratios in the four-cavity bandpass results in a total absorption that is about 80% of the total absorption in the two-cavity bandpass.

7.9.8.1.4 Narrowing of the spectral bandwidth by increasing V_m

This topic is treated empirically in §7.6.4.7. The approach below is more mathematical. A maximally flat bandpass is considered—one whose spectral transmittance is characterized by Eq. (7-27c). The goal is find the fractional spectral bandwidth of a prototype with a maximally flat shape. The starting point is Eq. (7-27c) which expresses the R/T of such a filter containing q cavities. The center of the passband occurs when $m\pi = \beta_0$ in that equation, where m is an integer. The R/T ratio when the transmittance is –3 dB is 1.0 and this defines the variable F_r The phase thickness of the prototype is $\beta_r \equiv \beta_0 + \Delta\beta_r$, at the –3 dB level. Solving Eq. (7-27c) for $\Delta\beta_r$ results in

$$\Delta\beta_r = \arcsin[(F_r/F_{max})^{1/(2q)}] \ . \qquad (7\text{-}48e)$$

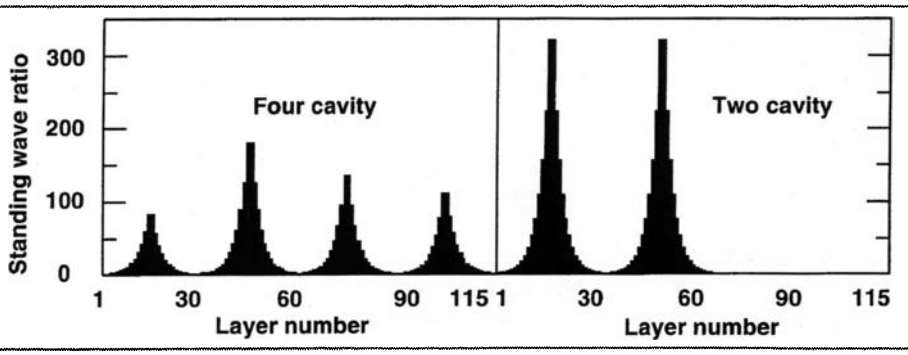


Fig. 7-83 The standing wave ratio at the emergent side of layers in the four-cavity and two-cavity bandpasses cited in Table 7-21.

In other words, $\Delta\beta_r$ is one-half the phase thickness of the passband. The next step is to approximate the arcsin in Eq. (7-48e) with its argument. The right of Eq. (7-28a) is substituted for F_{max}. The result is that the fractional bandwidth B [a dimensionless parameter that is the bandwidth ratio-ed to the center wavelength λ_0] is

$$B = 2\,\Delta\beta_r\ (\pi)^{-1} = 2\ (\pi)^{-1}\ (4\,F_r/V_m)^{1/(2q)} \ . \qquad (7\text{-}49a)$$

Taking the log of both sides of Eq. (7-49a) results in

$$\eta = -2\,q\,|\log(B\,\pi\,/\,2)| + \text{constant} \ . \qquad (7\text{-}49b)$$

This shows that a plot of fractional bandwidth with log abscissa should produce a straight line, with "double the cavity count" (i.e., 2 q) as the negative slope.

Some bandpasses were designed containing three cavities and four cavities with refractive indices for the layers as close to 1.45 and 2.08 as possible. Given the "$T = 50\%$ bandwidth" of each design, the fractional bandwidth was found. When η is plotted vs log of the fractional bandwidth, the linear relationship [see Eq. (7-49b)] is shown in Fig. 7-84.

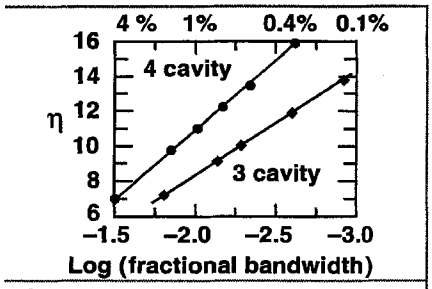


Fig. 7-84 η [see Eq. (7-36a)] is plotted on the *lower abscissa* vs decadic log of the fractional bandwidth [see Eq. (1-9)] and on the *upper abscissa* vs fractional bandwidth for all dielectric bandpasses with three cavities (*lower*) and four cavities (*upper*) in which the layer indices are close to 1.45 and 2.08.

What is the use of such a plot? For one thing, it provides the designer with an idea of the trade-offs between bandwidth and number of layers. It is useful to utilize the argument of §7.9.8.1.2 that η is proportional to the layer count.

For example, suppose the goal is to decrease the bandwidth by a factor of ten and four-cavity bandpasses are used. Inspection of Fig. 7-84 shows that the straight line for the four cavity case increases from $\eta = 7$ to $\eta = 15$ as the fractional bandwidth decreases from 3.2% (log = –1.5) to 0.32% (log = –2.5). The conclusion is that, in order to reduce the bandwidth tenfold, at least double the η is needed and hence the bandpass must contain twice the number of layers.

If a three-cavity bandpasses were used to accomplish the same goal as in the last paragraph, then a smaller increase in the number of layers is needed, due to the smaller slope of the three-cavity line in Fig. 7-84. The number of layers would scale as the ratio 3:4, which are the ratios of the "2q" slopes in Eq. (7-49b) for three-cavity and four-cavity bandpasses.

7.9.9 Bandwidth and spectral slope of a maximally flat prototype

The goal of this section is to compare the spectral bandwidth and slope of the transmittance of prototype maximally flat bandpasses.

7.9.9.1 Spectral bandwidth of the prototype

Given the R/T ratio F in Eq. (7-27c), at a given transmittance level (as for example, –3 dB or –0.5 dB) the R/T is computed, which defines F_r. It follows from Eq. (7-27c)

$$F_r^{1/(2q)} \; F_{max}^{-1/(2q)} = \sin\beta_r \qquad (7\text{-}50a)$$

where β_r is the phase thickness at the aforementioned transmission level. F_{max} appears in Eq. (7-28a). To cite examples, F_r is 1.00 at the 50% transmission level and F_r is 999 at the –30 dB level. The spectral bandwidth (in phase units) of the prototype is found by taking the arcsin of Eq. (7-50a).

7.9.9.1.1 A comparison of two prototype bandpasses

Consider next two prototype bandpasses, each containing a different number of cavities but manifesting the same spectral bandwidth. This implies that the right of Eq. (7-50a) is identical for these bandpasses. If the bandpasses contain $q1$ and $q2$ cavities and the

corresponding "F_{max}" values are F_{max1} and F_{max2}, respectively, then it follows from Eq. (7-50a) that

$$F_{max1}^{-1/(2q1)} = F_{max2}^{-1/(2q2)} \quad . \tag{7-50b}$$

The conclusion is that F_{max} of prototypes with a different cavity count are different and scale according to the foregoing equation. As an example, the prototype bandpass (containing five cavities) whose transmission curve appears in Fig. 7-43 manifests a F_{max1} of nearly 10^6. If a ten-cavity bandpass were constructed that manifests the same spectral bandwidth, Eq. (7-50b) shows that its F_{max2} should equal the square of F_{max1}, or 10^{12}. The effect of this larger F_{max} is to reduce by six decades the minimum offband transmittance, as shown in Fig. 7-85.

7.9.9.2 Spectral slope

In order to find the spectral slope, Eq. (7-27c) is differentiated with respect to the phase thickness β and evaluated at β_1. The result is

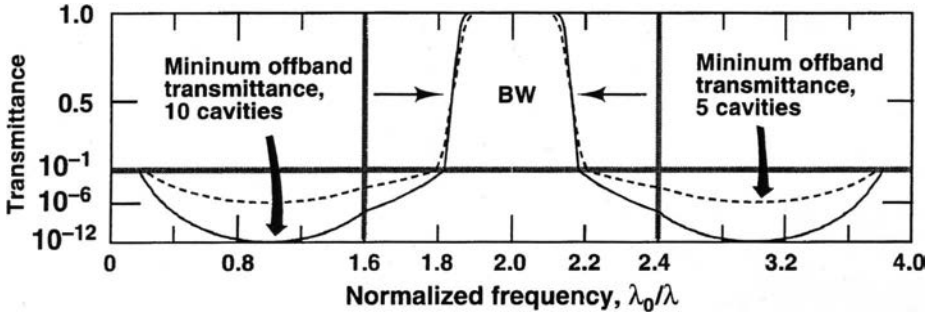


Fig. 7-85 Versus normalized frequency, transmittance of prototype bandpasses. For the five cavity (*dashed*) F_{max} [see Eq. (7-28a)] is 10^6. The F_{max} is 10^{12} for the ten cavity (*solid*). The abscissa changes scale at 0.8 and 1.2. The ordinate scale changes from linear to log at 0.10. The transmittance follows from Eqs. (2-24d) and (7-27c). The phase thickness β in Eq. (7-27c) is 180° at $\lambda_0/\lambda = 2.00$.

$$\frac{dF}{d\beta} = 2q\, F_{max} \cos\beta\ (\sin\beta)^{2q-1} = 2q \cos\beta\, (\sin\beta)^{-1}\, [F_{max}\ (\sin\beta)^{2q}\,] \quad . \tag{7-50c}$$

Equation (7-50e) converts the foregoing to a derivative with respect to transmittance T. The foregoing equation is evaluated at a specific $\beta = \beta_r$, which establishes the spectral bandwidth of the passband. According to Eq. (7-50b), the term in the square brackets equals F_r, which is established by the transmission level at which the bandwidth is measured. The result is

$$\frac{dF}{d\beta} = 2q\, F_r\ \cot\beta_r \quad . \tag{7-50d}$$

The following conclusions may be made about the slope of F (as a function of frequency, which is proportional to β)

(1) The slope is proportional to the number of cavities q.
(2) For a given bandwidth β_r, the slope is independent of V_m .
(3) For β_r values small compared to one, the slope is inversely proportional to the bandwidth, to the extent that $\cot\beta_r$ may be approximated by $(\beta_r)^{-1}$.

To cite some examples, suppose that the slope is evaluated at the –3 dB transmission level. This establishes F_r = 1.00. For a given bandwidth, a ten-cavity bandpass would manifest a slope that is twice as large as that produced by a five-cavity bandpass. The solid curve (for the ten-cavity bandpass) in Fig. 7-85 manifests greater slope than the dashed curve (for the five-cavity bandpass).

A final equation relates the derivative of F to the derivative of transmittance T

$$\frac{dT}{d\beta} = -T^2 \frac{dF}{d\beta} \quad . \tag{7-50e}$$

7.9.10 Design of a bandpass with a large number of cavities

The question is, what is a "large" cavity count? Is a 25 cavity bandpass possible? This is, indeed, possible, with the *caveat* that the number of layers will be quite large. The design procedure is based upon the quarterwave stack. The discussion of this structure begins in §5.1.2.1.1.

7.9.10.1 Design procedure

Given that the number of cavities has been established, the design procedure is:

(1) A starting design, which consists of a prototype bandpass is produced. This is discussed in §7.9.10.3.

(2) The refractive indices of the aforementioned starting design are computer optimized until the desired spectral bandwidth is achieved. An example appears in §7.9.10.5.

(3) The result of step (2) is that the standing wave ratios at the interfaces of the prototype are known. Multilayer reflectors are designed that produce those standing wave ratios. §7.6.5.6.1 and §7.6.5.6.2 cite examples

(4) Each reflector from step (3) is sandwiched between spacers of halfwave optical thickness.

7.9.10.2 Starting design of the prototype bandpass

The starting design of the prototype bandpass resembles a quarterwave stack. That is, with the exception of the outer layers at the beginning and end of the stack, the interior layers have the same optical thickness. All layers alternate between a low refractive index and a high refractive index.

Optimization on a digital computer is subsequently used to adjust the refractive indices of the layers, with the goal of improving the passband transmittance.

7.9.10.3 Stop band and passband of a quarterwave stack

Figure 5-15 illustrates that in the spectral region of the stop band, the reflectance of a quarterwave stack increases monotonically as more **HL** two-layer periods are added. A corollary is that the transmittance monotonically decreases in that spectral region. The spectral region outside of the stop band is the passband of the quarterwave stack bandpass filter. The quarterwave stack is designed so that it achieves the prescribed passband spectral width.

7.9.10.3.1 High-frequency edge of the stop band

Following the notation of §5.2.4.3.1, the symbol g denotes normalized frequency. The layer thicknesses are chosen such that the frequency center of the passband is at g = 2.00

and the stop band is centered at $g = 1.00$. Given the standing wave ratio V (see §5.3.2.1) at the interface between the **H** and **L** layers of the basic period, Eq. (5-5c) is modified to yield the normalized frequency g_1 at the high frequency stop band edge

$$g_1 = 1 + \frac{2}{\pi} \arcsin\left(\frac{V-1}{V+1}\right). \tag{7-50f}$$

7.9.10.3.2 Center of the passband and its bandwidth

The passband is centered at a normalized frequency of 2.00. The semiwidth B of the passband is

$$B = (2 - g_1) = 1 - \frac{2}{\pi} \arcsin\left(\frac{V-1}{V+1}\right). \tag{7-50g}$$

Inversion of the foregoing equation results in

$$V = \frac{1 + \cos(\frac{\pi}{2}\mathrm{BW})}{1 - \cos(\frac{\pi}{2}\mathrm{BW})} \tag{7-50h}$$

where BW is the fractional bandwidth of the passband, as measured to the stop band edges. Similar to the parameter FBW appearing in Eq. (1-9), BW is dimensionless. The V is the standing wave ratio at the interface between the high-index and low-index layers of the quarterwave stack.

7.9.10.3.3 Quarterwave stack — containing conventional coating materials

Consider the spectral width of the stop band of the quarterwave stack in Fig. 5-15. The *conventional coating materials* used in this stack have refractive indices of 2.31 and 1.38. Table 7-23 shows that its index ratio of 1.67:1. The first-order stop band is centered at a normalized frequency of 1.00 and the passband is centered at a frequency of 2.00. Table 7-23 shows that the semiwidth of the passband is slightly in excess of 16% and that the semiwidth of the passband is 1.00 – 0.162 = 0.838 = 83.8%.

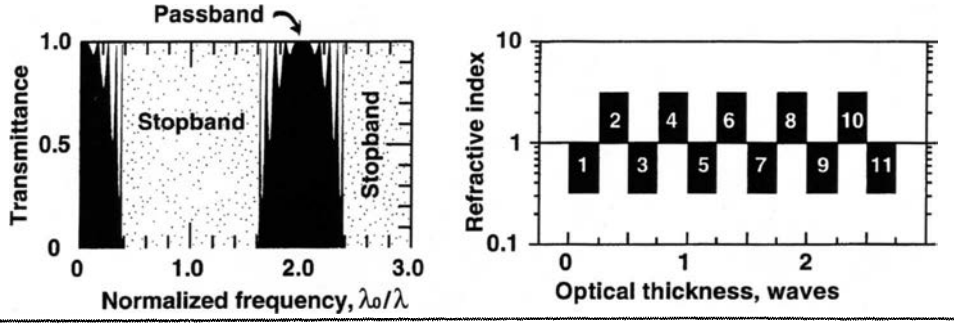


Fig. 7-86 (*Right panel*) Versus optical thickness, refractive indices of an eleven layer prototype bandpass in which the refractive index ratio if 10:1. Each layer has an optical thickness of $\lambda_0/2$ at a wavelength corresponding to a normalized frequency of 2.00. Its transmittance vs normalized frequency appears in the *left panel*. The stop band spectral region is dotted.

7.9.10.3.4 Quarterwave stack with fictitious indices — ten-to-one index ratio

The left panel in Fig. 7-86 depicts the transmittance versus normalized frequency of an eleven-layer quarterwave stack containing fictitious refractive indices. The ratio of the

high index to the low index is 10:1. As compared to the stack cited in §7.9.10.3.3, this higher index ratio broadens considerably the stop band spectral width. As compared to the stack cited in the previous paragraph, the passband semiwidth is shrunk to approximately 40%.

Table 7-23 Quarterwave stacks—spectral properties

Figure	Index ratio	Semiwidth of stop band	Semiwidth of passband
5-15	1.67:1	0.162	0.838
7-86	10:1	0.610	0.390
7-87	1165:1	0.966	0.0373

For quarterwave stacks with the refractive index ratios of the layers cited in the *Index ratio* column and with first-order stop bands centered at a normalized frequency of 1.00, the semiwidth of the stop band and the semiwidth of the passband. Transmittance curves appear in the *Figure* column.

7.9.10.3.5 Quarterwave stack with fictitious indices—1165:1 index ratio

The previous section showed that the use of fictitious indices in the quarterwave stack reduced the bandwidth. Even higher index ratios are used for the fictitious layers in this section. Figure 7-87 depicts the transmittance vs normalized frequency of a quarterwave stack in which the interface standing wave ratio (as defined in §7.6.2.2) is 1165 for all except the outer layers. Table 7-23 shows that the semiwidth of the stop band now exceeds 96% and the semiwidth of the passband has shrunk to less than 4%. As cited in the caption to Fig. 7-87, there are two "antireflection layers" on the exterior of the stack. These are similar to (but not identical to) the antireflection layers cited in §7.4.2. The caption to Fig. 7-87 cites 0.0293 as the refractive index of the low-index component. The refractive index of the antireflection layer is the $\sqrt{0.0293} = 0.171$.

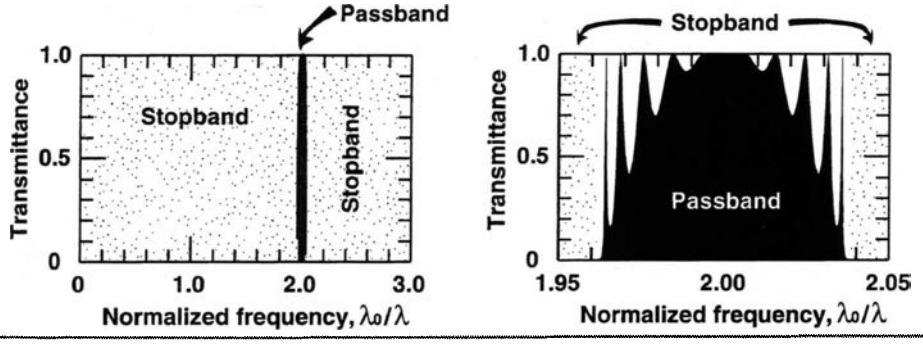


Fig. 7-87 The *right panel* is an expanded view of the passband in the *left panel.* Transmittance vs normalized frequency of an eleven-layer prototype bandpass of the design:

air **A H** (**L H**)^4 **A** *air*

where the refractive indices of **A**, **L** and **H** are 0.171, 0.0293 and 34.12, respectively. Each layer has an optical thickness of $\lambda_0/2$ at a wavelength corresponding to a normalized frequency of 2.00. The stop band spectral region is dotted and the passband is solid black.

7.9.10.4 Starting design for the final prototype bandpass

The goal is a bandpass with a spectral width of approximately 12 nm at a center wavelength of 1552.5 nm, which results in a fractional (i.e., percentage) bandwidth of 0.77 percent. The final bandpass contains third-order spacers and refractive indices of 1.47 and 2.065 are used in the coating. Substituting $m = 3$ into Eq. (7-36e), the inverse *shrink factor* is 5.47 and therefore the prototype bandpass should manifest a spectral bandwidth BW of approximately 4.2 percent. Substitution of this value into Eq. (7-50h) yields a standing wave ratio V for the quarterwave stack *starting design* of 920. Its properties are not distant from those of the 1165:1 stack appearing in Fig. 7-87.

7.9.10.5 Adjustment of the refractive indices of the layers of the prototype

Given the prototype starting design that evolved in §7.9.10.4, a computer code scans its transmittance, as, for example, the curve in Fig. 7-87. Starting at a normalized frequency of 2.00, the code finds the lowest frequency in the passband at which the derivative changes sign. That value determines the trial spectral passband width.

The next step is to establish some transmittance targets in the spectral region of the passband. The transmittance vs frequency should emulate a Chebyshev polynomial, as outlined in §7.10.1. A 0.1 percent ripple in the reflectance in the passband is the objective. The roots of a Chebyshev polynomial (of order 17) are the targets for 100 percent transmittance. Relative to the passband edge, frequencies of roots are 0.1837, 0.3612, 0.5264, 0.6737, 0.7980, 0.8952, 0.9618 and 0.9957. At the lower extrema of the polynomial, the transmittance is targeted at 99.9 percent. This would correspond to relative frequencies of 0.09227, 0.2737, 0.4457, 0.6026, 0.7390, 0.8502, 0.9325 and 0.9830.

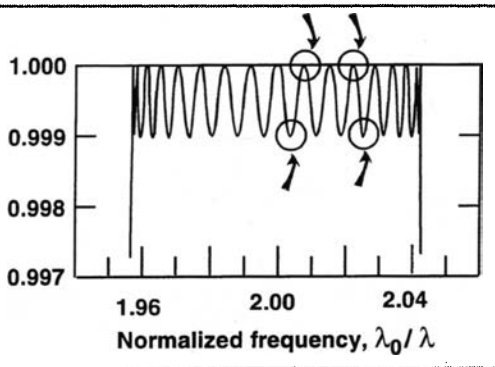


Fig. 7-88 Versus normalized wavelength, transmittance of the design air **A B C D E F G H J H G F E D C B A** air where the refractive indices of **J, H, G, F, E, D, C, B** and **A** are 30.0, 0.0356, 29.85, 0.036, 29.23, 0.0374, 26.85, 0.0467, respectively. Each layer has an optical thickness of a half wavelength at the wavelength corresponding to a frequency of 2.00. The arrowed circles show a few of the optimization targets.

The refractive indices of the starting design are then adjusted by an optimization process, with the objective of achieving the aforementioned passband transmittance. The layer optical thicknesses remain constant. Figure 7-88 depicts the transmittance after a commercial software package [Software Spectra 02] optimized it. The arrows and circles in Fig. 7-88 identify some of the optimization targets.

The optimization is repeated until a prototype of the proper passband width is achieved. The refractive indices of the final prototype appear in Fig. 7-89.

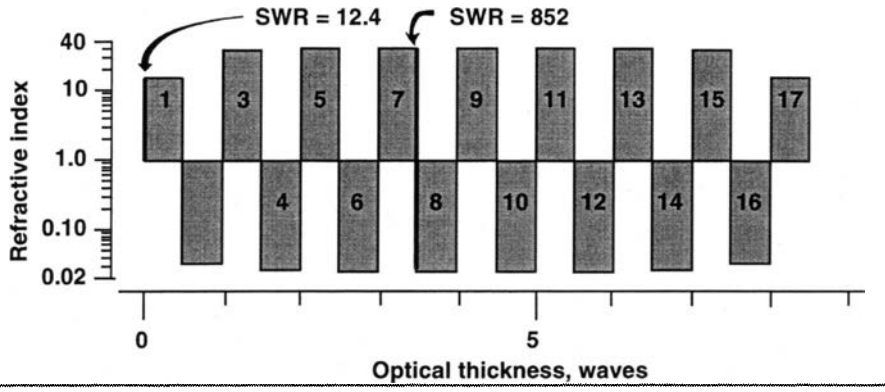


Fig. 7-89 Versus optical thickness (in waves), refractive index (on a logarithmic ordinate scale) of the layers of a seventeen-cavity prototype bandpass.

7.9.10.6 Conversion of the prototype bandpass into a multilayer bandpass

The final step is to convert the prototype bandpass into an actual multilayer bandpass. The refractive indices at the ith interface are known for the prototype. The standing wave ratio V_i at that interface may then be calculated, using equations similar to Eq. (7-31). The task is to produce a multilayer reflector that simulates V_i. This reflector contains layers of quarterwave optical thickness of high refractive index (symbol: **H**) and low refractive index (symbol: **L**). Refractive indices of 2.065 and 1.47 are used in this example. The spacer layers of halfwave (or three halfwaves) optical thickness are low index.

For the multilayer bandpass, *air* is assumed to be both the incident medium and emergent medium. After the refractive index of the actual glass substrate is known, a design correction is made for the substrate.

The first example is the outer reflector. The required standing wave ratio V_0 is 12.36. After that reflector is isolated (see §7.6.5.1), the incident medium is *air* and the emergent medium index is 1.47, i.e., the index of a low-index spacer. A trial design is the following five-layer quarterwave stack reflector

air **E** $(\mathbf{LH})^2$ *low-index-pseudo-substrate*

where **E** is a layer of quarterwave optical thickness (or odd multiple thereof) and the *low-index-pseudo-substrate* manifests an index of 1.47. The standing wave ratio V_0 of this reflector is

$$V_0 = 12.36 = n_E^2 n_L^{-5} n_H^4 (1.0003)^{-1} \qquad (7\text{-}50i)$$

where the index 1.0003 for *air* is retained for pedagogical reasons. Manipulation of Eq. (7-50i) shows that the refractive index of the **E** layer is 2.16. A standard computer code such as TFCalc [Software Spectra 02] enables the user to synthesize an equivalent layer of that index. The procedure in §2.11.3 may also be used. The result is the three-layer sequence 0.544**H** 1.917**L** 0.544**H**.

Figure 7-89 shows that the seventh reflector requires a standing wave ratio V_7 of 852. A trial reflector is

low-index-pseudo-incident-medium $(\mathbf{HL})^4$ **H E H** $(\mathbf{LH})^4$ *low-index-pseudo-substrate*

where **E** is a layer of quarterwave optical thickness. Because low-index spacers are used, both the incident medium and substrate manifest an index of 1.47. The standing wave ratio V_7 of this reflector is

$$V_7 = 852 = n_E^{-2}\, n_L^{-18}\, n_H^{20} \qquad (7\text{-}50j)$$

when the **E** layer refractive index is 1.507. A standard computer code such as TFCalc [Software Spectra 02] may be used to synthesize an equivalent layer, such as the three-layer sequence 0.118**H** 0.751**L** 0.118**H**.

The foregoing procedures are continued until reflecting stacks are designed for all of the nine independent reflectors in the bandpass. As mentioned in §7.6.5.2, there exists a symmetry that reduces the number of independent reflectors. The multilayer design of the final bandpass appears in §7.10.3.8 and its transmittance is graphed in Fig. 7-43. Third-order spacers are used to achieve the required bandwidth. Table 7-11 cites the layer count as 331.

Table 7-24 Spectral requirements of a WDM bandpass

Transmittance level	Spectral width
–0.5 dB	> 0.22 nm
–3.0 dB	< 0.3 nm
–20 dB	< 0.6 nm

Required spectral properties of a 50 GHz bandwidth WDM bandpass filter.

7.9.11 Pulse propagation through a WDM bandpass filter

7.9.11.1 Introduction

The following problem [Thelen 02] was posed at an International Conference on Optical Interference Coatings: Design a bandpass that meets the spectral specifications cited in Table 7-24.

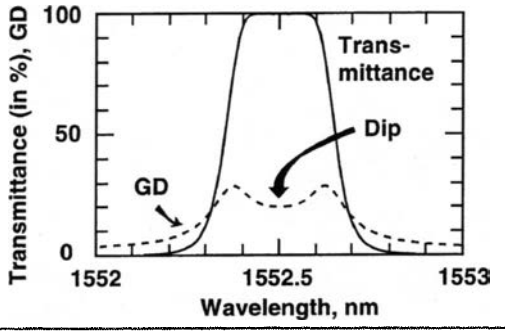


Fig. 7-90 Transmittance (in percent, *thin solid curve*) vs wavelength of a 50 GHz bandpass whose design appears in §7.10.3.9. "GD" is the group delay in ps. Further details appear in the text.
After [Software Spectra 02a].

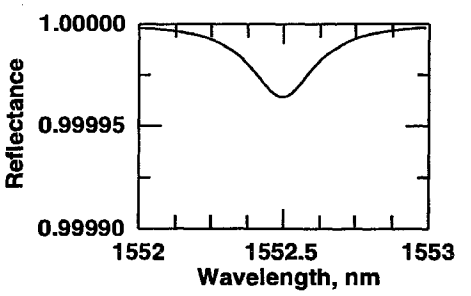


Fig. 7-91 Reflectance versus wavelength of an 88 layer phase correcting coating. Its design appears in §7.10.3.10.
After [Software Spectra 02b].

In this example, the center wavelength of the bandpass is at 1552.5 nm, which is reasonably close to the "conference design problem" wavelength of 1550 nm.

An additional requirement is that the group delay GD [see §1.6.7(g)] in the spectral region of the passband must be relatively flat. The degree of "flatness" depends upon the repetition rate of the pulses that propagate through the bandpass. If the rate is 10 Gb/s, a fluctuation in the GD of 10 ps is tolerable. At a more rapid rate of 40 Gb/s, however, the GD ripple must be less than 2.5 ps.

7.9.11.2 WDM bandpass with 50 GHz bandwidth

Figure 7-90 depicts the transmittance of a bandpass of 50 GHz spectral width. It was designed by A. D. Noe [Software Spectra 02a] and is released as part of the documentation for the computer code TFCalc, version 3.5b11. The bandpass contains five cavities and the spacer layers are of high refractive index.

7.9.11.2.1 Group delay of the WDM bandpass

Although the bandpass meets the spectral requirements cited in Table 7-24, the group delay GD (*dashed curve* in Fig. 7-90) fails to meet the specification cited in §7.9.11.1. The GD clearly manifests a dip at the spectral center of the passband. The GD fluctuates from 20 ps (at the dip minimum) to 28 ps at the spectral edges of the passband. As mentioned in §7.9.11.1, this difference (28 ps – 20 ps) of 8 ps is unacceptably large when pulses at a 40 Gb/s pulse rate propagate through the bandpass. The solution is to combine the bandpass with a reflective phase corrector.

7.9.11.3 Reflective phase corrector

The signal that traverses the bandpass is reflected from a coating that has these attributes:

- Its reflectance should exceed 99.99 percent.
- The phase shift upon reflection should compensate for the dip cited in §7.9.11.2.1. This dip is arrowed in Fig. 7-90.

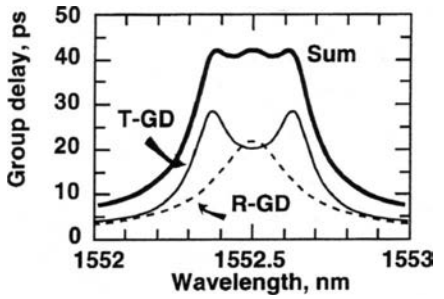


Fig. 7-92 (*Thin solid*, T-GD) Transmissive group delay [see Eq. (1-11e)] of the bandpass whose $T(\lambda)$ curve appears in Fig. 7-90. (*Dashed*, R-GD) Reflective group delay of the phase compensation coating whose $T(\lambda)$ curve appears in Fig. 7-91. (*Heavy solid*) Sum of the transmissive and reflective group delays.

A. D. Noe [Software Spectra 02a] designed this coating with computer optimization. Its reflectance appears in Fig. 7-91 and its group delay (for the reflected signal) appears as the dashed curve in Fig. 7-92. It is evident that the sum of the reflective and transmissive phase shifts produce a group delay GD that meets the ±2.5 ps cited in §7.9.11.1. The corrected phase shift is also depicted as a heavy solid curve in Fig. 7-90.

7.10 Miscellaneous topics and appendices

7.10.1 Design of a prototype bandpass with Chebyshev transmittance in its passband

This section outlines the equations and algorithms used to find the standing wave ratios of the reflectors of a prototype filter. The transmittance in the passband resembles that of a Chebyshev polynomial (see Table 4-5).

7.10.1.1 Statement of the problem

The problem is: Given the number of cavities, the maximum standing wave ratio V_m and allowed ripple (expressed as a percent) of the reflectance in the passband R_x, find the standing wave ratio of the reflectors in the prototype bandpass. In the limit as the ripple of the Chebyshev approaches *zero* passband ripple, it evolves to a *maximally flat* design.

7.10.1.2 Example—five-cavity bandpass

The aforementioned problem is solved for a five-cavity system. The reader can adapt the algorithm to other multiple-cavity systems. The mathematical form of the R/T of the prototype (see §7.6.2) is

$$F = F_{max} \frac{\left[\mathcal{T}_5\left(\frac{\sin\beta}{\sin\beta'}\right)\right]^2}{\left[\mathcal{T}_5\left(\frac{1}{\sin\beta'}\right)\right]^2} \tag{7-51a}$$

where F_{max} is obtained by substituting V_m into Eq. (2-45f). $\mathcal{T}_5(x)$ is a fifth-order Chebyshev polynomial (see Table 4-5) of the first kind

$$\mathcal{T}_5(x) \equiv 16\,x^5 - 20\,x^3 + 5\,x \ . \tag{7-51b}$$

β is the phase thickness of the prototype. β' is the phase thickness at the edge of the passband. This is illustrated in Fig. 7-93, which plots the passband transmittance of the prototype vs β. The passband is centered at a phase thickness of 180°. The ripple in the spectral region of the passband does not drop below T_x until the phase attains β' at its passband edge. β' is 163° in Fig. 7-93. R_x is the reflectance of the ripple, which, because the coating is nonabsorbing, $R_x \equiv 1 - T_x$.

The next step is to take the square root of Eq. (7-51a) and multiply the numerator and denominator by $\sin^5\beta'$. The result is

$$\sqrt{F} = \sqrt{F_{max}}\ \ G^{-1}\ (\ 16\sin^5\beta - 20\sin^3\beta\sin^2\beta' + 5\sin\beta\sin^4\beta') \tag{7-51c}$$

where the normalizing constant G is

$$G \equiv 16 - 20\sin^2\beta' + 5\sin^4\beta' \ . \tag{7-51d}$$

The Kard polynomial for a five layer stack is now written. The layers have equal optical thickness. The incident medium and emergent medium have the same refractive index in a prototype with an odd number of cavities—hence u_0 [see Eq. (2-29c)] is zero. Symmetry (see §7.6.5.2) also requires that $n_1 = n_5$ and $n_2 = n_4$. This means that $u_1 = u_5$ and $u_2 = u_4$, which causes the B_1 term in Eq. (2-51d) to vanish. A polynomial similar to Eq. (2-51c) is written for this five layer prototype. Taking the square root of that equation results in

$$\sqrt{F} = a_1\sin\beta\cos^4\beta - a_3\sin^3\beta\cos^2\beta + a_5\sin^5\beta \ , \tag{7-51e}$$

which reduces to

$$\sqrt{F} = a_1\sin\beta\,(1 - \sin^2\beta)^2 - a_3\sin^3\beta\,(1 - \sin^2\beta) + a_5\sin^5\beta \ . \tag{7-52a}$$

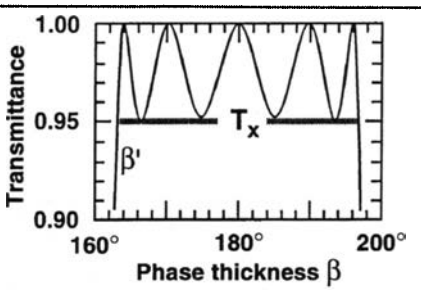


Fig. 7-93 Transmittance vs phase thickness β of the prototype Chebyshev bandpass *air* **A B C B A** *air* where the refractive indices of *air*, **A**, **B** and **C** are 1.00, 4.71, 0.226 and 7.314, respectively. The phase thickness of each layer is the same. T_x is the maximum ripple in the passband transmittance. β' is the phase thickness at the edge of the passband.

The foregoing is equated to Eq. (7-51c) (with $x \equiv \sin\beta/\sin\beta'$) and the coefficients of the terms are matched. The results are

$$a_1 = 5\sqrt{F_{max}}\ G^{-1}\sin^4\beta' \ , \tag{7-52b}$$

$$a_3 = 20\sqrt{F_{max}}\ G^{-1}\sin^2\beta' - 2\,a_1 \tag{7-52c}$$

and

$$a_5 = 16\ \sqrt{F_{max}}\ G^{-1} - a_1 - a_3 \ . \tag{7-52d}$$

7.10.1.3 Kard coefficients in terms of log of the refractive index

The coefficients a_1, a_3 and a_5 are functions of u_1, u_2 and u_3 which, in turn, are related to the indices n_1, n_2 and n_3 using Eq. (2-53b). n_s is 1.00 for a bandpass with an odd number of sections. The coefficients in Eqs. (7-52b) to (7-52d) are

$$a_1 = -\sinh(2\,u_3) - 2\,[\sinh(2\,u_1) + \sinh(2\,u_2)] \ , \tag{7-53a}$$

$$a_3 = 2\,\{\sinh[2\,(-u_1 + u_2 - u_3)] + \sinh(-2\,u_1) + \sinh(-4\,u_1 + 2\,u_2) + \sinh[2\,(-u_1 - u_2 + u_3)]\} + \sinh(-4\,u_1 + 2\,u_3) + \sinh(-4\,u_2 + 2\,u_3) \tag{7-53b}$$

and

$$a_5 = \sinh(-4\,u_1 + 4\,u_2 - 2\,u_3) \ . \tag{7-53c}$$

7.10.1.4 Procedure — solving for the refractive indices

The following procedure is used to solve the foregoing equations for indices n_1, n_2 and n_3:

1. Given the maximum passband ripple R_x, express the ripple in terms of R/T

$$F_x \equiv R_x\,(1 - R_x)^{-1} \ . \tag{7-54a}$$

2. Solve for sinβ'. A property of the qth order Chebyshev polynomial is that $\mathcal{T}_q(1) \equiv 1$, irrespective of q. When $\beta = \beta'$, the numerator of Eq. (7-51a) is 1.00 and (by definition) the left equals F_x. This provides

$$\mathcal{T}_5\left(\frac{1}{\sin\beta'}\right) = \left(\frac{F_{max}}{F_x}\right)^{1/2} \ . \tag{7-54b}$$

Manipulation of Eq. (7-54b) leads to

$$(16 - 20\sin^2\beta' + 5\sin^4\beta')\left(\frac{F_x}{F_{max}}\right)^{1/2} = \sin^5\beta' \ . \tag{7-54c}$$

The foregoing equation is solved recursively, with $\beta' = 0$ for a starting value. That is, given the value of $\sin\beta'$ from a previous iteration, this is substituted into the left. The fifth root of the left is taken and a new value of $\sin\beta'$ is computed. After three or four iterations a stable value of $\sin\beta'$ is obtained. The left of Eq. (7-54c) is the normalizing factor G in Eq. (7-51d).

3. Solve for the values of the coefficients a_1, a_3 and a_5 in Eqs. (7-53a), (7-53b) and (7-53c). These are termed the *target coefficients* in the ensuing discussion.
4. The next task is to find those values of u_1, u_2 and u_3 that satisfy the equations obtained in step (3) above. As a first guess at those index values, the standing wave ratios of a maximally flat bandpass are used (see §7.10.1.8).
5. Given the standing wave ratios in the previous step, the refractive indices of the prototype are found. Figure 7-36 illustrates that these indices alternate between values greater than 1.00 and fractional values. The refractive index of the incident medium and substrate is always 1.00 for any prototype with an odd number of cavities. For the ith layer (for which the standing wave ratio at the emergent side of its interface is V_i), its index n_i is

$$n_i = n_{i-1}\ V_{i-1}^{\pm 1} \qquad (7\text{-}54d)$$

where the exponent is chosen so that the index alternates between values greater than 1.00 and fractional values.

6. The indices in step (5) are converted into the natural log of the index using Eq. (2-29c). These are the *starting values* that are used in the next step.
7. An optimization type of computer code is then used to modify the u values until the target coefficients [see step (3) above] are obtained.
8. Equation (2-53b) is used to convert the u values into refractive indices — which subsequently are translated into standing wave ratios by inversion of Eq. (7-54d).

7.10.1.5 Numerical example — five-cavity bandpass

A five-cavity prototype is designed with an η [see Eq. (7-36a)] of 7 and passband ripple of 5%. F_x [see Eq. (7-54a)] is 0.053. Iterative solution of Eq. (7-54c) yields a $\sin\beta'$ of 0.291. The target coefficients a_1, a_3 and a_5 [step (3) above] are –3.89, –178 and –1588, respectively. According to step (4) above, the standing wave ratios of the reflectors of a maximally flat coating (with $\eta = 7$) are 2.81, 19.2 and 58.6, for V_0, V_1 and V_2, respectively. The refractive indices of the maximally flat prototype are 1.00, 2.81, 0.146, 8.58 and 1.00 for n_0, n_1, n_2, n_3 and n_s respectively. After the optimization is unleashed, the standing wave ratios are 4.71, 20.8 and 32.3 for V_0, V_1 and V_2, respectively. The refractive indices of the prototype appear in the caption to Fig. 7-93.

7.10.1.6 Procedure for a prototype bandpass with an even number of cavities

The foregoing procedure is followed for a bandpass with an even number of cavities, with several exceptions. The refractive index n_s of the emergent medium of the prototype is determined by the passband ripple. Given R_x, this is computed from

$$n_s = [\,1 + \sqrt{R_x}\,]\,[\,1 - \sqrt{R_x}\,]^{-1}\ . \qquad (7\text{-}55a)$$

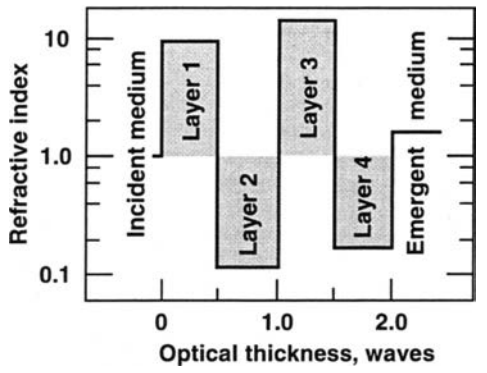


Fig. 7-94 Refractive index vs phase thickness of a four-cavity prototype bandpass with a passband spectral shape that emulates a Chebyshev polynomial. Its η is 7.85 and there is 5.3% reflectance ripple in its passband.

The example consists of a prototype with an η of 7.85 and 5.3% ripple in its passband. The index of n_s is found from Eq. (7-55a)

$$n_s = \frac{1+\sqrt{0.053}}{1-\sqrt{0.053}} = 1.60 \ . \qquad (7\text{-}55b)$$

The indices appear in Fig. 7-94. It is evident that the incident medium and emergent medium have a different refractive index. The symmetry conditions on the refractive indices are different for the even-numbered cavity and odd-numbered cavity systems. For the four-cavity, the constraints upon the indices and log indices are given in Eqs. (2-52e) and (2-52f), respectively.

The R/T polynomial is

$$\sqrt{F} = u_0 \cos^4\beta - a_2 \sin^2\beta \cos^2\beta + a_4 \sin^4\beta \qquad (7\text{-}56a)$$

where

$$a_2 = \sinh(3\,u_0 - 4\,u_1) + \sinh(3\,u_0 - 4\,u_2) + 2\,\{\sinh[3\,u_0 - 2\,(u_1 + u_2)] + \sinh[u_0 - 2\,(u_1 - u_2)]\} \qquad (7\text{-}56b)$$

and

$$a_4 = \sinh[u_0 - 4(u_1 - u_2)] \ . \qquad (7\text{-}56c)$$

The procedure in §7.10.1.4 is used to find the refractive indices of the prototype.

7.10.1.7 Standing wave ratios of the reflectors of a prototype bandpass with a Chebyshev passband transmittance

The first step is to define the log of the ripple_in_percent as

$$z \equiv \log(ripple\_in\_percent) \ . \qquad (7\text{-}57a)$$

The following empirical relationships obtain for the standing wave ratio of the ith isolated reflector

$$V_i = 10^{a + b\,\eta + c\,\eta^2} \ . \qquad (7\text{-}57b)$$

Each of the foregoing constants depends upon the passband ripple, as for example,

$$a \equiv a_0 + a_1\,z + a_2\,z^2 + a_3\,z^3 \ . \qquad (7\text{-}57c)$$

Similar equations exist for the b and c coefficients in Eq. (7-57b). The foregoing equations have been translated into ANSI C language and appear in §7.10.4.

7.10.1.8 Standing wave ratios of the reflectors of a prototype bandpass with a maximally flat passband transmittance

In contrast to the Chebyshev bandpass cited above, these equations are relatively simple and may be found on a hand-held calculator. For a maximally flat bandpass, a plot of the standing wave ratio of a reflector vs η on a semilog scale shows that it may be approximated by an equation similar to Eq. (7-57b), with the c term omitted. For the ith reflector (i = 0, 1, 2, ...), the approximation takes the form

$$V_i \approx B_0\,10^{\kappa\eta} \ . \qquad (7\text{-}57d)$$

Table 7-25 lists B_0 and κ. As an example, for the outer reflector of a seven-cavity maximally flat bandpass when η is 8,

$$V_0 \approx 0.655\ 10^{0.0533 \times 8} = 1.75 \qquad (7\text{-}57d)$$

and is accurate to better than 1%.

There are exact equations for the standing wave ratios of a two-cavity maximally flat bandpass

$$V_0 = V_m^{1/4} = \sqrt{V_1} \qquad (7\text{-}57e)$$

and

$$V_1 = \sqrt{V_m}\ . \qquad (7\text{-}57f)$$

Table 7-25
Constants for the standing wave ratios of the reflectors of a bandpass whose passband transmittance is maximally flat

	Reflector zero		Reflector one		Reflector two		Reflector three		Reflector four	
q	B_0	κ	B_0	κ	B_0	κ	B_0	κ	B_0	κ
3	0.803	0.1661	1.245	0.334						
4	0.696	0.1246	1.000	0.250	2.062	0.2509				
5	0.628	0.0936	0.758	0.199	2.028	0.208				
6	0.624	0.0714	0.645	0.162	1.602	0.179	2.40	0.175		
7	0.655	0.0533	0.566	0.134	1.204	0.159	2.24	0.154		
8	0.707	0.0385	0.532	0.111	0.913	0.143	1.88	0.139	2.40	0.136

Constants B_0 and κ appearing in Eq. (7-57d) vs the number of cavities q.

7.10.2 Shape factor of a prototype WDM bandpass

The shape factor appears in Eq. (1-11d). The goal is find the shape factor of a prototype filter with a maximally flat spectral shape. The starting point is Eq. (7-27c) which expresses the R/T of such a filter containing q cavities. The center of the passband occurs when $m\pi = \beta_0$ in that equation, where m is an integer. The R/T ratio when the transmittance is –0.5 dB is 0.122 and this defines the variable F_r. The phase thickness is $\beta_r \equiv \beta_0 + \Delta\beta_r$, at the –0.5 dB level. Solving Eq. (7-27c) for $\Delta\beta_r$ results in

$$\Delta\beta_r = \arcsin[(F_r/F_{max})^{1/(2q)}]\ . \qquad (7\text{-}58a)$$

F_{max} is usually quite large—a million or greater. In such instances, the arcsin may be approximated by its argument

$$\Delta\beta_r \approx (F_r/F_{max})^{1/(2q)}\ . \qquad (7\text{-}58b)$$

When the transmittance is $10^{-2.5}$, F equals (by definition) F_{25}, which is 315.2. The phase thickness is $\beta_{25} = \beta_0 + \Delta\beta_{25}$. Solving Eq. (7-27c) for $\Delta\beta_{25}$ leads to

$$\Delta\beta_{25} \approx F_{25}^{1/(2q)}\, F_m^{-1/(2q)}\ . \qquad (7\text{-}58c)$$

The *shape factor* [see Eq. (1-11d)] is the ratio of the foregoing $\Delta\beta$ values

$$\textit{shape factor} \equiv \Delta\beta_{25}/\Delta\beta_r \approx (F_{25}/F_r)^{1/(2q)} \qquad (7\text{-}59)$$

and, to the extent that the approximation in Eq. (7-58b) is valid, is independent of maximum standing wave ratio of the prototype filter. The foregoing equation also shows that the shape factor decreases as the number of cavities q increases.

7.10.3 Designs of bandpass filters and other coatings

There are bandpasses whose designs are so verbose that the caption would consume too much space in the text. These designs are below.

7.10.3.1 Bandpass designs for 3, 4, 5 cavities, the $T(\lambda)$ of which appear Fig. 7-41

H and **L** represent layers of optical thickness $\lambda_0/4$ at λ_0 of 1552.5 nm and their refractive indices are 2.065 and 1.47, respectively. The *glass* index is 1.50. For the three-cavity bandpass, the design is

air 0.429**L** 0.135**H** 0.429**L** (**L H**)^{11} 2**L** (**H L**)^{11} 0.202**L** 0.579**H** 0.202**L** **L** (**H L**)^{12} 2**H** (**L H**)^{12} **L** 0.202**L** 0.579**H** 0.202**L** (**L H**)^{11} 2**L** (**H L**)^{11} 0.194**L** 0.596**H** 0.194**L** *glass* ,

for the four-cavity bandpass

air **H** (**L H**)^{7} 2**L** (**H L**)^{17} **H** 2**L** (**H L**)^{18} **H** 2**L** (**H L**)^{17} **H** 2**L** (**H L**)^{8} 0.3**L** 0.38**H** 0.3**L** *glass*

and for the five-cavity bandpass

air 0.172**L** 0.639**H** 0.172**L** **L** (**H L**)^{5} 2**H** (**L H**)^{6} **L** 0.262**L** 0.457**H** 0.262**L** (**L H**)^{7} 2**L** **H** (**L H**)^{14} 2**L** **H** (**L H**)^{14} 2**L** (**H L**)^{7} 0.262**L** 0.457**H** 0.262**L** **L** (**H L**)^{6} 2**H** (**L H**)^{6} **L** 0.348**L** 0.289**H** 0.348**L** *glass* .

7.10.3.2 Designs of WDM bandpasses—definitions of symbols

For the bandpasses in the remainder of §7.10.3.2, **H** and **L** represent layers of optical thickness $\lambda_0/4$ at λ_0 of 1552.5 nm. The refractive indices of *glass*, **H** and **L** are 1.50, 2.065 and 1.47, respectively.

7.10.3.2.1 Bandpass, the $T(\lambda)$ of which appears in Fig. 7-58

The spacer thicknesses are first order for a three-cavity bandpass (*solid*)

air 0.26**L** 0.462**H** 0.26**L** (**L H**)^{8} 2**L** (**H L**)^{18} **H** 2**L** (**H L**)^{18} **H** 2**L** (**H L**)^{9} 0.41**L** 0.17**H** 2.41**L** *glass*

and for the *shaded* curve, the spacer thicknesses are third order

air 0.26**L** 0.462**H** 0.26**L** (**L H**)^{8} 6**L** (**H L**)^{18} **H** 6**L** (**H L**)^{18} **H** 6**L** (**H L**)^{9} 0.41**L** 0.17**H** 2.41**L** *glass* .

The symbols in the design appear in §7.10.3.2.

7.10.3.2.2 Bandpass, the $T(\lambda)$ of which appears in Fig. 7-59

There are three bandpasses, each containing four cavities and first-order thickness spacers. For the *dashed* curve the design is

air **H** (**L H**)^{7} 2**L** (**H L**)^{17} **H** 2**L** (**H L**)^{18} **H** 2**L** (**H L**)^{17} **H** 2**L** (**H L**)^{8} 0.3**L** 0.38**H** 2.3**L** *glass* ,

for the *solid*

air 0.338**L** 0.308**H** 0.338**L** (**L H**)^{8} 2**L** (**H L**)^{18} **H** 2**L** (**H L**)^{19} **H** 2**L** (**H L**)^{18} **H** 2**L** (**H L**)^{8} **H** *glass*

and for the *shaded* curve, the design is

air **H** (**L H**)^{8} 2**L** (**H L**)^{19} **H** 2**L** (**H L**)^{20} **H** 2**L** (**H L**)^{19} **H** 2**L** (**H L**)^{9} 0.25**L** 0.48**H** 2.25**L** *glass* .

The symbols in the design appear in §7.10.3.2.

7.10.3.2.3 Bandpass , the $T(\lambda)$ of which appears in Fig. 7-60

The four-cavity bandpass has second-order spacers

air **H** (**L H**)^{7} 4**L** (**H L**)^{17} **H** 4**L** (**H L**)^{18} **H** 4**L** (**H L**)^{17} **H** 4**L** (**H L**)^{8} 0.3**L** 0.38**H** 2.3**L** *glass* .

7.10.3.2.4 Bandpass, the $T(\lambda)$ of which appears in Fig. 7-61

The design of a five-cavity bandpass (*solid* curve) is

air 0.446**L** 0.568**H** 0.893**L** 0.568**H** 0.446**L** (**L H**)^{7} 2**L** (**H L**)^{17} **H** 2**L** (**H L**)^{18} **H** 2**L** (**H L**)^{18} **H** 2**L** (**H L**)^{17} **H** 2**L** (**H L**)^{8} 0.25**L** 0.48**H** 2.25**L** *glass*

and for the *shaded* curve the design is

air 0.29**L** 0.4**H** 0.29**L** (**L H**)^{8} 2**L** (**H L**)^{18} **H** 2**L** (**H L**)^{19} **H** 2**L** (**H L**)^{19} **H** 2**L** (**H L**)^{18} **H** 2**L** (**H L**)^{9} 0.436**L** 0.122**H** 2.436**L** *glass* .

The symbols in the design appear in §7.10.3.2.

7.10.3.2.5 Bandpass, the $T(\lambda)$ of which appears in Fig. 7-62

The design of a five-cavity bandpass with excessive ripple in the passband transmittance (*solid* curve) is

air 0.363**L** 0.260**H** 0.363**L** (**L H**)^{8} 2**L** (**H L**)^{17} **H** 2**L** (**H L**)^{18} **H** 2**L** (**H L**)^{18} **H** 2**L** (**H L**)^{17} **H** 2**L** (**H L**)^{8} **H** *glass*. .

The design for the *shaded* curve incorporates phase-dispersion narrowing of the passband

air 0.363**L** 0.260**H** 0.363**L** (**L H**)^{8} 2**L** (**H L**)^{16} **H L** 3**H** 4**L** (**H L**)^{18} **H** 2**L** (**H L**)^{18} **H** 4**L** 3**H L** (**H L**)^{16} **H** 2**L** (**H L**)^{8} **H** *glass* .

The symbols in the design appear in §7.10.3.2.

7.10.3.3. Definitions of symbols

For the designs that appear in §7.10.3.3, the refractive indices of *glass*, **L** and **H** are 1.52, 1.45 and 2.15, respectively. Optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 633 nm.

7.10.3.3.1 Design of a six-cavity bandpass, the $T(\lambda)$ of which appears in Fig. 7-57

air 0.25**L** 0.473**H** 2.25**L** (**H L**)^{3} 2**H L** (**H L**)^{4}2**H** (**L H**)^{2} 0.353**H** 0.276**L** 0.353**H** (**H L**)^{3} 2**H** (**L H**)^{3} 0.353**H** 0.276**L** 0.353**H** (**H L**)^{2} 2**H L** (**H L**)^{4} 2**H** (**L H**)^{3} 2**L** 0.576**H** 1.859**L** 0.576**H** *glass* .

7.10.3.3.2 Design of a seven-cavity bandpass, the $T(\lambda)$ of which appears in Fig. 7-57

air 0.416**L** 0.16**H** 2.416**L** **H** 1.344**L** 0.28**H** 1.344**L** **H** 2**L** (**H L**)^{3} **H** 2**L** (**H L**)^{2} 0.257**L** 0.462**H** 0.257**L** (**L H**)^{2} 2**L** (**H L**)^{2} 0.257**L** 0.462**H** 0.257**L** (**L H**)^{2} 2**L** (**H L**)^{3} **H** 2**L** **H** 1.344**L** 0.28**H** 1.344**L** **H** 2.2**L** 0.573**H** 0.2**L** *glass* .

7.10.3.4 Designs of four-cavity bandpasses

Refractive indices of *glass*, **L** and **H** are 1.52, 1.45 and 2.05, respectively. Optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 640 nm.

7.10.3.4.1 Bandpass that is related to Fig. 7-53

The design of the four-cavity bandpass with alternate refractive index spacers is

air 0.233**L** 0.515**H** 0.233**L** **L H** 2**L** (**H L**)^{5} 2**H** (**L H**)^{6} 2**L** (**H L**)^{5} 2**H L H L** 0.172**L** 0.639**H** 0.172**L** *glass* .

7.10.3.4.2 Bandpass, the $T(\lambda)$ of which appears in Fig. 7-54

The design of the four-cavity bandpass (high-index spacers) is

air 0.327**L** 0.329**H** 0.327**L** **L H L** 2**H** (**L H**)5 **L** 2**H** (**L H**)^{6} **L** 2**H** (**L H**)5 **L** 2**H L H L** 0.535**H** 1.933**L** 0.535**H** *glass* .

7.10.3.5 Designs of five-cavity bandpasses in Fig. 7-47

The refractive indices of *glass*, **H** and **L** are 1.52, 2.05 and 1.45, respectively. **H** and **L** represent layers of optical thickness $\lambda_0/4$ at λ_0 of 633 nm.

7.10.3.5.1 Heavy-solid curve in Figs. 7-47 and 7-40

The design is

air 0.542H 1.92L 0.542H 2L (H L)^3 H 2L (H L)^4 H 2L
(H L)^4 H 2L (H L)^3 H 2L H L 0.238L 0.505H 0.238L *glass* .

7.10.3.5.2 Short-dashed curve

air 0.285L 0.412H 0.285L L H 2L (H L)^4 H 2L
(H L)^5 H 2L (H L)^5 H 2L (H L)^4 H 2L H L 0.526H 1.951L 0.526H L *glass* .

7.10.3.5.3 Long-dashed curve

air 0.538H 1.927L 0.538H L H 2L (H L)^5 H 2L (H L)^6
H 2L (H L)^6 H 2L (H L)^5 H 2L H L H L 0.237L 0.508H 0.237L *glass* .

7.10.3.5.4 Light-solid curve

air 0.287L 0.408H 0.287L L H L H 2L (H L)^6 H 2L
(H L)^7 H 2L (H L)^7 H 2L (H L)^6 H 2L H L H L 0.551H 1.903L 0.551H L *glass* .

7.10.3.6 Designs of five-cavity bandpasses related to Fig. 7-32

The refractive indices of *cement, glass*, H and L are 1.52, 1.52, 2.30 and 1.45, respectively. H and L represent layers of optical thickness $\lambda_0/4$ at λ_0 of 633 nm.

7.10.3.6.1 Solid curve (Chebyshev)

cement 0.422L 0.141H 0.422L L H L H 2L (H L)^4 2H L
(H L)^4 2H L (H L)^4 2H (L H)^4 2L H L H L 0.422L 0.141H 0.422L *glass*

7.10.3.6.2 Dashed curve

cement 0.215L 0.537H 0.215L L H 2L (H L)^2 0.361L 0.254H 0.361L
(L H)^2 2L (H L)^2 H 0.39L 0.2H 0.39L (H L)^2 H 2L (H L)^2 H 0.39L 0.2H 0.39L (H L)^2
H 2L (H L)^2 0.361L 0.254H 0.361L (L H)^2 2L H L 0.215L 0.537H 0.215L *glass*

7.10.3.7 Designs of bandpasses with three cavities and seven cavities

The refractive indices of *cement, glass*, H and L are 1.52, 1.52, 1.95 and 1.45, respectively. H and L represent layers of optical thickness $\lambda_0/4$ at λ_0 of 633 nm.

7.10.3.7.1 Three-cavity design

(*Solid* in Figs. 7-42 and 7-44)

cement (H L)^4 2H (L H)^9 2L (H L)^9 2H (L H)^4 *glass* .

7.10.3.7.2 Seven-cavity bandpass with first-order spacers and larger V_m

(*Solid* in Fig. 7-42) The coating contains 123 layers and the spacers are first order

cement (H L)^3 2H L (H L)^7 2H (L H)^9 2L 3H(L H)^9 2L (H L)^9 3H 2L (H L)^9 2H L (H L)^7
2H (L H)^3 *glass* .

7.10.3.7.3 Seven-cavity bandpass with first-order spacers and smaller V_m

(*Solid* in Fig. 7-44) The coating contains 59 layers and the spacers are of first order

cement (H L)^3 2H L (H L)^7 2H (L H)^9 2L (H L)^9 H 2L (H L)^9 H 2L (H L)^9 2H L (H L)^7 2H
(L H)^3 *glass* .

7.10.3.7.4 Seven-cavity bandpass with order 14 spacers and smaller V_m

(*Solid* in Fig. 7-45)

cement 0.528H 1.964L 0.528H 28L H L 0.3L 0.387H 0.3L L H L 28H (L H)^4 28L (H L)^4 H 28L (H L)^4 H 28L (H L)^4 28H L H L 0.3L 0.387H 0.3L L H 28L 0.528H 1.964L 0.528H *glass* .

7.10.3.8 Designs of a nine-cavity bandpass and a seventeen-cavity bandpass

H and L represent layers of optical thickness $\lambda_0/4$ at λ_0 of 1552.5 nm. The refractive indices of *glass*, H and L are 1.50, 2.065 and 1.47, respectively. For the seventeen cavity

air 0.544H 1.917L 0.544H L H L H 6L (H L)^4 0.419L 0.153H 0.419L (L H)^4 6L (H L)^4 0.353L 0.28H 0.353L (L H)^5 6L (H L)^4 H 0.236H 0.51L 0.236H (H L)^4 H 6L (H L)^4 H 0.18H 0.623L 0.18H H (L H)^4 6L (H L)^4 H H (L H)^4 6L (H L)^4 H 0.129H 0.728L 0.129H H (L H)^4 6L (H L)^4 H 0.118H 0.751L 0.118H H (L H)^4 6L (H L)^4 H 0.113H 0.761L 0.113H H (L H)^4 6L (H L)^4 H 0.113H 0.761L 0.113H H (L H)^4 6L (H L)^4 H 0.118H 0.751L 0.118H H (L H)^4 6L (H L)^4 H 0.129H 0.728L 0.129H H (L H)^4 6L (H L)^4 H 0.147H 0.691L 0.147H H (L H)^4 6L (H L)^4 H 0.18H 0.623L 0.18H H (L H)^4 6L (H L)^4 H 0.236H 0.51L 0.236H (H L)^4 H 6L (H L)^5 0.353L 0.28H 0.353L (L H)^4 6L (H L)^4 0.419L 0.153H 0.419L (L H)^4 6L H L H L 0.544H 1.917L 0.544H *air*

and for the nine-cavity bandpass the design is

air 0.363L 0.26H 0.363L (L H)^3 2L (H L)^8 H 2L (H L)^9 H 4L (H L)^{10} 2H (L H)^9 3L H 2L H 3L (H L)^9 2H (L H)^{10} 4L (H L)^9 H 2L (H L)^8 H 2L (H L)^3 H *glass* .

7.10.3.9 Design of bandpass with 50 GHz bandwidth

The design of the 50 GHz bandwidth bandpass whose transmittance appears in Fig. 7-90 is

air 1.271L 0.3968H (L H)^6 L 6H L(H L)^{15} 8H L (H L)^{19} 4H L (H L)^{19} 8H L (H L)^{15} 6H (L H)^6 L *glass*

where *glass* refractive index is 1.52. The optical thickness of L and H is $\lambda_0/4$ at λ_0 of 1552.5 nm and their indices are 1.465 and 2.065, respectively.

7.10.3.10 Reflective phase corrector

The design of this phase correcting coating is

air 269.87L 185.60H 247.13L 174.44H 248.08L 178.04H 253.16L 180.34H 254.13L 179.76H 252.63L 179.26H 254.49L 183.78H 267.11L 1136.98H 268.19L 188.42H 265.03L 187.96H 264.93L 187.95H 264.93L 187.95H 264.93L 187.95H 264.93L 187.95H 264.93L 187.95H 264.93L 187.95H 264.93L 187.95H 264.93L 187.95H 264.93L 187.95H 264.93L 187.95H 264.93L 187.95H 264.93L 187.96H 264.98L 188.10H 265.80L 1506.00H 265.47L 186.74H 261.93L 185.48H 261.47L 185.72H 262.20L 186.32H 263.01L 186.83H 263.61L 187.17H 263.99L 187.38H 264.22L 187.50H 264.35L 187.57H 264.43L 187.62H 264.48L 187.64H 264.51L 187.66H 264.54L 187.69H 264.57L 187.71H 264.61L 187.75H 264.67L 187.80H 264.76L 187.87H 264.87L 187.96H 265.01L 188.07H 265.20L 752.08H *silver*

where the optical constant of the silver is $0.49 - j\,9$. The layer thicknesses are metric, in nanometer. The refractive indices of L and H are 1.465 and 2.065, respectively.

7.10.4 Standing wave ratios of the reflectors of a prototype bandpass with a Chebyshev passband transmittance —the ANSI C source code

The objective of the code is to furnish the standing wave ratios of the two or three independent reflectors of a bandpass . There are two inputs to the code. One is the variable <<lvm>>, which is "log V_m" or "eta", η, that appears in Eq. (7-36a). The range of <<lvm> is greater than 6.0 but less than 30.0.

The other input variable is the "ripple_in_percent"— the percentage ripple in the reflectance in the passband. The range of "ripple_in_percent" is in excess of 0.01 but less than 6.0. As an example, the "ripple_in_percent" in Fig. 7-92 is 5.0.

Table 7-26 ANSI C code for the design of a bandpass containing either 3, 4 or 5 cavities

```
// section one  SWRs of a three cavity BP with Chebyshev response
void x3_cavity(double lvm, double ripple_in_percent, double vee[] )
{ // modified  June 18 2000
double a[]= {-0.035651,  0.053884 ,  0.024172,  0.0049826 };
double b[]= { 0.16648,    1.5592e-5,  2.9109e-6, -5.1513e-8 };
double c[]= { 5.093e-6,  -4.1917e-7, -7.741e-8,  1.7103e-9 };
double sum ;
sum  =   poly3(a, ripple_in_percent )        ;
sum +=   poly3(b, ripple_in_percent )  * lvm ;
sum +=   poly3(c, ripple_in_percent )  * lvm * lvm ;
vee[0]  = pow(10., sum) ;
vee[1]  = pow(10., 0.5 * lvm ) / vee[0] ;
}
```

```
// section two    SWRs of a four cavity BP with Chebyshev response
void x4_cavity(double lvm, double ripple_in_percent, double vee[] )
{
double a0[]={-.0727033664, 0.072206209,  0.0281748793,  0.005590774933};
double b0[]={0.123985688,   0.000228238333, 3.6557419339e-5, 0.0 };
double c0[]={2.7041054063e-5, -6.681978848e-6, -1.0917840107e-6, 0.0 } ;
double a1[]={-0.021741379466, -0.027043493792, -0.015915777224, -0.0037424466469};
double sum ;
sum  =   poly3(a0, ripple_in_percent )        ;
sum +=   poly3(b0, ripple_in_percent )  * lvm ;
sum +=   poly3(c0, ripple_in_percent )  * lvm * lvm;
vee[0] = pow(10., sum) ;
sum  =   poly3(a1, ripple_in_percent ) ;
vee[1] = pow(10., sum + 0.25 * lvm) ;
vee[2] = pow(10.,lvm  ) / ( vee[0] * vee[1] * vee[0] * vee[1] ) ;
}
```

```
// section three    SWRs of a five cavity BP with Chebyshev response
void x5_cavity(double lvm, double ripple_in_percent, double vee[] )
{ //  new data 2000 07 08
double a0upper[]={1.0177757053, 0.088464423249, 0.029914997744, 0.0054691904425 } ;
double b0upper[]={0.099808856285, 4.3221425785e-5, 5.0726486898e-6, 0.0  } ;
double c0upper[]={7.9969879277e-6, -1.8036644152e-6, -2.1281725115e-7, 0.0 } ;
double a0lower[]={1.018773189, 0.088232316023, 0.029878834766, 0.0054647730657, 0.};
double b0lower[]={-0.10037605341,9.8712264099e-5,3.7437757893e-5,1.054569511e-5,0.};
double c0lower[]={0.00057972420175, -0.00013536847058, -1.6404858631e-5, 0.0 } ;
double allong[]={-0.068268162535,-0.0032752337187,-0.0085152967872,  -0.006849652907};
double a1short[]={-0.068300569167,-0.001177425433,-0.0071760518144, -0.001653209928};
double b1[] =  {0.19858993332,    0.00030620330687,  1.9293030139e-5, 0.0 } ;
double c1[] =  {7.0611318622e-5, -1.60937098415e-5, -1.0325778788e-6, 0.0 } ;
double d1[] =  {-1.1377888994e-6, 2.6919356869e-7,   1.7184613607e-8, 0.0 } ;
double sum , x2, transition_lvm = 11.3  ;
x2 = fabs(lvm -  transition_lvm) ;
if(lvm > transition_lvm )
{
 sum  =   poly3(a0upper, ripple_in_percent )        ;
 sum +=   poly3(b0upper, ripple_in_percent )  * x2 ;
 sum +=   poly3(c0upper, ripple_in_percent )  * x2 * x2 ;
}
else
{
 sum  =   poly3(a0lower, ripple_in_percent )        ;
 sum +=   poly3(b0lower, ripple_in_percent )  * x2 ;
 sum +=   poly3(c0lower, ripple_in_percent )  * x2 * x2 ;
}
vee[0] = pow(10., sum) ;
if(ripple_in_percent < 1.00)
 sum  =   poly3(a1short, ripple_in_percent ) ;
else
 sum  =   poly3(a1long, ripple_in_percent ) ;
x2 =  lvm * lvm ;
sum +=   poly1(b1, ripple_in_percent )  * lvm ;
sum +=   poly1(c1, ripple_in_percent )  * x2 ;
sum +=   poly1(d1, ripple_in_percent )  * x2 * lvm ;
vee[1] = pow(10., sum) ;
vee[2] = pow(10.,lvm/2. ) / ( vee[0]*vee[1]) ;
}
double poly1(double a[], double ripple_in_percent)
{
return a[0] + a[1] * log10(ripple_in_percent) ;
}
double poly3(double a[], double ripple_in_percent)
{
double sum, z ; //  third order polynomial
sum = a[0] ;
z = log10(ripple_in_percent) ;
  sum += a[1] * z ;
  if(a[2]) sum += a[2] * z * z ;
  if(a[3]) sum += a[3] * z * z * z ;
return sum ;
}
```

Input and output of the function in section one

In ANSI C language, the function <<x3_cavity>> that returns approximate values for the standing wave ratios V_0 and V_1 of the two reflectors of a three-cavity bandpass with a quasi-Chebyshev passband transmittance.

Input: "lvm" which is equivalent to "eta" [η in Eq. (7-36a)]. The range is $6 \leq \text{lvm} \leq 30$.

Input: "ripple_in_percent" in excess of 0.01 but less than 6.
Output: V[0] and V[1] . The function <<poly3>> appears in Section three below.

Input and output of the function in section two

In ANSI C language, the function <<x4_cavity>> that returns approximate values for the standing wave ratios V_0, V_1 and V_2 of the three reflectors of a four-cavity bandpass with a quasi-Chebyshev passband transmittance.
Input: "lvm" which is equivalent to "eta" [η in Eq. (7-36a)]. The range is $6 \leq \text{lvm} \leq 30$.
Input: "ripple_in_percent" in excess of 0.01 but less than 6.
Output: V[0] , V[1] and V[2] . The function <<poly3>> appears in Section three below.

Input and output of the function in section three

In ANSI C language, the function <<x5_cavity>> that returns approximate values for the standing wave ratios V_0, V_1 and V_2 of the three reflectors of a five-cavity bandpass with a quasi-Chebyshev passband transmittance.
Input: "lvm" which is equivalent to "eta" [η in Eq. (7-36a)]. The range is $6 \leq \text{lvm} \leq 30$.
Input: "ripple_in_percent" in excess of 0.01 but less than 6.
Output: V[0] , V[1] and V[2] .
The functions <<poly1>> , <<poly3>> and <<sum>>are included.

7.10.5 Historical notes

The single-cavity filter was invented by Perot and Fabry at the beginning of the twentieth century. In 1942 Geffcken [42] patented a metal-dielectric bandpass filter. The design of an all-dielectric two-cavity bandpass was first published by A. F. Turner in U.S. Government reports in the early 1950s. A vector diagram was used to design such a bandpass. Smith [58] used partitioning (see §2.6.4) to analyze the performance of a two-cavity bandpass. Thelen [66b] developed the periodic-structure design method and applied it to three-cavity filters. Seeley [61] related optical filters to their microwave equivalents. In his book, Thelen [89] applied the microwave theory of Ribblet [57] to bandpass design. Thelen also clearly understood the effects of phase dispersion upon the narrowing of the spectral bandwidth of the passband, although those exact words do not appear in the text.

Chapter 8
Coatings that contain absorbing layers

8.1 Introduction

8.1.1 Uses of absorbing films

Every effort is made to minimize, if not eliminate, absorption in many of the multilayers described previously. §5.3.10.5 demonstrates, for example, that the performance of the multilayer reflector is degraded by absorption. There are numerous instances in which the absorption of metals and other layers is a welcomed effect. A few examples are:

- Bandpass filter. The absorption of the metallic films in a bandpass filter provides superb off-band attenuation (see Fig. 1-93).
- Absorbing coating. In many instances, the goal is a highly absorbing coating that manifests a small radiant reflectance.
- Selective absorber. This coating is highly absorptive at certain wavelengths and highly reflective at others.

Quite often the absorption of a bandpass filter is not undesirable. For example, if the bandpass filter rejects radiation by reflecting it, the reflected flux is liable to scatter from baffles in the optical system and thence leak through the bandpass filter, especially at high incidence angles. An absorbing coating (such as the type mentioned in §8.5) is coated on each baffle that reduces its reflectances to small values. The performance of the coating is usually superior to the black lacquer often applied to visual optics. The IR reflectance of such coatings is quite small. Another advantage is that they do not outgas in a vacuum, as does the organic lacquer. Selective absorbers are used extensively in photothermal collection systems.

8.1.2 Generic absorbing materials

In this chapter there are many references to a "metal." This "metal" may, indeed, be a metallic layer of aluminum, silver or titanium. There are other possibilities: (1) The "metal" might be an oxide semiconductor such as indium oxide, tin oxide or cadmium stannate. (2) The "metal" could be a semiconductor, such as germanium, that is absorbing because the photon energy exceeds its band gap. (3) The "metal" might be a "dielectric" coating material that manifests absorption in the near IR because of its reststrahlen bands [Turner 65, Martin 66]. Kagarise [61] used a bismuth (a semi-metal) to coat étalon plates in the IR.

8.1.3 Absorption controlled by the standing wave electric field

Suppose that an absorbing layer is immersed in a multilayer stack. The amount of absorption that it produces is determined by its position in the standing wave electric field of the stack. This phenomenon may be understood in terms of the standing wave field in a microwave waveguide. As depicted in Fig. 8-1, a cavity is created by inserting two iris diaphragm reflectors into a waveguide. The spacing between the irises is adjusted so that there is second-order interference in the cavity. The upper panels of Fig. 8-1 show schematically (as a solid black curve) the standing wave electric field that exists in the cavity.

The next step is to insert into the cavity a quite thin absorbing lamina, which may be moved to various positions. Its k value is 0.01 and its metric thickness is one thousandth of the length of the cavity.

The shaded curve in Fig. 8-1 graphs the normalized mean square electric field, as defined in §2.9.6.1. There is a node in the electric field at the center of the cavity. The absorption in the lamina, which is plotted on a logarithmic ordinate, varies by several orders of magnitude, depending upon its position in the standing wave electric field. For example, Fig. 8-1 shows that when the lamina is situated in the *antinode* of the standing wave electric field, the absorption is a maximum and is approximately 0.008. The movement of the lamina to the center of the cavity, where there exists a node of the standing wave electric field, causes the absorption to decrease by a factor of at least 100. Figure 8-42 shows that the placement of metal layers in the nodes of a standing wave electric field causes a reflection filter to be wavelength selective.

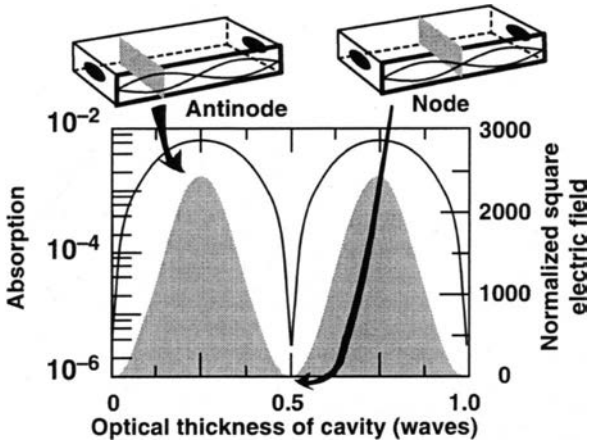


Fig. 8-1 (*Upper* panels) Standing wave electric field (sine wave, *solid*) in a microwave waveguide created by semireflecting iris diaphragms. (Left ordinate, *solid* curve) On a logarithmic ordinate, radiant absorption vs physical position of the absorbing lamina. (Right ordinate, *shaded*) Normalized square electric field [see Eq. (2-63g)] vs position in the cavity. The standing wave ratio of each reflector is 50. The lamina thickness is 0.001 of a wave and its k value is 0.01.

8.1.3.1 Construction of an absorbing coating

The foregoing paragraph suggests how an absorbing coating may be constructed: (a) Create a standing wave electric field; and (2) locate absorbing layers in the antinodes of the standing wave field. This subject is pursued further in §8.5.

8.1.3.2 Construction of a coating with minimum absorption

Contrary to the above, the absorption should be minimized in certain classes of coatings, such as a bandpass filter. The concept of induced transmission (see §8.2.3) may be viewed as a way of locating the node of the standing wave electric field such that the absorption in a metallic layer is minimized. Figures 8-8 and 8-9 illustrate this point.

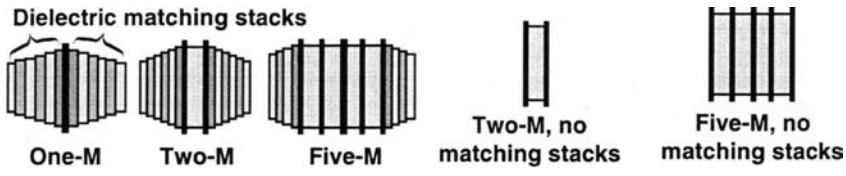


Fig. 8-2 Construction of One-M, Two-M and Five-M filters containing metal layers (*dark*) and dielectric spacer layers (*shaded*). In the left panels the filters are augmented with dielectric matching stacks and in the right they are not.

Table 8-1 Optical properties of bandpass filters

Cnt	Metal	BW(%)	T_{max}	T_{off}	λ_0, μm	Reference
1	S	3.7	0.67	2.5	0.49	Berning [57]
1	A	2.4	0.55	1.7	0.26	Costich [65b]
1	S	2.6	0.70	2**	2.0	Holloway [69]
1	S	10	0.85	1.4	Several*	Gardner [82]
2	S	1.6	0.22	2 - 3¶	0.65	Turner [50]
2	S	2.5	0.44	2 - 3¶	0.56	Turner [50]
2	S	1.0	0.20	2 - 3‡	0.60	Turner [50]
2	A	15	0.14	2 - 3	0.18	Bates [66]
2	A	3.2	0.67	1 - 2§	0.27	Maier [67]
3	A	10	0.25	4 - 6	0.25	Barr [74]
4	A	4	0.10	> 4	0.23	Anders [74]
5	A	10	0.28	3 - 5	0.23	Schroeder [62]

Measured spectral properties of bandpass filters containing layers of aluminum A or silver S. *Cnt* is the number of metallic layers. BW is the fractional spectral width of the passband (centered at λ_0) at 0.5 T_{max}. T_{max} is the maximum passband transmittance and T_{off} is the offband transmittance in absorbance units.
*Six different filters were fabricated, with λ_0 in the range 0.75 μm to 1.6 μm.
**Estimated from a graph in the publication. Transmission leaks occur on the short wavelength side of λ_0.
¶Estimated from thickness of the silver.
‡A second-order spacer produces a transmission leak at wavelengths longer than λ_0.
§Transmission leak of 30% occurs on the long wavelength side of the passband.

8.2 Bandpass filters—general properties

8.2.1 Overview of bandpass filters

This section categorizes bandpass filters and subsequent sections elaborate on design methods. §1.6.4 summarizes the attributes of bandpass filters, which include peak transmittance, offband rejection and bandwidth.

Figure 8-2 shows how bandpass filters containing metallic layers may be constructed. The vertical black lines represent the metallic layers—as few as one and as many as six layers are commonly used. The nonabsorbing (i.e., dielectric) layers between the metallic layers are called *spacers*. Each pair of metallic layers separated by a spacer

is termed a *cavity*. In order to enhance the peak transmittance in the passband, *matching stacks* consisting of dielectric layers are often added to the exterior of the stack.

A bandpass filter with two or more cavities surrounded with *matching stacks* is termed an *augmented filter*. Table 8-1 summarizes the attributes of the filters containing as many as five metal layers.

8.2.2 Trade-off between maximum transmittance T_{max} and offband rejection

Figure 1-95 depicts the attributes of an ideal bandpass filter — with its superb maximum transmittance T_{max} at the center of its passband and offband rejection T_{off} at a reasonably low level of 4×10^{-5}. In many applications, however, this is inadequate — offband transmittances as small as 10^{-12} are needed for UV filters that are solar blind. In this chapter, filters containing metal are designed that provide that offband attenuation.

The T_{off} (see Fig. 1-95) is lowered by adding more metal to the bandpass. An illustration is that a bandpass filter with a total thickness of the metallic layers of 50 nm has an offband transmittance of 3 absorbance units. The addition of 20 nm of metal increases this absorbance to 5. However, the T_{max} in the passband is also lowered as more metal is used. There is always this trade-off in the design of bandpass filters with metallic layers: If too little metal is used, the T_{max} is adequate but the offband rejection is insufficient. The inclusion of a substantial amount of metal provides superb offband rejection, but at the expense of a lowering of the transmittance in the passband.

8.2.3 Simple filters containing one or two metal layers

The chief motivation for the use of metal layers is to provide the offband rejection. The One-M and Two-M filters in §8.2.4 and §8.2.5 are largely of historical interest because they contain an insufficient amount of metal to provide that rejection. Notwithstanding the publications relating to One-M bandpass filter design [Berning 57, Baumeister 69, Lissberger 81], the scientific literature shows that only a few such filters have been actually incorporated in instruments [Gardner 82].

8.2.4 One-M bandpass filter

This bandpass filter contains a single metallic layer surrounded by dielectric stacks. Berning [57] called it an *induced transmission filter*. Because at least two other filters in Fig. 8-2 use the principle of *induced transmission* (see §2.9.4), this book classifies filters by the number of metallic layers, as in the *Cnt* column of Table 8-1.

8.2.4.1 Design procedure

The following design procedure results in a bandpass filter with its passband centered at λ_0. The procedure is simplified by assuming that the incident medium of the assembled bandpass filter is an optical cement of the same refractive index as the substrate.

(1) Select a metal for the bandpass filter. This is discussed further in §8.3.1.

(2) Use the procedure of §2.9.4.2 to compute Ψ_{max} for metal layers M of various thicknesses. A layer of less than 10 nm thickness may form an island film (see §3.3.2.3), in which case its optical constants are different from those of thicker layers. It is advisable that a metal film thinner than 10 nm not be used unless its optical constants have been measured. As explained in §8.2.4.2.2, the Ψ_{max} determined above is essentially the maximum transmittance of the bandpass.

(3) Use the procedures in §2.9.4.3, §2.9.4.4 or §2.9.4.5 to find the amplitude reflection coefficient (or admittance) of the dielectric matching stack that produces Ψ_{max}. The general design of this matching stack is

L' H (L H)^q L" H" *substrate*

if there is a low index adjacent to the metal or

H' (L H)^q L" H" *substrate*

when there is an **H** layer contiguous to the metal layer. **H** and **L** represent quarterwave optical thickness layers at λ_0. **H'**, **H"**, **L'** and **L"** are nonquarterwave thickness layers and q is an integer. The parameter q and the thicknesses of **H"** and **L"** are chosen so that the standing wave ratio that produces Ψ_{max} is attained. The optical thicknesses of **L'** and **H'** are adjusted so that the proper phase matching is achieved, as outlined in §8.2.4.2.1.

(4) Reflect the layers in step (3) about the metal film to form the One-M filter

cement **H" L" (H L)^q H L' M L' H (L H)^q L" H"** *substrate* or

cement **H" L" (H L)^q H' M H' (L H)^q L" H"** *substrate* .

(5) Investigate the effects of thickness errors upon the passband transmittance. The layers **L'** or **H'** that abut the metal are termed the *phase-matching layers*. Thickness errors in these layers quite easily degrade the peak transmittance of the bandpass filter. The designer should simulate errors that may occur in the optical coating process by varying the thicknesses of each phase-matching layer *independently* of the other. As an illustration, suppose that the left **L'** layer is deposited 2% thicker than required and that the right **L'** layer is laid down 1% thinner than its normal thickness. A computer code should be used to predict the degradation of the peak passband transmittance produced by such perturbations.

(6) Abandon the design if the thickness perturbations in step (5) degrade excessively the peak passband transmittance of the passband.

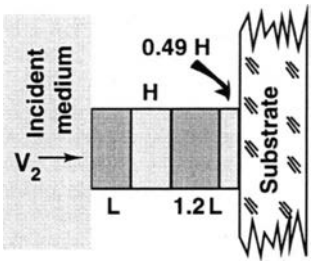


Fig. 8-3 Partial matching stack that produces Ψ_{max} at λ_0 for a metal layer of metric thickness of 30 nm whose $\hat{n}$ is 0.172 – j 4.196. The refractive indices of the incident medium, **H** and **L** are 2.25, 2.25 and 1.35, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$ at λ_0 of 650 nm. The standing wave ratio V is measured at the medium-stack interface.

8.2.4.2 Design example

The data in Table 2-20 are used to produce a bandpass filter with a maximum transmittance at λ_0 = 650 nm. Table 2-20 assumes that there is a high index layer (index 2.25) adjacent to the silver **M** of 30 nm metric thickness. The incident medium of the assembled bandpass filter is an optical cement of the same refractive index (1.52) as the glass substrate. This symmetry simplifies the design procedure.

8.2.4.2.1 Design of the matching stack

The first step is to design a dielectric matching stack that optimizes the net flux ratio Ψ (see §2.9.2) of the silver at λ_0. Table 2-20 lists the standing wave ratio V_2 that the matching stack must produce when the index of the pseudo-incident medium is 2.25 and the substrate is glass. Figure 8-3 shows that V_2 is measured at the interface between the matching layer and the stack on its emergent side.

The two-layer quarterwave stack **L H** manifests a V_2 of 4.11—this is too small. When all of the layers of the stack **L H** $\mathbf{L}_1$ $\mathbf{H}_2$ are quarterwaves, its V_2 of 11.4 is too large. A procedure similar to that of §5.3.2.2 is used to adjust the thickness of the layers $\mathbf{L}_1$ and $\mathbf{H}_1$ so that the *desideratum* V_2 of 8.8 is produced. The remaining task is to insert the high index layer next to the silver — this is the *phase-matching layer.*

Table 2-20 lists its phase thickness as 72.3°, which corresponds to 72.3°/90° = 0.8 of a QWOT. After addition of the silver layer **M**, the *partial* bandpass filter is

M 0.8**H L H** 1.2**L** 0.49**H** *glass* .

For reasons cited in §2.9.4.7, the foregoing design manifests Ψ_{max} at λ_0. Because the caption to Fig. 8-5 defines the metal layer as a metric thickness of 1000 nm, the metal layer appears as "0.03**M**" in that caption.

8.2.4.2.2 Design of the antireflection stack

The objective is a coating that manifests a transmittance T_{max} equaling Ψ_{max} at λ_0. This occurs if the dielectric multilayer antireflection coating reduces the reflectance to zero at that wavelength. That is, the *completed* design is

AR 0.03**M** 0.8**H L H** 1.2**L** 0.49**H** *glass* .

Figure 8-4 illustrates the nomenclature referenced in an appendix (§8.9.1), in which it is argued that this zero reflectance is necessary. The *completed* design (above) is divided into two partitions: the antireflection coating AR and the *partial* bandpass filter. The latter manifests a net flux ratio Ψ_2, which is the same as Ψ_{max}. §2.9.2.1 states that the net flux ratio Ψ_c of the *completed* bandpass filter is the product

$$\Psi_c \equiv \Psi_1 \, \Psi_2 \qquad (8\text{-}1a)$$

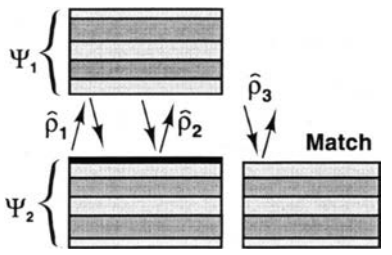


Fig. 8-4 (*Left panel*) The partitions of a One-M bandpass filter consist of a matching stack (including the metal layer — *solid*) and an upper antireflection (AR) stack. $\hat{\rho}_2$ is the amplitude reflection coefficient of the lower partition. The $\hat{\rho}_1$ of the AR is measured at its lower interface. (*Right panel*) $\hat{\rho}_3$ is the amplitude reflection coefficient of the matching stack.

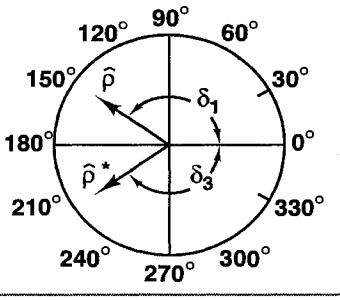


Fig. 8-5 At a wavelength λ_0, phase shift upon reflection δ_3 of the matching stack

medium 0.8**H L H** 1.2**L** 0.49**H** *glass*

and δ_1 of the partial bandpass filter

medium 0.03**M** 0.8**H L H** 1.2**L** 0.49**H** *glass*

where **M** represents a 1000 nm metric thickness layer of optical constant $\hat{n} = 0.172 - j\,4.196$. The refractive indices of *medium*, *glass*, **L** and **H** are 2.25, 1.52, 1.35 and 2.25, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$ at λ_0 of 650 nm.

where Ψ_1 is the net flux ratio of the antireflection AR, which is the other partition. The antireflection coating is a *dielectric* (i.e., nonabsorbing) stack and therefore Ψ_1 = 1.00 (see §2.9.2.2). Combining Eqs. (2-55f) and (8-1a) provides

$$\Psi_c = \Psi_{max} = \frac{T_c}{1 - R_c} \tag{8-1b}$$

where T_c and R_c are, respectively, the reflectance and transmittances of the completed bandpass filter. It follows that $T_c = \Psi_{max}$ if, and only if, the reflectance R_c of the completed bandpass filter is zero.

The final argument is that since the coating is symmetrical, the antireflection coating must, in fact, be identical to the stack on its emergent side. That is, the design of the *final* bandpass filter is

cement 0.49**H** 1.2**L H L** 0.8**H** 0.03**M** 0.8**H L H** 1.2**L** 0.49**H** *glass* .

8.2.4.2.3 Transformation of the admittance

The foregoing design illustrates the transformation of admittance mentioned in §2.9.4.6. This occurs only when the Ψ of the stack is Ψ_{max}. In such cases, the $\hat{\rho}_2$ of the *partial* bandpass filter is the complex conjugate of the $\hat{\rho}_3$ of the matching stack, as shown in Fig. 8-4. A formal proof appears in an appendix—§8.9.1. The aforementioned complex conjugate phase shifts are plotted in Fig. 8-5.

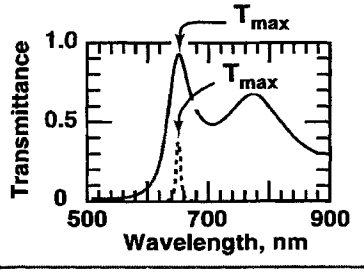


Fig. 8-6 $T(\lambda)$ of the One-M filters (*dashed*)
glass 0.226**H** 0.271**L** (**H L**)^3 **H** 1.81**L** 0.08**M** 1.81**L** (**H L**)^3 **H** 0.271**L** 0.226**H** *glass*
and (*solid*)
glass 0.5**H** 1.24**L H L** 0.8**H** 0.03**M** 0.8**H L H** 1.24**L** 0.5**H** *glass*
where **M** represents a layer of optical constant $\hat{n} = 0.172 - j\,4.196$ and 1.0 μm metric thickness. The refractive indices of *glass*, **L** and **H** are 1.52, 1.35 and 2.25, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$ at λ_0 of 650 nm.

8.2.4.3 Comparison of One-M bandpass filter designs

Figure 8-6 depicts the transmittance of the One-M bandpass filter designed in §8.2.3. Table 8-2 shows that its T_{max} is close to the Ψ_{max} in Table 2-20. The problem is that, although the T_{max} is in excess of 90%, a passband does not truly exist because of the relatively high transmittance on the long wavelength side of λ_0.

Berning [57] avoided this problem by using a thicker **M** layer. Table 8-2 lists the parameters used in the design of a One-M bandpass filter that incorporates an **M** layer of 80 nm metric thickness. Figure 8-6 shows that a passband truly exists and that the offband rejection is superb. Given that the user is willing to live with the "less-than-excellent" peak passband transmittance T_{max} of 37%, why not produce this bandpass filter?

The execution of step (5) in §8.2.4.1 answers this question. T_{max} decreases from 37% to 31% if a 1% thickness error is made in *just one* of the **L** matching layers adjacent to the silver. This strongly suggests that the yield in the production of such a bandpass filter would be unacceptably low. It is for this reason that §8.3 and §8.4 focus on the multiple-cavity bandpass filter, such as the **MDMDM**.

Table 8-2 Optical properties of One-M filters

Design of matching stack	h	Ψ_{max}	n_1	T	R %	A
0.8**H L H** 1.24**L** 0.5**H** *glass*	30	0.926	2.25	0.924	0.15	0.074
1.81**L H L H L H L H** 0.271**L** 0.226**H** *glass*	80	0.377	1.35	0.368	0.98	0.622

At λ_0 of 650 nm, Ψ_{max} (see §2.9.4) of a layer of optical constant $\hat{n} = 0.172 - j\ 4.196$ of metric thickness (in nm) shown in the h column. n_1 is the refractive index of the layer contiguous to both sides of the metal layer. The transmittance T, reflectance R and absorption A occur at λ_0 when the metal layer is sandwiched between the matching stack and the matching stack in reverse order. **L** and **H** represent layers of refractive index 1.35 and 2.25, respectively, of optical thickness $\lambda_0/4$. The *glass* index is 1.52.

Table 8-2 shows that the One-M bandpass filter with the 80 nm silver layer uses an **L** matching layer. Whereas the **H** matching layer used in §8.2.4.2.1 is less than a quarterwave at λ_0, the optical thickness of the **L** layer must be greater than a quarterwave in order to obtain the proper admittance (see §2.9.4.5) so that Ψ_{max} occurs at λ_0.

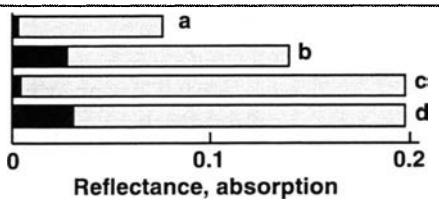


Fig. 8-7 At a wavelength λ_0 of 650 nm, reflectance (*black*) and absorption (*shaded*) of the One-M filters whose designs appear in Table 8-3.

8.2.4.4 Advantages of symmetrical matching stacks

There are two matching stacks contiguous to the metal layer. The absorption is minimized when they are essentially identical. Macdonald [71] suggested that a One-M bandpass filter may be designed by depositing antireflection coatings on both sides of the metal layer. Although this approach is valid, the absorption increases markedly if the *symmetry* of these antireflection coatings is not preserved. This is best illustrated by the One-M filters in Table 8-3. The first (a) has five-layer matching stacks and is the "nearly perfectly matched" One-M bandpass filter in §8.2.4.2. Figure 8-7 shows that its absorption is minimal—7.4%.

Table 8-3 Optical properties of silver-dielectric stacks

Design	Ref.	R %	R' %	A %	A' %	T %
0.49**H** 1.24**L H L** 0.8**H M** 0.8**H L H** 1.24**L** 0.49**H**	a	0.18	0.18	7.4	7.4	92.4
0.156**L** 0.654**H** 1.156**L** 0.73**H M** 0.73**H** 1.156**L** 0.654**H** 0.156**L**	b	2.9	2.9	11	11	86.1
0.7**H M** 0.65**H L H** 1.208**L** 0.546**H** 0.208**L**	c	3.0	0.06	16.8	19.8	80.2

At λ_0 of 650 nm, transmittance T, obverse reflectance R, reverse reflectance R', obverse absorption A and reverse absorption A' of the bandpass filters in the "Design" column. **M** represents a 30 nm metric thick layer of optical constant $\hat{n} = 0.172 - j\ 4.196$. **L** and **H** represent layers of refractive index 1.35 and 2.25, respectively, of optical thickness $\lambda_0/4$.

The matching stacks of a One-M bandpass filter (entry *b* in Table 8-3) not only contain fewer layers but, in addition, the layers are nonquarterwave in thicknesses. The result is that the silver layer is "undermatched" because the standing wave ratios of the

matching stacks are not sufficiently large to attain Ψ_{max}. This is manifested as residual reflectance of 2.9% at λ_0, as listed in Table 8-3. Figure 8-7 shows that its absorption is nearly 50% greater than that of the first entry in Table 8-3.

Another One-M bandpass filter (entry *c* in Table 8-3) also manifests a reflectance of close to 2.9%, but on only its *reverse* side. This occurs because the design is highly asymmetrical. There are six dielectric layers on one side of the silver and only one layer on its other side. Figure 8-7 illustrates that its absorption is at least double that of the first entry in Table 8-3.

The conclusion is that optimal transmittance occurs if the admittances of the dielectric matching stacks at the two sides of the metal layer are nearly identical. In other words, an observer sits on top of the metal layer and determines the amplitude reflection coefficient $\hat{\rho}_1$ of the matching stack that lies above it (see Fig. 8-4). This is compared to $\hat{\rho}_3$ of the matching stack that lies below it. The parameters $\hat{\rho}_3$ and $\hat{\rho}_1$ should be nearly the same.

8.2.4.5 Standing wave electric fields

Following Lissberger [81], Fig. 8-8 graphs the normalized square electric field (see §2.9.6.2) of a One-M bandpass filter. This is the symmetrical coating. It matches the metal layer "nearly perfectly." There is a node of the standing wave in the center of the metal layer.

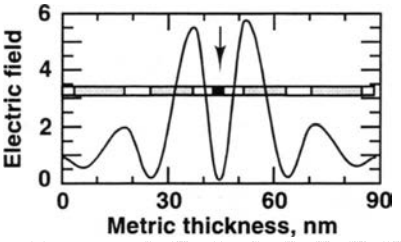


Fig. 8-8 At a wavelength of 650 nm, normalized square electric field [see Eq. (2-63g)] vs metric thickness of the One-M bandpass filter "a" in the "Ref." column of Table 8-3. The arrow designates the silver layer.

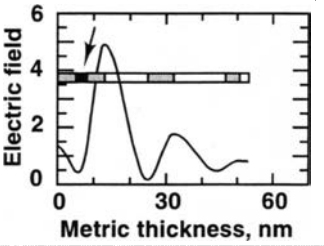


Fig. 8-9 The caption to Fig. 8-8 obtains, except that the coating is "c" in the "Ref." column of Table 8-3.

Table 8-4 Maximum net flux ratio of a single film

Tkns	Ψ_{max}	Tkns	Ψ_{max}	Tkns	Ψ_{max}
15.0	0.982	30.0	0.926	45.0	0.821
17.5	0.976	32.5	0.912	47.5	0.798
20.0	0.968	35.0	0.897	50.0	0.774
22.5	0.959	37.5	0.880	70.0	0.523
25.0	0.949	40.0	0.862	80.0	0.377
27.5	0.938	42.5	0.842	90.0	0.241

At a wavelength of 650 nm, Ψ_{max} of a single silver film of the indicated metric thickness (*Tkns* column, in nm) and optical constant $\hat{n} = 0.171 - j4.196$.

Figure 8-9 plots the electric field in the asymmetrical coating that is the final entry in Table 8-3. Although there is a node of the standing wave in the metal layer, it is at a higher level than in the coating mentioned in the previous paragraph. This is better illustrated in Fig. 8-10, which shows in detail the electric field in the interior of the silver layer for the three designs in Table 8-3. An interpretation of Eq. (2-63a) is that the absorption is proportional to the area under the electric field vs distance curve, as shown by the shaded region in Fig. 8-10. The dashed curves represent the asymmetric coatings in Table 8-3. It is patent that the area, and hence the absorption, is greater for such coatings.

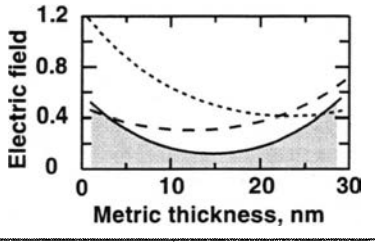


Fig. 8-10 Within the silver layers at a wavelength of 650 nm, normalized square electric field [see Eq. (2-63g)] vs metric thickness of the One-M filters appearing in the "Ref." column of Table 8-3 as "a" (*solid*), "b" (*long dash*) and "c" (*short dash*).

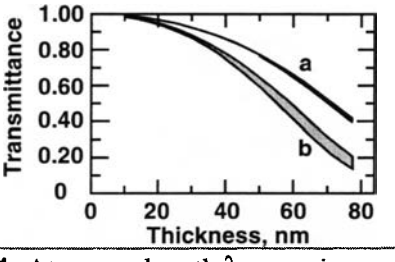


Fig. 8-11 At a wavelength λ_0, maximum transmittance of a One-M filter
matching **H' M H"** *matching*
vs metric thickness of the **M** layer of optical constant $\hat{n}$. Upper curves:
(a) λ_0 is 650 nm and $\hat{n}$ is $0.172 - j4.196$;
(b) $\hat{n}$ is $0.48 - j9.8$ at λ_0 of 1330 nm.
Lower curves: The thickness of the **H'** layer is 1% too large and the thickness of the **H"** layer is 1% too small.

8.2.4.6 Effects of thickness errors on Ψ_{max}

For thicknesses of the **M** silver layer ranging from 10 nm to 80 nm, Fig. 8-11 graphs the peak transmittances of the One-M bandpass filter

stack **H' M H"** *stack*

where the dielectric *matching stacks* (shown above as *stack*) provided the standing wave ratio *V* required to achieve Ψ_{max}. Table 8-4 also lists the Ψ_{max} values.

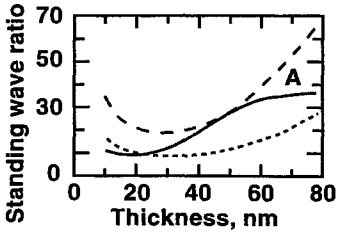


Fig. 8-12 At a wavelength λ_0, standing wave ratio V needed to produce Ψ_{max} for the designs
M H' *matching*
vs metric thickness of the M layer of optical constant $\hat{n}$. V is measured at the interface between the H' layer (of index n_H) and the matching stack. (*Short dash*) λ_0 is 650 nm, n_H is 2.25 and $\hat{n}$ is $0.172 - j\,4.196$. (*Long dash*) λ_0 is 1330 nm, n_H is 2.25 and $\hat{n}$ is $0.48 - j\,9.8$. (*Solid*) λ_0 is 260 nm, n_H is 1.50 and $\hat{n}$ is $0.19 - j\,2.85$.

Curve *a* in Fig. 8-11 is at 650 nm and *b* is for 1330 nm in the near IR. This Fig. also shows the effects of a thickness error in the H' layer. The lower transmittance occurs when there are thickness errors — the H' layer is 1% too thick and H" is 1% too thin. At 650 nm the effects of such errors are not significant, but for thicker layers in the IR, a 4% to 5% transmission loss results. This thickness error is mentioned in step (5) of §8.2.4.1.

Suppose that the One-M bandpass filter designed in §8.2.4.2.2 is subjected to such thickness tolerancing. The H' layer adjacent to the metal in the "correct" matching stack has a thickness of 0.8 QWOT. An H' layer that is 1% too thin has a thickness of 0.792 QWOT. The layer that is 1% too thick is 0.808 QWOT. In this instance, the peak transmittance is reduced by a quite small amount — from 0.9255 to 0.9233 — by such thickness errors.

8.2.4.7 Standing wave ratio needed to achieve Ψ_{max}

The standing wave ratio V of the matching stack (see §8.2.4.6) is measured at the interface between the remainder of the matching stack and its top layer, which is the H' layer in this example.

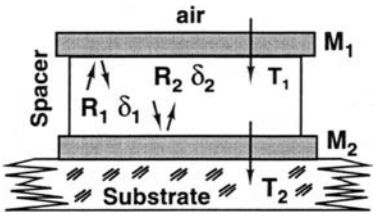


Fig. 8-13 A partitioned MDM consists of the metal layers M_1 and M_2 separated by a dielectric spacer. R_2 and δ_2 are, respectively, the radiant reflectance and phase shift upon reflection of M_2, as measured from inside the spacer. R_1 and δ_1 are the corresponding parameters for M_1, using air as its emergent medium. T_1 and T_2 are the radiant transmittances from the spacer into the surrounding media.

Suppose that the V that produces Ψ_{max} is plotted vs thickness of the M layer, as in Fig. 8-12. The curves are for silver at 650 nm and 1330 nm, as well as for aluminum at 254 nm. In each instance the curve drops to a minimum. This implies that V is a double-valued function of the layer thickness.

8.2.5 Properties of the Two-M bandpass filter

8.2.5.1 Design of the bandpass filter

This single-cavity bandpass filter is often called a *Fabry-Perot* for historical reasons. The coating whose $T(\lambda)$ is depicted in Fig. 1-80 is a simple Two-M bandpass filter containing silver layers. Following the procedure in §7.2, §8.2.5.4 lists methods of altering the width of the passband by changing the composition and thickness of the spacer. This is also accomplished by adjusting the reflectances of the reflectors.

8.2.5.2 Basic parameters

The optical properties of the Two-M bandpass filter are analyzed by using the method of *partitioning* (see §2.6.4.2 and §2.11.4). The coating contains a dielectric spacer layer S and two metal layers M_1 and M_2

air M_1 S M_2 *glass* .

Although the metal layers have the same thickness, one is bounded by air and the other by the glass substrate—this introduces asymmetry into the bandpass filter. Figure 8-13 shows an exploded view of such a bandpass filter. This Fig. elaborates upon Fig. 2-67 and illustrates at what positions in the stack the radiant reflectances R_1 and R_2 [see Eq. (2-45a)] and phase shifts δ_1 and δ_2 [see Eq. (2-44h)] are measured.

The spacer is cryolite (n = 1.35) in this example. T_1 is the transmittance from air through the silver film M_1 into the spacer. In other words, the emergent medium index is 1.35. T_2 is the transmittance of M_2 from the spacer into the substrate. Here, the incident medium index is 1.35 and the substrate of index is 1.52. R_1 and δ_1 are, respectively, the radiant reflectance and phase shift upon reflection for film M_1 using the spacer as an incident medium and air as the emergent medium. R_2 and δ_2 are the corresponding quantities for M_2 with an incident medium index of 1.35 and a glass substrate. Table 8-5 lists the foregoing parameters for metal layer thicknesses of 13 nm and 25 nm. Spacer layers of both high and low index are used.

8.2.5.3 Spacer thickness

The phase thickness β_c of a first-order spacer is

$$\beta_c = \delta_m \qquad (8\text{-}2a)$$

where the mean phase shift upon reflection δ_m appears in Table 8-5 and Eq. (2-44h).

As an example, consider the Two-M bandpass filter with a low index spacer layer and 13 nm thick silver layers. Table 8-5 shows that its mean phase shift is 125.5°. From Eq. (8-2a), the optical thickness of the spacer is 125.5°/360 λ_0 = 0.35 λ_0, which is equivalent to 0.35/0.25 = 1.39 QWOT. The design is

air M 1.39L M *glass*

where L represents a QWOT — a layer of optical thickness $\lambda_0/4$.

The foregoing first-order spacer is converted to a higher order by adding to it an optical thickness of $\lambda_0/2$, λ_0, $3\lambda_0/2$, and so on, where λ_0 is the wavelength of the passband center. A halfwave optical thickness is added in this example. The resulting design is

air M LL 1.39L M *glass* or *air* M 3.39L M *glass* .

Table 8-5 Optical properties of Two-M filters

	Spacer	$n = 1.35$	Spacer	$n = 2.25$	
Parameter	$h = 13$	$h = 25$	$h = 13$	$h = 25$	Augm't'd
R_1	0.491	0.783	0.437	0.720	0.655
R_2	0.423	0.743	0.343	0.667	0.600
R_m	0.456	0.763	0.387	0.693	0.632
T_1	0.424	0.144	0.469	0.186	0.288
T_2	0.510	0.192	0.580	0.249	0.354
δ_1	122°	137°	92°	113°	0°
δ_2	129°	139°	99°	115°	359.4°
δ_m	125.5°	138°	95.5°	114°	
T_{max}	0.731	0.492	0.724	0.491	0.753
$n_c\, h_c$	0.35	0.38	0.27	0.32	
Q	0.8	2.3	0.65	1.7	

At λ_0 of 650 nm, properties of Two-M bandpass filters containing silver layers ($\hat{n} = 0.172 - j4.196$) of metric thicknesses $h = 13$ nm and $h = 25$ nm for spacer refractive indices of 1.35 and 2.25. R_1, R_2, δ_1, δ_2, T_1, T_2, T_{max} and δ_m appear in Eqs. (2-44f) to (2-45c) and are explained in the text. $n_c\, h_c$ is the optical thickness of the spacer layer in waves at λ_0. Q is the resolution — neglecting the phase dispersion of the silver.

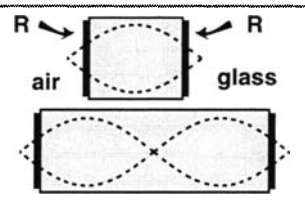


Fig. 8-14 For a Two-M bandpass filter that transmits at $\lambda_0 = 650$ nm
air **M L' M** *glass*
the standing wave electric field at λ_0 is superimposed on the optical thickness of the coating. (*Upper*) First-order spacer. (*Lower*) Second-order spacer. *R* represents the reflector at each end of the spacer.

Figure 8-14 depicts a superposition of a vacuum wavelength (i.e., standing wave) vs the *optical* thickness of bandpass filters with first-order and second-order spacers. The node of the standing wave electric field is well below the surface of the metal layer. This is related to the phase shift upon reflection, as mentioned in §2.11.4.1.2. Figure 2-99 shows that the node of the standing wave is at the surface of the cavity wall when the phase shift upon reflection is 180° — as contrasted with the phase shift of approximately 122° (see Table 8-5) in this example. This causes the node of the standing wave to shift below the surface of the metal layer — similar to that shown in the center panel of Fig. 2-99.

Table 8-6 Optical properties of silver

λ	*n*	*k*	λ	*n*	*k*	λ	*n*	*k*	λ	*n*	*k*
4000	1.65	25.9	1180	0.37	8.53	690	0.163	4.56	490	0.163	2.76
2860	1.32	20.5	1050	0.282	7.50	650	0.172	4.20	470	0.121	2.51
2220	1.04	16.6	950	0.223	6.67	610	0.183	3.87	440	0.100	2.26
1820	0.82	13.7	870	0.185	6.00	570	0.192	3.57	430	0.100	2.02
1540	0.63	11.5	800	0.165	5.44	540	0.194	3.29	410	0.100	1.78
1330	0.48	9.8	740	0.159	4.97	510	0.186	3.02	390	0.100	1.55

Optical constants *n* and *k* of silver, where λ is the wavelength in nm.

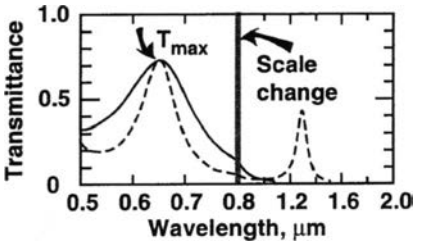


Fig. 8-15 $T(\lambda)$ of Two-M filters
air **M L' M** *glass*
where the indices of **L'** and *glass* are 1.35 and 1.52, respectively. **M** represents a layer of 13 nm metric thickness whose optical constants are in Table 8-6. The optical thickness of **L'** is (*solid*) 1.397 $\lambda_0/4$ and (*dashed*) 3.397 $\lambda_0/4$ at λ_0 of 650 nm.

8.2.5.4 Control of the spectral bandwidth

§7.2 and §2.11.4.2 list three methods of controlling the bandwidth of a single-cavity bandpass filter: (1) Adjustment of the spacer thickness; (2) modification of the reflectances of the reflectors; and (3) control of the phase dispersion. Two of these methods are applied to the single-cavity bandpass filter containing metal layers.

8.2.5.4.1 Thickness of the spacer

Bandpass filters containing first-order and second-order spacers are designed in §8.2.5.3. Fig. 8-15 depicts the transmittances of the Two-M bandpass filters

air **M L' L' M** *glass* and

air **M L' 2L L' M** *glass*

where **2L** represents a halfwave optical thickness layer at λ_0. The second-order spacer reduces the bandwidth. Figure 2-97 illustrates that the higher order spacer introduces passbands on the low frequency side of the main passband. Such a passband is evident at 1.35 μm in Fig. 8-15.

The $T(\lambda)$ in Fig. 8-15 is computed dispersively using the optical constants in Table 8-6. They are close to published data [Hass 72] and have been smoothed by a polynomial fit. If the same optical constant were used at all wavelengths, the transmittance in the IR would exceed that of the visible, rather than decreasing at longer wavelengths.

8.2.5.4.2 Reflectances of the reflectors

Consider first the Two-M bandpass filter. The reflectance is measured from inside the spacer (see Fig. 8-13). Table 8-5 shows that the reflectance $R_1 = 49.1\%$ of

low index **M** *glass*

is higher than the $R_1 = 43.7\%$ of

high index **M** *glass*

where *low index* and *high index* represent massive media of the same index as the low index and high index spacers, respectively. This higher reflectance increases the resolution Q—i.e., produces a narrower bandwidth. Although Fig. 8-16 shows that the narrowing is not immense, it becomes much more pronounced when low index spacers are used with many metal layers, as illustrated in Fig. 8-27.

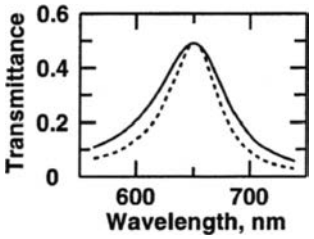


Fig. 8-16 $T(\lambda)$ of a Two-M bandpass filter (*dashed*)
air **M** 1.533**L M** *glass*
and (*solid*)
air **M** 1.267**H M** *glass*
where the indices of **L**, **H** and *glass* are 1.35, 2.25 and 1.52, respectively. **M** represents a layer of 25 nm metric thickness whose optical constants are in Table 8-6. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 650 nm.

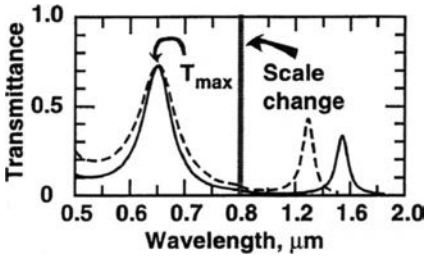


Fig. 8-17 $T(\lambda)$ of the **MDM** (*dashed*)
air **M L' L" M** *glass*
and (*solid*)
air **M L' H H L" M** *glass*
where the indices of **L"**, **L'**, **H** and *glass* are 1.35, 1.35, 2.25 and 1.52, respectively. **M** represents a layer of 13 nm metric thickness whose optical constants are in Table 8-6. The optical thickness of **L'**, **L"** and **H** are 0.676 $\lambda_0/4$, 0.721 $\lambda_0/4$ and $\lambda_0/4$, respectively, at λ_0 of 650 nm.

§5.7.1 describes a metal overcoated with two or more dielectric layers that enhance its reflectance. The same technique increases the reflectance of the semitransparent metallic layers in the reflectors of the bandpass filter [Turner 50]. The reflectance R_1 is increased by adding a quarterwave layer inside the spacer. The reflectance of

high index **L' M** *glass*

is 65.5%, as listed in the "Augm't'd" column of Table 8-5. Figure 8-17 shows that its bandwidth is narrower than the second-order bandpass filter of §8.2.5.4.1. At longer wavelengths in the IR—near 1.5 μm—the increased thickness of the dielectric layers between the **M** layers functions as a pseudo-spacer. A first-order transmission peak occurs.

8.2.5.5 Equivalent layer for the spacer — nonnormal incidence

The use of an equivalent layer as a spacer can reduce the polarization splitting of the passband at nonnormal incidence [Costich 70]. The manifestations of such splitting at 50° incidence angle are evident in Fig. 8-18. The *p* and *s* polarized passbands are centered at different wavelengths, which may be attributed to the polarization dependence of the phase shift upon reflection (see Fig. 2-30, for example).

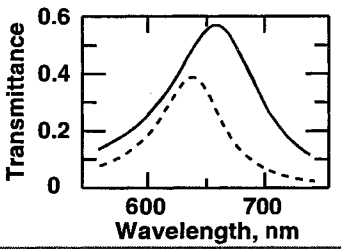


Fig. 8-18 At an incidence angle of 50°, $T(\lambda)$ for *p* polarized flux (*solid*) and *s* polarized (*dashed*) of the Two-M bandpass filter

air **M** 1.348**H M** *glass*

where the indices of **H** and *glass* are 2.25 and 1.52, respectively. **M** represents a layer of 25 nm metric thickness whose optical constants are in Table 8-6. The optical thickness of **H** is $\lambda_0/4$ at λ_0 of 650 nm.

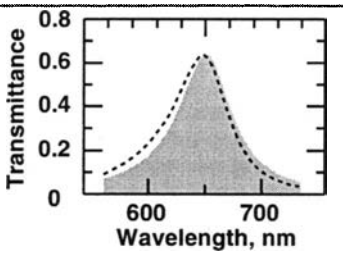


Fig. 8-19 At an angle of 50°, $T(\lambda)$ for *p* polarized flux (*shaded*) and *s* polarized (*dashed*) of the Two-M bandpass filter

air 1.03**L** 1.29**H M** 1.57**H** 0.1**L**
1.57**H M** 2.44**H** *glass*

where refractive indices of *glass*, **L**, **S** and **H** are 1.52, 1.35, 1.62 and 2.25, respectively. **M** is 25 nm in metric thickness and its optical constants appear in Table 8-6. The optical thickness of **H** and **L** is $\lambda_0/4$ at λ_0 of 650 nm, *matched* at 50°.

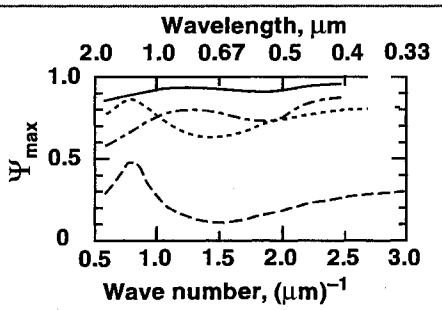


Fig. 8-20 Versus wave number, computed Ψ_{max} of layers of silver of 30 nm metric thickness (upper *solid*), 50 nm (lower *solid*), aluminum of 20 nm metric thickness (upper *dashed*) and 40 nm (lower *dashed*). The optical constants are from Hass [72].

The starting design is

air 0.025**M E** 0.025**M** *glass*

where **E** is an equivalent layer (see §2.6.7) that produces second-order interference. A guess at the phase thickness of the spacer is obtained from Eq. (8-2a), using δ_m of 114° for the high index spacer from Table 8-5. Since this is to be a second-order spacer, however, 180° of phase must be added, resulting in a phase thickness of 114° + 180° = 294°. A computer code synthesized several equivalent layers of 294° phase thickness—the equivalent index ranged from 2.4 to 3.3. An *HLH* type layer (see §2.6.7.2) is used. When the index of the equivalent layer is 2.50, its equivalent index is nearly the same for both polarization states.

The remaining problem is that, although the passband wavelength centers are nearly the same, the peak T_p exceeds T_s by 10%. The solution is to add matching stacks that equalize the transmittances. Thin layers were added to both ends of the stack (see §10.2.4.5) and an optimization computer code adjusted the transmittance only at 650 nm. The thicknesses of the silver and spacer layers were not altered. The final design is captioned in Fig. 8-19. The polarization splitting of the transmittance curves is minimal. The problem is that the coating is not producible because of the ultra-thin **L** layer—its thickness is 0.1 QWOT. One solution is to add a halfwave, so that the resulting design is

air 1.03**L** 1.29**H M** 1.57**H** 2.1**L** 1.57**H M** 2.44**H** *glass* .

The foregoing manifests more polarization splitting than the original in Fig. 8-19.

8.3 Design procedures for metal-dielectric bandpass filters

This section addresses the selection of the metals for a given spectral region, the number of cavities and the thicknesses of the metal layers.

Table 8-7 Metals recommended for metal-dielectric bandpass filters

Region	Metal	Note
120-300 nm	Aluminum	Thickness of the aluminum oxide passivation layer on the aluminum should be minimized at wavelengths below 170 nm.
300-400 nm	Aluminum	Silver has substantial transmittance at wavelengths less than 400 nm, but poor contrast.
400-800 nm	Silver	Silver must be environmentally protected.
800 nm to 1500 nm	Silver	Gold is a possible metal, if transmission leaks at $\lambda < 700$ nm can be filtered. When silver is used in the near IR, the small errors in the layer thicknesses of the matching stacks degrade considerably the passband transmittance.

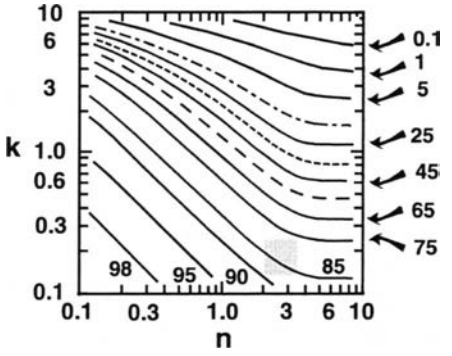


Fig. 8-21 Contours of constant Ψ_{max} (in percent) vs the n (the real part of $\hat{n}$ — *abscissa*) and k (the imaginary part of $\hat{n}$ — *ordinate*) of a film 0.1 wave in metric thickness. The shaded area indicates the n and k of a weakly absorbing dielectric layer. Dashed contours: *Long*—55%; *short*—35%; *long-short*—15%.

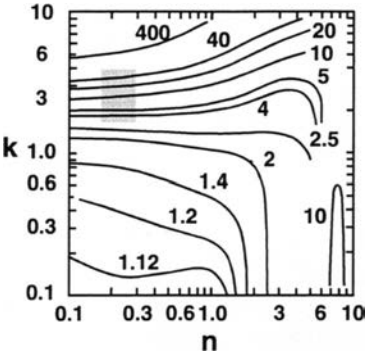


Fig. 8-22 Contours of constant contrast [see Eq. (8-2b)] for the film captioned in Fig. 8-21. The optical constants of silver lie in the shaded area.

8.3.1 Recommended metals for metal-dielectric bandpass filters

§8.2.4.3 illustrates that the Ψ_{max} of a metal film determines, in a large part, the peak passband transmittance. Figure 8-20 graphs Ψ_{max} vs wave number for metals with high electron density, such as silver or aluminum. Table 8-7 cites the recommended metals — principally aluminum in the UV and silver in the visible and near IR. It appears that it is relatively easy to produce a bandpass filter at $\lambda = 2$ μm containing silver films with a high peak transmittance. In fact, the matching stacks have a quite large standing wave ratio. The execution of steps (5) and (6) in §8.2.4.1 shows that the peak transmittance is degraded considerably by thickness errors in the matching stack.

8.3.1.1 Selection of metals — the contrast

The Ψ_{max} (see §8.3.1) is but one criterion for the selection of a metal for a bandpass filter. §8.2.2 discusses the trade-off between peak transmittance and offband rejection. Costich [65a] devised the contrast C as a merit function in the selection of an absorber for a metal-dielectric bandpass filter

$$C \equiv \Psi_{max} / T_0 \qquad (8\text{-}2b)$$

where T_0 is the radiant transmittance of the film with air as both its emergent medium and incident medium. Assuming a reasonable Ψ_{max}, metals with higher C values are more desirable to use in bandpass filters because of their better offband rejection.

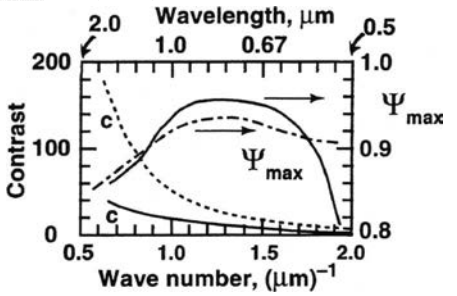


Fig. 8-23 Ψ_{max} (*right ordinate*) and contrast (*left ordinate* — curves marked *c*) of layers of silver (*dashed*) and gold (*solid*) of 30 nm thickness. The optical constants of silver are from Table 8-6 and those of gold are from Hass [72]. The contrast appears in Eq. (8-2b).

As an example, Fig. 8-21 depicts constant Ψ_{max} contours vs n and k for a film 0.1 wave thick. It might be tempting to select a material with an n of 1.5 and k of 0.2 — as shown by the shaded area. Its Ψ_{max} is in excess of 90%. This film is a weakly absorbing dielectric, similar to the layer whose R and T are graphed in Fig. 2-51. Although its Ψ_{max} is adequate, its offband rejection is poor. Examination of Fig. 8-22 establishes that the contrast of a film with an n of 1.5 and k of 0.3 is approximately 1.3. This Fig. also shows that a metal with $n < 0.2$ and $k > 3$ manifests a much larger contrast. Those optical constants are close to those of silver in the visible spectral region.

The next example compares the contrast and Ψ_{max} of gold and silver, as depicted in Fig. 8-23. Although the Ψ_{max} of these two metals is nearly the same at wavelengths longer than 600 nm, the contrast of the silver is superior at all wavelengths.

8.3.2 Trade-off between spectral bandwidth and peak transmittance

The goal is to produce a bandpass filter with passband centered at λ_0. Inspection of Table 8-1 leads to the conclusion that the peak transmittance is less than optimal when the spectral bandwidth approaches small values — say, a few percent. Although this sweeping generalization must be tempered — there are several metals and the wavelengths of the passbands range from the UV to the IR — the following design philosophy is still useful.

A recommended guideline is: Use all-dielectric bandpass filters (see Chap. 7) to attain relatively narrow bandwidth and to provide the offband attenuation in the spectral range $0.8\lambda_0 < \lambda < 1.2\lambda_0$. Use multiple-cavity filters with metal layers in conjunction with absorbers such as dyes or colored glass to attenuate in other spectral regions. This approach is advantageous because it allows the designer to optimize the peak transmittance in the passband and to minimize the offband transmittance — free from other constraints.

8.3.3 Overview of multiple-cavity bandpass filter design

The goal is to design bandpass filters with multiple cavities containing dielectric spacers S and metal layers M, such as

$M_1\ S_1\ M_2\ S_2\ M_3$	Three-M
$M_1\ S_1\ M_2\ S_2\ M_3\ S_3\ M_4$	Four-M
$M_1\ S_1\ M_2\ S_2\ M_3\ S_3\ M_4\ S_4\ M_5$	Five-M
$M_1\ S_1\ M_2\ S_2\ M_3\ S_3\ M_4\ S_4\ M_5\ S_5\ M_6$	Six-M

and so on. Appropriate matching stacks (see Fig. 8-2) may be added, if needed, to increase the maximum passband transmittance.

An example illustrates the advantage of the multiple-cavity bandpass filter. Suppose that at λ_0 = 650 nm, a One-M bandpass filter contains a single silver layer of 80 nm thickness. Its Ψ_{max}, as obtained from either Fig. 8-11 or Table 8-4, is 38%. Figure 8-24 illustrates that its offband rejection is superb — attaining nearly 4 absorbance units at selected wavelengths.

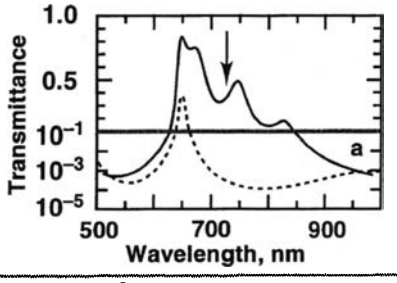


Fig. 8-24 $T(\lambda)$ of a Three-M bandpass filter (*solid*)
cement 1.215**H L** 0.82**H** $\mathbf{M_1}$ 0.82**H** 0.82**H** $\mathbf{M_1}$ 0.82**H** 0.82**H** $\mathbf{M_1}$ 0.82**H** **L** 1.215**H** 0.53**L** *glass*
and a One-M bandpass filter (*dashed*)
cement 0.704**L** 1.133**H L H L** 0.69**H** $\mathbf{M_2}$ 0.69**H L H L** 1.133**H** 0.704**L** *glass* .
The optical constants of $\mathbf{M_1}$ and $\mathbf{M_2}$ are in Table 8-6; their metric thicknesses are 26.7 nm and 80 nm, respectively. The refractive indices of **L**, **H**, *cement* and *glass* are, respectively, 1.35, 2.25, 1.52 and 1.52. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 650 nm. The ordinate changes from linear to log at 0.1.

Suppose that the 80 nm of silver mentioned above is split into three layers of equal thickness. The Ψ_{max} of a layer of 26.7 nm thickness is 94%. The optimal T_{max} of a Three-M bandpass filter containing three layers of that thickness is $\Psi^3_{max} = (0.94)^3 =$ 83%. The foregoing conclusion is valid because of the multiplicative properties of Ψ (see §2.9.2.1). This T_{max} of 83% is attained only if the proper dielectric matching stacks are added. The computed T_{max} of the Three-M bandpass filter in Fig. 8-24 is 80%, which is within 3 percent of Ψ_{max}. The use of three silver layers — in lieu of the single layer — more than doubles T_{max}.

The maximum passband transmittance is but one of the concerns. At wavelengths longer than λ_0, the agonizingly slow decrease of the transmittance of the Three-M bandpass filter is contrasted with the performance of the One-M bandpass filter, whose $T(\lambda)$ plunges much more rapidly to quite small values. Consider the arrowed wavelength in Fig. 8-24, which is about 100 nm greater than λ_0. The transmittance of the One-M bandpass filter is more than two decades below that of the Three-M. At wavelengths near 950 nm (near *a* in Fig. 8-24), however, the "interference effects" between the metal layers in the Three-M bandpass filter subside and the transmittance of both filters is nearly the same.

The Three-M bandpass filter would decrease more rapidly at wavelengths longer than λ_0 if: (1) Low index (rather than high index) spacer layers were used; and (2) the center metal layer were thickened.

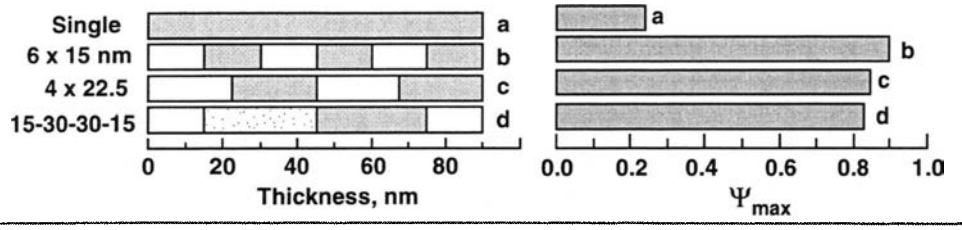


Fig. 8-25 (*Left panel*) Silver film of 90 nm thickness—(*a*); 15 nm thickness (*non-shaded*—*b* and *d*); 22.5 nm thickness (*shaded*—*c*); and 30 nm thickness (*dotted*—*d*). (*Right panel*) Ψ_{max} of: (*a*) single film of 90 nm thickness; (*b*) six films of 15 nm thickness; (*c*) four films of 22.5 nm thickness; and (*d*) films of thicknesses 15 nm, 30 nm, 30 nm and 15 nm.

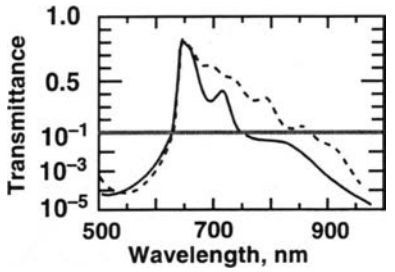


Fig. 8-26 $T(\lambda)$ of a Six-M bandpass filter (*dashed*)
air 1.992**L** 2.94**H M** 1.786**L M** 1.83**L M** 1.84**L M** 1.843**L M** 1.84**L M** 2.92**H** 0.934**L** 0.93**H** *glass*
and a Four-M bandpass (*solid*)
air 0.85**H** 0.885**L** 0.994**H M** 1.795**L** 2**M** 1.73**L** 2**M** 1.79**L M** 0.925**H** 0.965**L** 0.965**H** 0.965**L** *glass*
where the indices of **L**, **H** and *glass* are 1.35, 2.25 and 1.52, respectively. **M** represents a silver layer of 15 nm metric thickness whose optical constants appear in Table 8-6. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 650 nm. The ordinate changes from linear to log at 0.1.

8.3.4 Guidelines for the design of multiple-cavity bandpass filter

There are many ways of choosing the sequence of thicknesses of the metal layers. Guidelines are illustrated by several examples. Consider a bandpass filter in which the total thickness of the silver is 90 nm. Figure 8-25 shows four possible bandpass filter configurations: (a) A One-M; (b) a Six-M containing six equal thickness silver layers — each is 15 nm thick; (c) a Four-M in which four metal layers of 22.5 nm thickness are cascaded; and (d) a bandpass filter (designated as "15-30-30-15") in which the outer metal layers are double the thicknesses of the inner ones, as in the design

air Match $\mathbf{M}_{15}\ \mathbf{S}\ \mathbf{M}_{30}\ \mathbf{S}\ \mathbf{M}_{30}\ \mathbf{S}\ \mathbf{M}_{15}$ Match *glass*

where the subscripts designate the thicknesses of the silver layers in nm. **S** is a dielectric spacer. The dielectric matching stacks "Match" are optional. The foregoing "air-to-glass" design containing silver is probably environmentally unstable and is included for pedagogical reasons—to illustrate that the matching stacks are not identical when the bandpass filter is asymmetrical.

Figure 8-26 illustrates that the thickness ratios of the metal layers is important. This Fig. compares the transmittance of the Six-M bandpass filter (with the relatively thin layers) and the "15-30-30-15" Four-M. The transmittance of the former is often a decade higher than that of the latter on the long wavelength side of λ_0. §8.2.5.4.2 illustrates that the bandwidth of a Two-M bandpass filter with low index spacers is narrower than the

bandwidth of the Two-M with high index spacers. The same phenomenon is in evidence in Fig. 8-27, which depicts the transmittance of Three-M filters. The coating with the low index spacers manifests a decade more rejection at selected wavelengths.

8.3.5 Bandpass filter design procedures

The filters designed in previous sections with passbands centered at λ_0 furnish guidelines:

- The attenuation at wavelengths greater than 1.8 times λ_0 is essentially determined by the total amount of metal in the coating.
- The total metal thickness mentioned above is more effective in attenuating at longer wavelengths when it is partitioned so that thicker layers are used in the center of the bandpass filter.
- Better attenuation in the offband spectral region is produced by low index spacer layers, as compared to a bandpass filter with high index spacers.

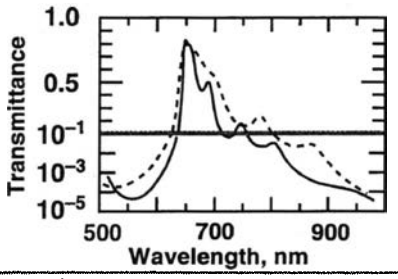


Fig. 8-27 $T(\lambda)$ of a Four-M bandpass filter with low index spacers (*solid*)
air 0.602**H** 0.982**L** 2.997**H** 1.980**L** **M** 1.778**L** **M** 1.778**L** **M** 1.771**L** **M** 1.931**L** 2.966**H** 0.986**L** 0.986**H** *glass*
and bandpass with high index spacers (*dashed*)
air **H** 1.93**L** **M** 1.562**H** **M** 1.637**H** **M** 1.632**H** **M** 1.924**L** 0.986**H** 0.98**L** 0.979**H** *glass*
where the indices of **L**, **H** and *glass* are 1.35, 2.25 and 1.52, respectively. Each **M** layer is 22.5 nm in metric thickness with the optical constants listed in Table 8-6. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 650 nm. The ordinate changes from linear to log at 0.1.

Two design methods that are discussed are:
(1) optimization; and
(2) analytical.

8.3.5.1 Optimization

The bandpass filters in Figs. 8-26 and 8-27 were designed by an optimization computer code that maximized T at λ_0. The code was not permitted to alter the thicknesses of the metal layers. The starting design consisted of: (a) a metal-dielectric stack alternating silver films and halfwave optical thickness spacer layers; and (2) quarterwave dielectric layers exterior to the metal-dielectric stack. The latter functioned as the matching stacks. Any layer whose thickness became negative was discarded.

8.4 Bandpass filter design examples

8.4.1 Introduction

Figure 8-2 depicts completed four cavity bandpass filters — with and without external matching layers. Table 8-8 depicts a disassembled stack as it appears during the various stages of design. The steps are mentioned below. The letters M_1, S_1, etc., refer to the metal and spacer layers, respectively.

The bandpass filters in the design examples are intended to produce substantial attenuation at wavelengths as long as 1200 nm — where the silicon photodiode responsivity is essentially zero. Most bandpass filters manifest transmission leaks on the short wavelength side of λ_0, which may be removed by multilayer blockers, absorbing glass or organic dyes.

Although the examples pertain to Five-M and Six-M filters, the same principles are applicable to other designs. In each case, the final bandpass filter is sandwiched between glass and optical cement of the same refractive index.

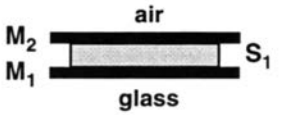
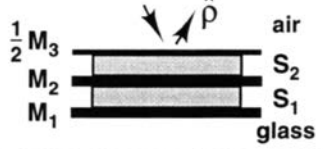
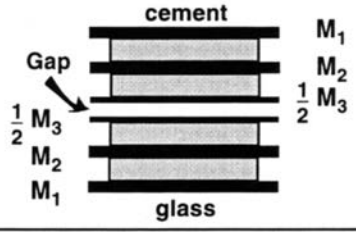
8.4.2 Design procedure — the number of cavities is even

The first example consists of a Five-M bandpass filter with no matching layers on its exterior. A subsequent design includes a few matching layers.

8.4.2.1 Selection of the thickness of the outer metal layer

The first step is to choose a thickness for the outer metal layer $\mathbf{M}_1$. In general, the thinner it is, the better the peak transmittance. A constraint is that if the layer is too thin — say, thinner than 10 nm — an island film is formed that fails to manifest the bulk optical constants of Table 8-6. A 15 nm thickness is chosen. Table 8-4 shows that its Ψ_{max} is quite large, namely 0.982.

Table 8-8 Construction of a Five-M bandpass filter

Config. before	What is added	What is adjusted	Design after
glass This is the uncoated substrate.	$\mathbf{M}_2\ \mathbf{S}_1\ \mathbf{M}_1$ The first two metal layers and spacer.	Adjust optical thickness of spacer $\mathbf{S}_1$ — maximize Ψ at λ_0.	
$\mathbf{M}_2\ \mathbf{S}_1$ $\mathbf{M}_1$ *glass*	$\frac{1}{2}\mathbf{M}_3\ \mathbf{S}_2$ Next spacer and one-half of the center metal layer.	Adjust thickness of spacer $\mathbf{S}_2$ until phase shift upon reflection δ at λ_0 is 180°.	
Bottom one-half of the filter.	Reflect design about surface of one-half thickness $\mathbf{M}_3$ layer.	Nothing.	

Five-M bandpass with silver layers **M** (*solid black*) and dielectric spacer layers **S** (*shaded*)
cement $\mathbf{M}_1\ \mathbf{S}_1\ \mathbf{M}_2\ \mathbf{S}_2\ \mathbf{M}_3\ \mathbf{S}_2\ \mathbf{M}_2\ \mathbf{S}_2\ \mathbf{M}_1$ *glass*.

The "Gap" in the bottom right panel disappears when $\frac{1}{2}\mathbf{M}_3$ layers are coalesced.

8.4.2.2 Selection of next metal-dielectric pair

The next step is to add a dielectric spacer $\mathbf{S}_1$ and metal layer $\mathbf{M}_2$, resulting in the design

air $\mathbf{M}_2\ \mathbf{S}_1\ \mathbf{M}_1$ *glass* .

For reasons mentioned in §8.3.4, $\mathbf{M}_2$ should be thicker than $\mathbf{M}_1$—a thickness of 25 nm is selected. Table 8-8 depicts the partially completed bandpass filter.

The optical thickness for $\mathbf{S}_1$ is determined by varying its optical thickness between $\lambda_0/4$ and $\lambda_0/2$. The optimum thickness maximizes the net flux ratio Ψ. Ψ is independent

of the index of the incident medium (see §2.9.2.4)—this is the reason air may be chosen for that medium during this design step. The result is that S_1 is 1.62 QWOTs when its index is 1.35.

8.4.2.3 Construction of the half bandpass filter

There is a metal film M_3 at the center of the bandpass filter, which is bisected (using the symbol $\frac{1}{2}M_3$). In other words, the final bandpass filter is *partitioned* (see §2.6.4) into two identical sections. The addition of the spacer S_2 and bisected center layer produces the *half filter*

$$\tfrac{1}{2}M_3\ S_2\ M_2\ S_1\ M_1\ \textit{glass}\ .$$

The M_3 layer of 30 nm thickness is chosen so that it is thicker than M_2. The optical thickness of S_2 is adjusted so that the phase shift upon reflection δ of the half filter (shown above) is 180° or 0°. Examination of Eq. (2-45h) shows that the transmittance of the entire bandpass filter is a maximum when the admittance of the *half filter* is pure real. Inspection of Eq. (2-16h) shows that this occurs when the phase shift upon reflection is 0° or 180°.

This method of designing a half bandpass filter with the correct admittance is an approximation — and a good one. In fact, the T_{max} in Eq. (2-45h) also varies with wavelength and thus the optimum T occurs at *nearly* — but not *exactly* — the wavelength at which the admittance is pure real. An optical thickness of 1.72 QWOTs for S_2 produces the aforementioned *desideratum* phase shift upon reflection of 180°.

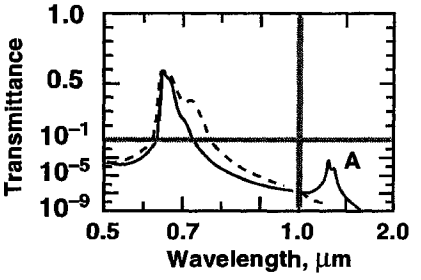


Fig. 8-28 $T(\lambda)$ of the Five-M bandpass filter (*dashed*)
cement 0.015M 1.63L 0.025M 1.72L 0.03M 1.72L 0.025M 1.63L 0.015M *glass*
and (*solid*)
cement 0.015M 1.63L 0.025M 3.72L 0.03M 3.72L 0.025M 1.63L 0.015M *glass* .
The scale of the abscissa changes at 1.0 μm.

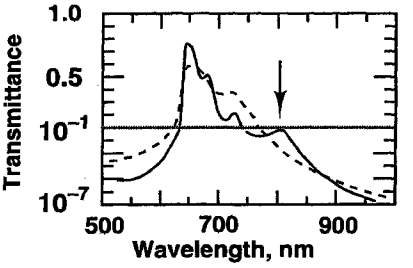


Fig. 8-29 $T(\lambda)$ of a Five-M bandpass filter with matching layers(*solid*)
cement H L H 1.93L 0.015M 1.82L 0.025M 1.73L 0.03M 1.73L 0.025M 1.82L 0.015M 1.93L H L H *glass*
and with no matching layers (*dashed*)
cement 0.015M 1.63L 0.025M 1.72L 0.03M 1.72L 0.025M 1.63L 0.015M *glass* .

For both Figs. the ordinate changes from linear to log at 0.1. The indices of L, H, *glass* and *cement* are 1.35, 2.25, 1.52 and 1.52, respectively. M represents a layer of 1.0 μm metric thickness whose optical constants are in Table 8-6. The optical thickness of L and H is $\lambda_0/4$ at λ_0 of 650 nm.

8.4.2.4 Reflection of the half bandpass filter

The half bandpass filter is reflected about its outer surface and the resulting bandpass filter is

cement $M_1\ S_1\ M_2\ S_2\ M_3\ S_2\ M_2\ S_1\ M_1$ *glass* .

Its transmittance is graphed in Fig. 8-28.

8.4.2.5 Five-M bandpass filter with matching stack The T_{max} of the foregoing four cavity bandpass filter is 59% at λ_0. How close is this to optimal? For silver films of 15 nm, 25 nm and 30 nm thickness, Table 8-4 shows that the Ψ_{max} attains values of 0.982, 0.949, and 0.926, respectively. The theoretical maximum transmittance is, therefore, $(0.982)^2 \times (0.949)^2 \times 0.926 = 80.3\%$. In conclusion, it is possible to enhance the transmittance by an additional 21.3%, while maintaining the thicknesses of the metal layers.

A matching stack consisting of four dielectric films is appended to the ends of the Five-M bandpass filter. As discussed in §8.2.4.2.1, this stack serves to increase—if not optimize—Ψ at the interface of the outer silver layer of 15 nm thickness. It is necessary to recalculate the thicknesses of the spacer layers—using the procedures of §8.4.2.2 and §8.4.2.3. The design of the resultant bandpass filter is captioned in Fig. 8-29. Figure 8-29 compares the $T(\lambda)$ of the nine-layer Five-M bandpass filter (*dashed*) with the bandpass filter that includes the matching stacks (*solid*). The T_{max} of the latter is 79%, which is reasonably close to the optimum of 80.3% cited in the foregoing paragraph.

The coating with the matching stacks is not without its disadvantages, however. As shown in Fig. 8-29, there is an annoying long wavelength leak (arrowed) produced by the matching stacks. In addition, the transmittance spike near the peak of the passband is quite narrow—about 3% semiwidth. A consequence is that any errors in the layer thicknesses introduced in the fabrication of the bandpass filter could easily reduce the peak transmittance. The Five-M bandpass filter without the matching stacks has a relatively broad passband and is more tolerant of thickness errors.

8.4.2.6 More rapid increase of offband rejection

One of the deficiencies of the Four-M bandpass filter is that its transmittance (*dashed* in Fig. 8-28) does not decrease rapidly enough on the long wavelength side of λ_0. A remedy (see §8.2.5.3) is to insert higher order spacers. A substantial transmission leak at longer wavelengths occurs if all spacers are second order. A compromise is to convert only the two center spacers to second order. The solid curve in Fig. 8-28 shows that the transmittance at 850 nm is decreased by several decades.

The inevitable long wavelength leak at approximately $2\lambda_0$ occurs, however. This is near the letter *A* in Fig. 8-28. Although the peak transmittance of the leak is only 2×10^{-4} at 1.2 μm, it might be sufficiently large to preclude the use of this bandpass filter in conjunction with a silicon photodiode detector.

8.4.3 Design procedure—the number of cavities is odd

8.4.3.1 Introduction

The thicknesses for the metal layers are the same as those in §8.4.2, with the exception that two 30 nm silver films, rather than one, are inserted into the center of the bandpass filter. All spacers are low index. A few of the steps in the design of this Six-M bandpass filter are identical to those in §8.4.2.

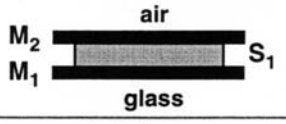
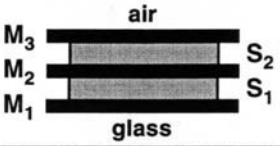
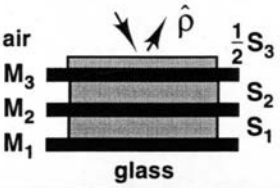
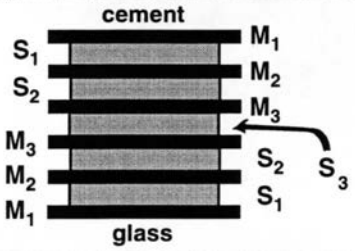
8.4.3.2 Design of the outer cavity

The first step is to add a metal and a metal-dielectric pair so that the partial bandpass filter is *air* $M_2\,S_1\,M_1$ *glass*. The thickness of S_1 is adjusted until Ψ is maximized. The result is identical to that in §8.4.2.2.

8.4.3.3 Addition of the second cavity

Another metal-dielectric pair is added and the partial bandpass filter design is *air* $M_3\,S_2\,M_2\,S_1\,M_1$ *glass*. Ψ is optimized by varying the thickness of S_2. The result is that the optical thickness of S_2 is 1.74 QWOTs.

Table 8-9 Construction of a Six-M bandpass filter

Configuration before	What is added	What is adjusted	Design after
glass The uncoated glass substrate.	$M_2\,S_1\,M_1$ Two metal layers and spacer.	Adjust the thickness of spacer S_1 until Ψ at λ_0 is a maximum.	
$M_2\,S_1\,M_1$ *glass*	$M_3\,S_2$ The next metal layer and spacer.	Adjust the optical thickness of spacer S_2 until Ψ at λ_0 is a maximum.	
$M_3\,S_2$ $M_2\,S_1\,M_1$ *glass*	$\frac{1}{2}S_3$ Add one-half of the center dielectric spacer layer.	Adjust the thickness of the half spacer $\frac{1}{2}S_3$ until the phase shift upon reflection δ at λ_0 is 0° or 180°.	
$\frac{1}{2}S_3\,M_3\,S_2$ $M_2\,S_1\,M_1$ *glass*	Reflect the design about the half spacer layer, which, in the final design, is the center spacer layer.	Nothing.	

Steps in the design of a Six-M bandpass using silver layers M and dielectric spacers S
cement $M_1\,S_1\,M_2\,S_2\,M_3\,S_3\,M_3\,S_2\,M_2\,S_2\,M_1$ *glass*.

8.4.3.4 The half bandpass filter — thickness of the center spacer

The thickness of the center spacer S_3 is chosen by converting the foregoing design into a half bandpass filter

air $\frac{1}{2}S_3\,M_3\,S_2\,M_2\,S_1\,M_1$ *glass*

where $\frac{1}{2}S_3$ represents a cryolite film of one-half of the optical thickness of the layer in the completed bandpass filter. The justification of the foregoing procedure is similar to the

arguments in §8.4.2.3. The thickness of $\frac{1}{2}S_3$ is varied until the admittance of the half bandpass filter is pure real, which is equivalent to stipulating a phase shift upon reflection of 0° or 180°. The result is an optical thickness of 0.86 QWOTs for $\frac{1}{2}S_3$.

8.4.3.5 Performance of the completed bandpass filter

The design is reflected about the outer surface of the "half spacer." The final bandpass filter is depicted in the lower right corner of Table 8-9. The solid curve in Fig. 8-30 graphs $T(\lambda)$ of the Six-M bandpass filter. The $T(\lambda)$ of the Five-M bandpass filter designed previously is included for comparison. The addition of another 30 nm silver layer to the Five-M bandpass filter (to form the Six-M bandpass filter) decreases its peak T—T_{max}—from 58.8% to 54%. This loss of T_{max} is offset by a decade of reduction in T at 1.0 μm.

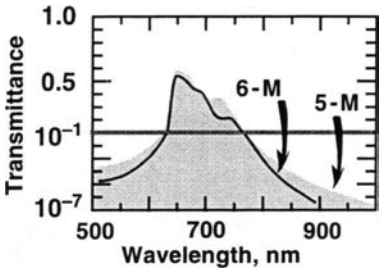


Fig. 8-30 $T(\lambda)$ of a Five-M bandpass filter (*shaded*)
cement 0.015**M** 1.631**L** 0.025 1.723**L** 0.03**M** 1.723**L** 0.025**M** 1.631**L** 0.015**M** *glass*
and Six-M bandpass filter (*solid*)
cement 0.015**M** 1.63**L** 0.025**M** 1.74**L** 0.03**M** 1.72**L** 0.03**M** 1.74**L** 0.025**M** 1.63**L** 0.015**M** *glass*
where the indices of **L**, **H**, *glass* and *cement* are 1.35, 2.25, 1.52 and 1.52, respectively. **M** represents a layer of 1.0 μm metric thickness whose optical constants are in Table 8-6. The optical thickness of **L** is $\lambda_0/4$ at λ_0 of 650 nm. The ordinate changes from linear to log at 0.1.

8.4.4 General discussion

The examples in §8.3 and §8.4 illustrate that the design of metal-dielectric bandpass filters involves trade-offs and, for that reason, it is difficult to specify any universal design rules. One of the constant problems is the control of the optical thicknesses of the layers and how much the peak transmittance (or bandwidth) is degraded by thickness errors during production. Only the experience in the coating machine can provide answers to such a question.

Caution should be exercised on the placement of the silver layers in the bandpass filter. If they are close to the outer surface, optical cements (such as epoxy) may react chemically with the silver and degrade it.

8.4.5 Design of an asymmetrical three-M filter

The bandpass filters designed in §8.4.2 and §8.4.3 are symmetrical. The bandpass filter developed in this section is asymmetrical and, for reasons mentioned in §8.4.5.7, would function well if the layer thicknesses were wedged, as shown in Fig. 1-232, but otherwise with difficulty. The passband center wavelength is at λ_0 of 2200 nm.

The low index layers (symbol **L**) and high index layers (symbol **H**) manifest an optical thickness of $\lambda_0/4$ at λ_0 of 2200 nm and refractive indices of 1.35 and 2.25, respectively. The three silver layers M_1, M_2 and M_3 grow progressively thicker, from 11 nm to 18 nm and to 26 nm.

8.4.5.1 The matching stack for the thinnest silver layer

The metric thickness of the thinnest silver layer M_1 is chosen as 11 nm. Its Ψ_{max} (see §2.9.4.2) at λ_0 is 0.9805. Assume that a high index film (i.e., the phase-matching layer) is contiguous to the emergent side of the silver. The aforementioned Ψ_{max} is achieved when the standing wave ratio V of the stack on the emergent side of the phase-matching layer is 53. Given that the substrate is glass of refractive index 1.51, a nine-layer matching stack is requisite in order to achieve Ψ_{max}. A compromise is to use a five-layer matching stack, which results in the design

medium M_1 0.98**H L H L H** *glass.*

The Ψ of the above design is 0.954, which is less than 3% below Ψ_{max}. The layer "0.98**H**" above is the phase-matching layer, the thickness of which is adjusted to maximize Ψ.

8.4.5.2 Addition of the center silver layer

The next task is to add the center silver layer M_2 of thickness 18 nm. A calculation shows that when a high index layer is contiguous to M_2 on its emergent side, at λ_0 a V of 36.8 is needed to achieve Ψ_{max}. A still larger V of 61 is needed when a low index layer is contiguous to M_2.

Computer analysis shows that the design

medium M_1 0.98**H L H L H** *glass*

produces at λ_0 a V of 14 and 23 when the *medium* manifests a refractive index 2.25 and 1.35, respectively. Since these values fall short of the *desideratum* of 36.8, we have no choice but to use a two-layer "admittance-step-up-transformer" stack to boost the V.

A calculation shows that at λ_0 the stack

medium 0.965**H** 0.978**L** M_1 0.98**H L H L H** *glass*

manifests a V of 37.7 when the *medium* index is 2.25. In the foregoing design the layer "0.978**L**" alters the admittance of the layers on its emergent side so that the admittance is reasonably close to being pure real. The layer "0.965**H**" is the phase-matching layer whose thickness is varied to optimize Ψ.

Was the foregoing design approach successful? The answer is found by computing Ψ (at λ_0) after addition of the M_2 layer, whose Ψ_{max} is 0.9472. After addition of the M_2 layer, the Ψ is 0.9036. Since the Ψ of the M_1 layer is 0.9543, the largest value that Ψ may achieve for these two silver layers is Ψ = (0.9543) (0.9472) = 0.9039. This is quite close to the 0.9036 cited above. The foregoing calculation assumes the five-layer matching stack on the emergent side of the M_1 layer, as developed in §8.4.5.1.

8.4.5.3 Addition of the thickest silver layer

The next task is to add the thickest silver layer M_3 of 26 nm thickness. At λ_0 its Ψ_{max} is 0.8876. The foregoing value is achieved when the standing wave ratio V of the stack on its emergent side is 31.7 (when there is a high index layer on the emergent side of M_3) or 52.6 (when a low index layer is on the emergent side of M_3). In this instance, there is no need to use a low index spacer, since a V of 39 is produced by

medium M_2 0.965**H** 0.978**L** M_1 0.98**H L H L H** *glass*

when the medium index is 2.25. It is noted that this silver layer M_3 is slightly *over-matched*, since the V of 39 exceeds the target V of 31.6. The thickness of the **H** layer between M_3 and M_2 is adjusted to maximize Ψ and the design becomes

medium M_3 1.91**H** M_2 0.965**H** 0.978**L** M_1 0.98**H L H L H** *glass*

which at λ_0 manifests a Ψ of 0.80.

How close is the foregoing design to the optimum? If the M_2 and M_3 layers were each optimally matched so that Ψ_{max} was achieved, the Ψ of the aforementioned three-M filter is

$\Psi = (0.9543)\ (0.9472)\ (0.8876) = 0.8023.$

The conclusion is that the M_2 and M_3 layers in the present stack are quite close to being optimally matched.

8.4.5.4 Antireflection of the stack

The remaining task is to add an all-dielectric stack that reduces the reflectance to nearly zero at λ_0. The design of the seven-layer antireflection coating is

air **H' L H L H L** 0.944**H**

where the optical thickness of the **H'** layer is slightly thicker than a quarterwave. The final design contains 18 layers and is

air 1.10**H L H L H L** 0.944**H** M_3 1.91**H** M_2 0.965**H** 0.978**L** M_1 0.980**H L H L H** *glass* .

Its transmittance appears in Fig. 8-31. At the tuned wavelength λ_0, the reflectance from both the air and glass sides of the bandpass filter is below one percent. The peak passband transmittance is 78%.

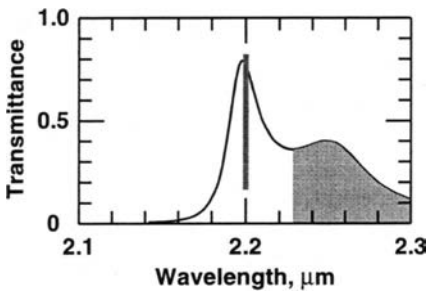


Fig. 8-31 $T(\lambda)$ of an asymmetrical Three-M bandpass filter of the design

air 1.10**H L H L H L** 0.944**H** 26**M** 1.91**H** 18**M** 0.965**H** 0.978**L** 11**M** 0.980**H L H L H** *glass* .

A design objective is to reduce the transmittance in the shaded spectral region.

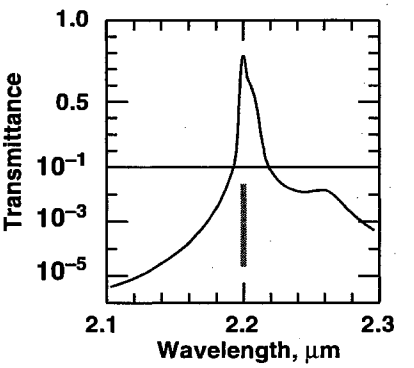


Fig. 8-32 $T(\lambda)$ of an asymmetrical Three-M bandpass filter of the design

air 3.008**H L** 5.0127**H** 1.003**L** 1.002**H** 5.677**L** 4.412**H** 26**M** 5.913**H** 18**M** 3.363**H** 6.406**L** 11**M** 0.9854**H** 1.002**L H L H** *glass* .

The scale of the ordinate changes from linear to log at 0.1.

The refractive indices of **L**, **H**, and *glass* are 1.35, 2.25 and 1.51, respectively.

M represents a silver layer of one nm metric thickness whose optical constants appear in Table 8-6. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 2200 nm.

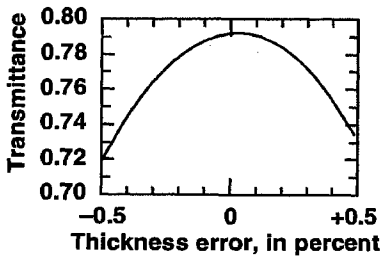


Fig. 8-33 Peak transmittance vs thickness error (in percent) for the layer "4.412H" that is a component of the bandpass filter captioned in Fig. 8-32.

8.4.5.5 Computer optimization—starting design for optimization

The remaining task is to eliminate the long wavelength transmittance "tail," which is identified as the shaded area in Fig. 8-31. The starting design for the computer optimization is the design achieved in §8.4.5.4, with the exception that the thicknesses of a few layers in the matching stack have been increased by adding either a full wave or two full waves to their optical thicknesses.

8.4.5.6 Computer optimization—results

The transmittance of the optimized design appears in Fig. 8-32. The thicknesses of the three silver layers were not altered nor was the layer count increased. The optimization targets consisted of a maximum transmittance at 2200 nm and a linear decrease in transmittance from 0.45 to 0.01 in the spectral region 2208 nm to 2216 nm. The resulting design is

↓

air 3.008**H L** 5.0127**H** 1.003**L** 1.002**H** 5.677**L** 4.412**H** 26**M** 5.913**H** 18**M** 3.363**H** 6.406**L**
11**M** 0.9854**H** 1.002**L H L H** *glass* .

A subsequent section analyzes how the passband transmittance is influenced by thickness errors in the layer that is arrowed above.

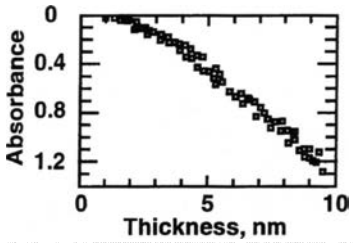


Fig. 8-34 At a wavelength of 633 nm, absorbance vs metric thickness (estimated from the sputtering conditions) of tantalum metal layers. After Dew [92].

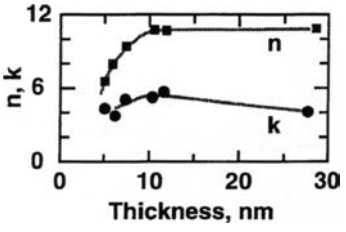


Fig. 8-35 At a wavelength of 2.8 μm, refractive index *n* (*upper, squares*) and extinction coefficient *k* (*lower, circles*) vs metric thickness of the nickel film. After Goede [92].

8.4.5.7 Dependence of the performance upon layer thickness

It is easily demonstrated that the passband transmittance is quite sensitive to the thicknesses of the dielectric layers adjacent to the exterior silver layers—these are often termed the "phase-matching layers." As an example, Fig. 8-33 plots the peak transmittance of the bandpass filter vs the thickness change for the high index layer that is on the incident side of the outer silver layer "26M". This high index layer is arrowed in §8.4.5.6. The Fig. illustrates that a mere one-half percent change in the layer thickness decreases the peak passband transmittance by five percent. This suggests that it would

not be practical to produce such a filter because of its low yield. However, consider the situation when the layer thicknesses are wedged, as shown in Fig. 1-232, in which case there is a lateral position on the coated surface where the correct thickness for "phase-matching layers" occurs and the peak transmittance is obtained.

8.5 Dark mirror absorber

This coating is opaque and hence it absorbs virtually all of the incident flux when its reflectance is small. §8.5.1 discusses the general problems associated with its design and production. Subsequent sections elaborate upon the optimization and analytical design approaches.

8.5.1 Characterization of the metal

The coatings described below often require layers as thin as 5 nm—layers that are island films (see §3.3.2.3) and that do not manifest the bulk optical constants. They may be used in the coating if their optical constants are reproducible from run to run, even if not known.

An example is the transmittance vs thickness of thin tantalum layers depicted in Fig. 8-34 [Dew 92]. The thickness was estimated from the power settings and pressure during the sputtering process. The scatter of the data points indicates the reproducibility of the process.

Figure 8-35 depicts the optical constants (at a wavelength of 2.8 μm) of a nickel layer. The accuracy of the data is confirmed by the spectral reflectance of the final design, which contains nickel layers as thin as 7 nm.

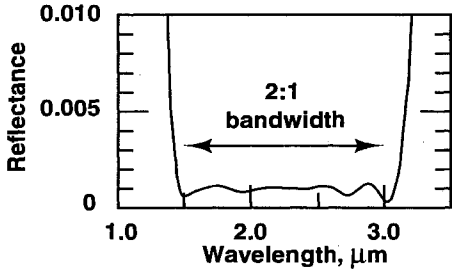


Fig. 8-36 $R(\lambda)$ of a dark mirror of the design

air 307.04**L** 12.31**M** 136.47**L** 18.63**H** 98.14**L** 34.52**H** 29.66**M** 228.53**H** 77.23**M** 157.75**H** *metal*

where the optical constants of **M** and *metal* are those of bulk titanium, as listed in §8.9.2. The refractive indices of **L** and **H** are 1.45 and 2.30, respectively, and *metal* is an opaque layer. The layer thicknesses are metric, in nm.

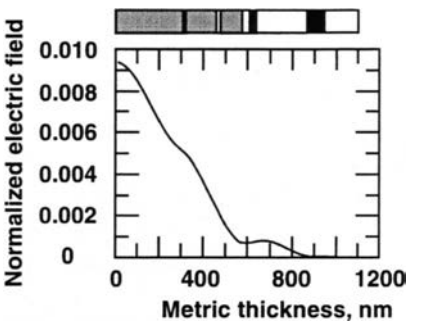


Fig. 8-37 At a wavelength of 2.0 μm, normalized standing wave electric field [see Eq. (2-63g)] vs metric thickness of the coating captioned in Fig. 8-36. The upper panel shows the layer thicknesses, with the substrate on the right

8.5.2 Optimization as a design method

The design procedure is: (1) Choose the metal(s) to be used in the coating, as well as the dielectrics. It is useful if a high index and low index are available. (2) Use a starting design of the Coursen variety (see §10.2.4.5), alternating between dielectric and metal layers. (3) Use an optimizing computer code to reduce the reflectance over the selected bandwidth. (4) If the final design contains layers thinner than 10 nm and if the optical constants of such thin layers are unknown, either discard them or anchor them at a greater thickness. (5) Re-optimize if either of the procedures in step (4) was executed. A useful design tool is an optimization computer code that alters the optical constants as the thickness changes.

As an example, the goal is a dark mirror over a 2:1 bandwidth in the near IR — from 1500 nm to 3000 nm. The starting design is

air **L** (0.02**M H**)^6 *metal*

where the **M** layers are titanium and the *metal* is an opaque substrate of the same material. The "0.02" means that each layer is 0.02 μm (20 nm) thick. Each of each dielectric layer (i.e., **H** and **L**) is a QWOT at 1 μm.

After optimization, the thirteen layers in the starting design were reduced to the ten in the final design, which is captioned in Fig. 8-36. The foregoing layer count excludes the opaque titanium metal layer that is contiguous to the substrate. A 2:1 bandwidth is achieved, in which the reflectance does not exceed 0.15 percent. Each of the metal layers is thicker than 10 nm, which means that the optical constants of the layers resemble those of the bulk metal.

Figure 8-37 graphs the normalized standing wave electric field, which, for the most part, decreases monotonically from 1.0 at the outer surface to zero at the boundary of the opaque layer on the substrate.

8.5.3 Analytical design procedure

The first step is to select a metal (or absorber). A spacer layer is inserted between the thin absorber layer and the opaque metal substrate. §8.5.3.2 discusses the spacer and the dielectric antireflection stack coated on top of it.

8.5.3.1 Selection of the absorber

The absorber can be a metal or a lossy material such as iron oxide in the visible part of the spectrum. It is important that both the opaque metal substrate and the semitransparent absorbing layers have a relatively low reflectance.

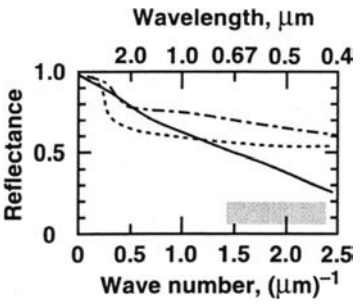


Fig. 8-38 $R(\lambda)$ of the bulk opaque metals: nickel (*solid*), titanium (*short dash*) and platinum (*long dash*). The visible portion of the spectrum is shaded. After Weaver [81].

Consider the possible metals for the near IR—the spectral region near 2.0 μm. If used alone as single layers, gold, silver and aluminum are poor choices for the metal because of their relatively high reflectance.

Figure 8-38 graphs the reflectance of nickel, titanium and platinum—metals that have a reflectance of less than 75% in the near IR. Platinum is quite chemically stable and easy to sputter, although more expensive to deposit by vacuum evaporation. Titanium is used in the next example, which consists of two dielectric layers and a thin metal layer.

8.5.3.2 Spacer layer and the antireflection coating

The H layer—the spacer—is sandwiched between the thin metal layer and the opaque substrate, as shown in Fig. 8-39. The phase shift upon reflection of the opaque substrate causes the node of the standing wave to be buried beneath the surface (see Fig. 2-27).

The thin metal layer is at the antinode of the standing wave where the electric field peaks and hence the absorption [see Eq. (2-62d)] is a maximum. A single layer of low index L reduces the reflectance to nearly zero. Figure 8-40 graphs the $R(\lambda)$ of the trilayer coating. Designs similar to the foregoing are often used as write-once read-only storage media in data recording. A focused laser beam burns through the outer metallic layer, which causes the reflectance to rise abruptly.

Goede [88] describes a dark mirror containing nickel layers buried beneath a dielectric stack. The mirror reflects strongly at λ_0 and is highly absorbing at $2\lambda_0$.

8.6 Reflectors

Although rather specialized, both the asymmetrical reflector and the reflection bandpass filter have been used in optical systems.

8.6.1 Reflector with different reflectance from air and substrate sides

This coating consists of one or more layers of semitransparent metal surrounded by dielectrics. The dielectrics enhance the reflectance on one side and reduce the reflectance on the other. In other words, one surface is shiny and the opposite surface is dark. Figure 8-41 shows the reflectance of a thin film of a transition metal surrounded by six dielectric layers. Its reflectance at 633 nm (arrowed) is quite low.

Smith [51] produced such a coating and used it as a mirror in a Fizeau interferometer. The fringes in *reflection* were similar to those seen in *transmission* when an all-dielectric mirror is used in the same interferometer. Pohlack [59] coated eye glasses with this type of reflector and Apfel [72] suggested its use on window glass.

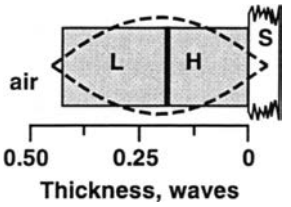


Fig. 8-39 The design of a dark mirror
air 1.88**L M** 1.464**H** *metal*
where the optical thicknesses are plotted. A standing wave (at a wavelength of 2 μm in vacuum) is superimposed.
The symbols **M**, **H** and **L** are defined in the caption to Fig. 8-40.

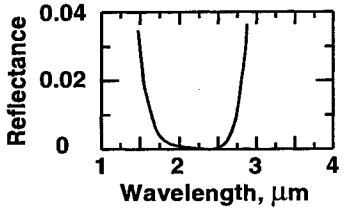


Fig. 8-40 $R(\lambda)$ of the dark mirror whose design appears in Fig. 8-39. The refractive indices of **L** (layer *a*) and **H** (layer *b*) are 1.45 and 2.30, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 1 μm. **M** represents a titanium layer (*dark line in Fig. 8-39*) of 14 nm metric thickness.

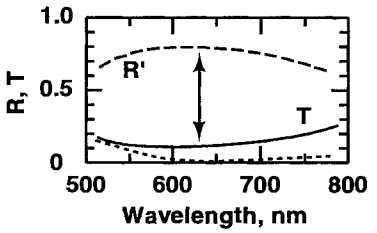


Fig. 8-41 $T(\lambda)$ (*solid*), $R(\lambda)$ from the obverse side (*short dash*) and $R(\lambda)$ from the reverse side (*long dash*) of
air **L** 0.58**H M** 0.85**L H L H** *glass*
where the indices of **H**, **L** and *glass* are 2.30, 1.52 and 1.45, respectively. The optical constant of **M** is $2.76-j\,1.80$ and its metric thickness is 20 nm. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 633 nm.

8.6.2 Reflection filter

This coating consists of alternate layers of metal and dielectric. As shown in Fig. 8-42, the position of the node of the standing wave electric field determines the wavelengths at which it reflects. Although an all-dielectric stack could produce a similar reflection spectrum, the advantage of the metal-dielectric stack is its relatively few number of layers. An example is Fig. 1-141, which depicts the $R(\lambda)$ of

air **M L** *aluminum* .

Schmitt [66] used more than one metal-dielectric period in a coating that absorbed in the visible spectral region and reflected the IR. Stelmack [83] extended the coating to the ultraviolet and Berning [87] used it to authenticate a document.

8.6.2.1 Standing waves in a reflection filter

Leeb [76] and Anthony [86] used intracavity reflection filters in lasers to isolate spectral lines. At the wavelength of maximum reflectance the node of the standing wave is at the center of the metal layers, as shown in Fig. 8-42. This minimizes the loss in the same way that it does in the One-M bandpass filter described in §8.2.4.5. The higher loss at other wavelengths reduces the reflectance.

Figure 1-142 shows the axial modes of an injection laser and also the spectral reflectance of a single-period mirror that was at one end of the cavity. Its reflectance was sufficiently low such that only one of the modes lased (arrowed in Fig. 1-142).

8.7 Beamdivider containing silver

The general form of this coating is a semitransparent layer of silver surrounded by dielectrics. It is immersed (see §1.8.1.1) and typically functions at a 45° angle of incidence. As depicted in Figs. 1-162 and 1-163, an advantage of this coating is that the polarization splitting is minimal in both the reflected and transmitted beams. A

comparison of it with the coatings in §6.5.5.2 illustrates that the range of angles over which this coating will perform is superior to that of an all-dielectric beamdivider. The disadvantage is that its *RT* product [see §1.8.1.2(g)] is approximately 21%, in comparison to the 25% *RT* product of an all-dielectric system.



Fig. 8-42 Standing wave electric field of a double-period reflection filter at the wavelength of maximum reflectance (*right*) and off wavelength (*left*). Metal layers are solid black and dielectric spacer layers are shaded. After Anthony [86].

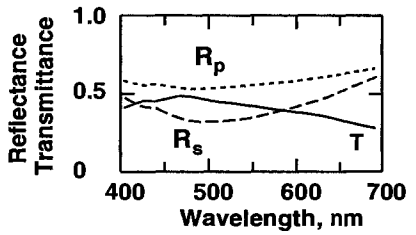


Fig. 8-43 At 45° incidence in *cement*, $R_p(\lambda)$ (*short dash*), $R_s(\lambda)$ (*long dash*) and $T_a(\lambda)$ (*solid*) of
cement **H' M H'** *glass*
where the indices of *cement*, *glass* and **H** are 1.517, 1.517 and 2.25, respectively. The optical constants of the **M** layer appear in Table 8-6. The metric thicknesses of the **M** and **H'** layers are 30 nm and 46 nm, respectively.

The simplest form of this beamdivider contains only a single silver layer sandwiched between high index matching layers

cement 0.8**H M** 0.8**H** *glass*

where the optical *cement* has the same refractive index as the *glass*. The **M** silver layer is 30 nm in thickness. Figure 8-43 shows that its average reflectance is close to 50% and that polarization splitting (see §6.1.2.5), while not zero, is minimal.

8.8 Neutral density coatings

The purpose of the neutral density (ND) coating is to attenuate the incident flux. In its simplest form, it consists of a single metallic layer such as a nickel-chromium alloy [Banning 47]. Its thickness is adjusted so that a prescribed transmittance is attained. The design and production of a neutral density are more difficult when additional specifications are imposed, such as: (1) It must withstand a high irradiance; (2) its absorbance is tightly controlled (example: 2±0.05); (3) its reflectance must be quite low; and (4) its absorbance vs wavelength is essentially constant over an extended bandwidth. Each of these topics is discussed separately.

8.8.1 Effects of inserting the coating into a beam of high irradiance

The high irradiance (i.e., radiant flux per unit area) of a laser beam can destroy quite easily a neutral density when the absorbed flux overheats a small area. The overheating is reduced when a substrate of relatively high thermal conductivity, such as silicon, diamond or sapphire, is used. It is advantageous to use platinum, rhodium or other refractory metals that do not oxidize at higher temperatures.

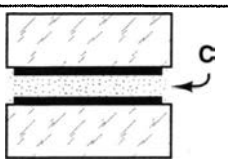


Fig. 8-44 A composite neutral density (ND) filter is constructed by joining separate ND coatings (*solid black)* with an optical cement (*dotted*).

8.8.2 Precise control of absorbance

Figure 8-34 shows the statistical variations in transmittance that occur when single layers of tantalum are deposited. One obvious method of increasing the yield is to tighten the deposition process.

Figure 8-44 shows that a composite neutral density is fabricated by cementing two separate neutral density coatings "back-to-back." The advantage is that the neutral density coatings of several coating runs may be sorted by absorbance — thereby increasing the yield. A caveat is that the absorbances do not add if the individual coatings are reflective (§10.3.5.1 has an example). A composite neutral density would not function properly if it were irradiated with a laser. The reason is that the coherence length of the laser usually exceeds the thickness of the optical cement (see §10.3.2.2).

8.8.3 Coating with relatively low reflectance

Two designs with relatively low reflectance are: (1) Multiple pairs of dielectric and metal layers; and (2) the gradient metal.

The design of the "multiple pairs" neutral density is

$$air\ \ \mathbf{D}\ (\mathbf{M}\,\mathbf{D})^{q}\ \ glass$$

where **M** is the metal layer and **D** a dielectric. The neutral density described by Fay [73] is an alloy of nickel, chromium and iron that was sandwiched between layers of silicon monoxide. One approach is to deposit the outer **M** layers so thin (example: 3 nm) that they are island films. Such layers usually have a low reflectance.

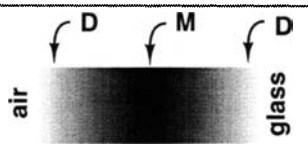


Fig. 8-45 Construction of a gradient metal neutral density in which the concentration of the metal (*dark shading*) peaks at the center *M* and is minimal at both ends, where there are dielectric materials *D*. After Anders [67a].

The multiple pairs described above have the advantage that they may be deposited in a conventional coating chamber. The gradient metal coating shown in Fig. 8-45 requires co-evaporation similar to the apparatus of Anders [67a]. The deposition begins with a pure dielectric. The concentration of the metal increases, reaches a maximum and then decreases until its concentration is nearly zero at the outer interface. Figure 8-46 shows that the reflectance of such a coating is quite low, especially from the reverse side. Figure 8-47 shows the transmittances of coatings of several thicknesses.

8.8.4 Constant absorbance vs wavelength

It is difficult to achieve an absorbance that is spectrally "flat." It is serendipitous, for example, that the dashed curve in Fig. 8-47 is spectrally flat in the near IR. One method of flattening the transmittance curve is to modify the reflectance by altering the thicknesses of one or more dielectric layers on top of the stack. Another approach is to combine two or more metals.

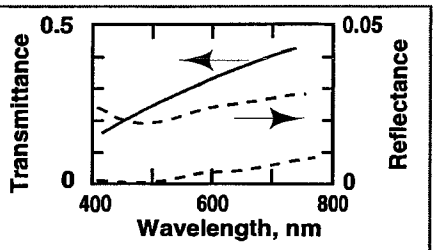


Fig. 8-46 $T(\lambda)$ (*solid*, left ordinate), obverse reflectance (*upper dashed*, right ordinate) and reverse reflectance (*lower dashed*, right ordinate) of a few of the coatings captioned in Fig. 8-45. After Anders [67a].

As an example, Fig. 8-48 shows that in the near IR the transmittance of an alloy of nickel, chromium and iron is increasing with wavelength, whereas the opposite is true for platinum. A spectrally flat transmittance is obtained by using both metals in the neutral density—thus balancing the transmittance increase of the former with the decrease of the latter.

8.9 Miscellaneous topics

8.9.1 Amplitude reflection coefficient of an overcoated matching stack

Consider a single metal surrounded by symmetrical matching stacks. Given that the dielectric matching stacks shown in Fig. 8-4 produce Ψ_{max}, the proof that $\hat{\rho}_2^* = \hat{\rho}_3$ follows. (1) Due to the symmetry of the stack, the amplitude reflection coefficients $\hat{\rho}_1$ and $\hat{\rho}_3$ are identical. (2) Following the notation of Fig. 8-4, the One-M filter is partitioned into the upper antireflection dielectric stack and the partial filter containing the metal layer (see §8.2.4.2.1). (3) Equation (2-45g) is rewritten

$$T_f = \frac{T_1 T_2}{\{1 - R\exp[-j(\delta_1 + \delta_2)]\}\{1 - R\exp[j(\delta_1 + \delta_2)]\}} \tag{8-3a}$$

where T_f is the transmittance of the One-M filter, T_1 is the transmittance of the upper antireflection dielectric stack—from the incident medium into the space between the partitions, T_2 is the transmittance partial filter—from the space between the partitions into the emergent medium. R is defined as

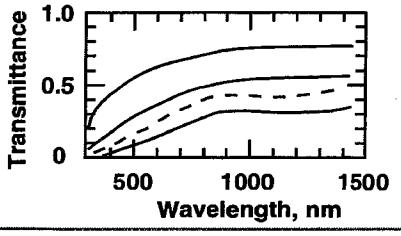


Fig. 8-47 Spectral transmittance of coatings captioned in Fig. 8-45. After Anders [67a].

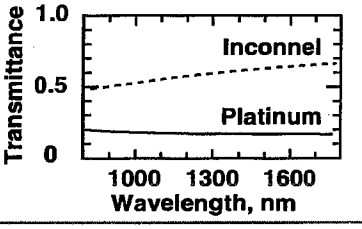


Fig. 8-48 Measured $T(\lambda)$ of a thin layer of platinum (*lower*) and an alloy of nickel, chromium and iron (*upper*). Coursen [92] sputtered the platinum. The alloy was thermally evaporated at Coherent [92].

$$R \equiv \hat{\rho}_1^* \hat{\rho}_1 \equiv \hat{\rho}_2^* \hat{\rho}_2 \tag{8-3b}$$

and is the reflectance of either the metal-layer side of the partial filter or the antireflection stack.

(4) If we *assume* that $\hat{\rho}_2^* = \hat{\rho}_1$, it follows that $\delta_1 = 2\pi - \delta_2$ and Eq. (8-3a) reduces to

$$T_f = \frac{T_1 T_2}{(1-R)^2} \,. \tag{8-3c}$$

(5) The antireflection stack is all dielectric and hence its reflectance $R_1 = R$ is the same from both sides. This leads us to rewrite Eq. (8-3c) as

$$T_f = \frac{T_1}{1-R_1}\ \frac{T_2}{1-R} = \Psi_1 \Psi_2 . \qquad (8\text{-}3d)$$

(6) The first term on the right of Eq. (8-3d) is the Ψ_1 of the antireflection stack, which is unity (see §8.2.4.2.2). (7) The second term is the Ψ_2 of the partial filter which, according to Eqs. (8-1a) and (8-1b), is the Ψ of the complete One-M filter. (8) It follows that the assumption in step (4) is consistent. Because $\hat{\rho}_1$ and $\hat{\rho}_2$ are complex conjugates, a consequence is that $\hat{\rho}_2$ and $\hat{\rho}_3$ are also conjugates.

8.9.2 Optical constants of titanium

Table 8-10 lists the optical constants of titanium metal [Weaver 81]. This is a component of a dark mirror.

Table 8-10 Optical properties of titanium

λ, nm	*n*	*k*	λ, nm	*n*	*k*	λ, nm	*n*	*k*
1033.1	3.35	3.3	1589.5	4.11	3.85	2530.1	4.29	3.15
1078.1	3.4	3.35	1631.3	4.18	3.87	2582.9	4.11	3.1
1127.1	3.47	3.4	1675.4	4.25	3.88	2637.8	3.89	3.08
1180.7	3.54	3.46	1721.9	4.32	3.89	2695.2	3.62	3.12
1239.8	3.62	3.52	1771.1	4.38	3.89	2755.0	3.35	3.25
1265.1	3.65	3.55	1823.2	4.44	3.9	2817.7	3.1	3.44
1291.4	3.69	3.57	1878.4	4.52	3.9	2883.2	2.88	3.68
1318.9	3.73	3.6	1937.1	4.6	3.88	2951.8	2.69	3.98
1347.6	3.77	3.62	1999.6	4.66	3.83	3023.8	2.56	4.32
1377.5	3.8	3.65	2066.3	4.71	3.77	3099.4	2.49	4.68
1408.8	3.85	3.69	2137.5	4.75	3.71	3178.9	2.48	5.02
1441.6	3.93	3.73	2213.9	4.75	3.62	3262.6	2.46	5.35
1475.9	3.94	3.77	2295.9	4.73	3.5	3350.7	2.43	5.65
1511.9	3.94	3.8	2384.2	4.62	3.35	3443.8	2.41	5.93
1549.7	4.04	3.82	2479.5	4.43	3.22	3542.2	2.39	6.21

8.9.3 Historical notes

The procedures in §8.3 and §8.4 use the concepts of *net flux ratio, optimization* and *partitioning* to design bandpass filters. Many designs were extant several decades before these concepts appeared in the literature. In 1939, the pioneer Geffcken [42] fabricated a Two-M bandpass filter—a Fabry-Perot with a solid spacer—by evaporating both the magnesium fluoride spacer and silver layers. This was later studied by Hadley [47]. Several publications [Dufour 50, Hermansen 55, Wolter 56, Kitchens 74] describe methods of designing two cavity and three cavity bandpass filters. Schroeder [62] was particularly successful in applying these concepts to the ultraviolet. Although Berning [57] hinted that dielectric matching stacks could be added to a MDM filter to enhance its peak transmittances, Maier [67] was the first to publish both a theoretical analysis and experimental data on such filters.

Chapter 9
Coating deposition

9.1 Introduction

Reviews such as Bunshah [82] and Pulker [84] describe numerous methods of laying down films, such as chemical deposition, sputtering and vacuum evaporation. It is certain that chemical vapor deposition and plasma chemical vapor deposition are destined for widespread use in future optical coatings, although their use is in its infancy now. Excluding the coatings on architectural glass, most are presently deposited by vacuum evaporation. It is for this reason that this chapter focuses on this method.

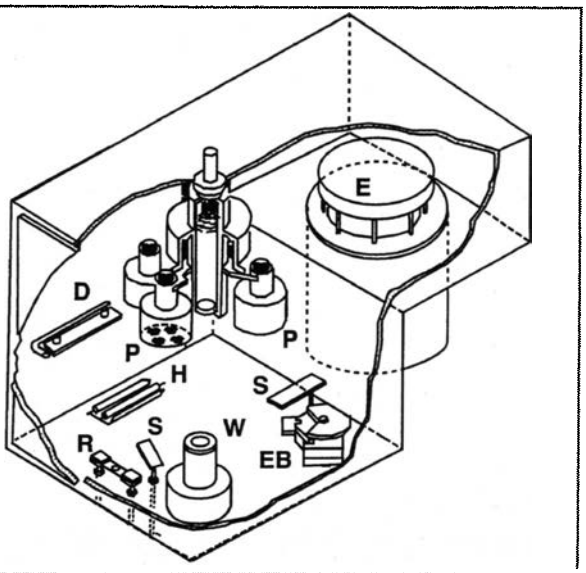


Fig. 9-1 A box coating tank for thin film deposition, showing pumps to evacuate the gas from the tank *E*; planetary racks holding substrates to be coated *P*; radiant heaters *H*; window to permit optical monitor probe beam to enter the system *W*; mechanical shutters S; electrically heated source *R*; insulated electrode for glow discharge *D*; and electron beam gun *EB*. After Morton [77].

9.1.1 Overview of the vacuum evaporation process

Figures 3-2 and 9-1 depict vacuum tanks in which films are deposited. Although the cost and level of sophistication can vary widely, a generic procedure common to many chambers and deposition processes can be listed:

1. Prepare the chamber for deposition. Remove any flakes of coating materials from the chamber walls. Load the evaporants into the resistively heated sources or electron gun crucibles.
2. Clean substrates and load them into the chamber.
3. Seal the chamber and activate pumps that remove air until the pressure is 10^{-4} Torr (13 mPa) or lower. §9.3.2 discusses the importance of the rate at which air is removed from the chamber.
4. Start deposition of the first layer—its thickness may be controlled by any of several methods. Terminate deposition when a predetermined thickness is attained.
5. Continue deposition of additional layers. As few as one and as many as 100 may be deposited.
6. At the end of deposition, re-admit air to the chamber and remove substrates.
7. Evaluate the performance of the coating by measuring its transmittance on a spectrophotometer. Subject the coating to environmental testing, if required. These topics are discussed in Chap. 10.

Further details are given in the remainder of this chapter.

9.2 Cleaning of a substrate prior to coating it

There are at least two reasons to clean a substrate: (1) contaminants (such as organic materials on glass) are removed that cause the coating to adhere poorly to the substrate; and (2) particulates and other foreign matter are removed that reduce the performance of the coating. Several examples are: (1) dust that engenders pinholes in a bandpass filter and permits stray light to leak through it; (2) contamination that increases the scatter of light by the coating; and (3) defects in the coating that degrade the laser damage threshold..

The next question: What is "clean?" As several reviews [Adams 79, Mattox 78, Karow 82] emphasize, there is no single answer. The selection of the cleaning process and condition of the substrate after cleaning depend upon the substrate composition, the intended coating and its ultimate use. Karow [93] devotes more than six pages to the cleaning of optical elements, including details on an ultrasonic "clean line."

With the advent of semiconductor integrated circuit production, it is possible to purchase a sophisticated cleaning line contained in a class 100 clean room. Only a few optical coatings, such as laser mirrors, however, are costly enough to merit such treatment. Regardless of whether the cleaning is sophisticated or a simple scrub with calcium carbonate, many consider cleaning an art rather than a science. In my experience, there are at least three essential steps in the cleaning process: (1) a *scrub* that removes particulates mechanically from the substrate; (2) immersions and rinses in several liquids; and (3) drying and removal of the liquid. Each is discussed separately.

9.2.1 Removal of particulates on the substrate by mechanical means

Two methods of mechanical particulate removal are: (1) the surface is scrubbed with a moistened pad; and (2) a high pressure spray is directed to the surface. As Stowers [78] has pointed out, the spray exerts a shear force that removes particles whose size is below a micrometer. Ultrasonic cleaning may loosen larger particles, but is ineffective in the removal of the smaller particles that tenaciously adhere to the substrate.

9.2.2 Cleaning of the substrate in liquids

9.2.2.1 Purification of liquids

Regardless of which solvent is used in the process—whether it be water, isopropanol, acetone or others—it is important that it be continuously purified. The most widely used purification method is distillation, usually in a degreaser type of tank.

9.2.2.2 Cleaning and rinsing solutions

Many aqueous cleaning solutions are marketed under the generic term *detergent* (an is example Neodol [75]). The ingredients of the detergent usually include:

- a surfactant—this is a wetting agent;
- an emulsifier—this suspends grease, oil; and
- a sequestrant—this inhibits precipitation.

In addition, the detergent may contain a base such as trisodium phosphate that attacks grease. Many glasses are etched by a basic solution, however.

The substrate is rinsed several times in de-ionized (DI) water after its immersion in the detergent. At this point, there are several options on what to do next. Many prefer water for the final rinse (a warning: ultra-pure water etches many types of glass!). Others are partial to isopropanol, methanol or acetone—or azeotropic mixtures of these and

similar solvents. Methanol vapor is a health hazard. Acetone has several disadvantages. First, it cools as it evaporates, thus causing water vapor in the air to condense as a fog of water droplets—each of which may contain a dust particle. The explosiveness of acetone vapor is a safety hazard.

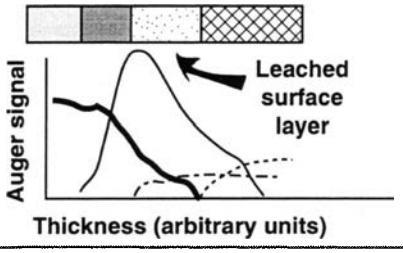


Fig. 9-2 Depth profiling with Auger electron spectroscopy of Schott BaK4 glass coated with two layers. The glass was etched for 50 s in a pH = 5 acidic solution at 50 °C. Thick solid black: signal from an **L** layer. *Thin solid black*: signal from an **H** layer. *Long-short dash*: signal from silicon. *Short dash*: signal from barium. The *cross-hatched* area is the glass substrate. After Günther [81a, 81d].

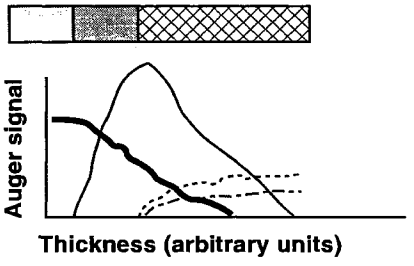


Fig. 9-3 The caption to Fig. 9-2 obtains, but the glass was not etched.

9.2.3 Drying of the substrate

The most important step in the process is to remove the liquids from the surface so that no drops remain upon it. Any drops that evaporate leave small spots of residue. A properly designed degreaser tank does not leave drops.

At the University of Rochester we used a costly but effective method of cleaning fused silica substrates. An immersion in warm chromic acid that removed organic contaminants was followed by a rinse in DI water. After removal, the disk was tipped and the water driven to one edge with a gentle stream of oil-free ionized nitrogen (the ionizer was produced by [Simco]). If the surface was truly clean, interference fringes were observed in the light reflected in the thin film of water. A clean surface permitted the water to drain in a uniform sheet. No drops remained on the surface.

9.2.4 Chemical etching of the substrate surface

The chemicals used in a cleaning process may degrade an optical surface. Günther [81a, 81d], for example, coated Schott BaK4 glass with several nonabsorbing layers. Prior to coating, the polished glass was immersed in an acidic solution (of pH = 5) at 50 °C. The glass surface was etched during its brief sojourn of 50 s in the bath. This is illustrated by Fig. 9-2, which depicts the Auger electron signal as the coated glass was depth profiled by sputter etching. This shows the presence of both silicon and barium in the bulk of the glass, but the absence of barium in a physical region near the surface. The interpretation is that the barium was etched chemically from this surface region. Figure 9-3 shows the presence of both silicon and barium in a glass sample that was not immersed in acid.

9.3 Tooling, initial pumpdown, ion bombardment and heating of substrates

After the substrates are cleaned, they are loaded into the tooling. The chamber is then pumped. Ion bombardment is often used to further "clean" the substrates. The heating of the chamber is also important.

9.3.1 Tooling — introduction

Tooling is the generic term for the substrate holders, as, for example, the rack in Fig. 3-2 or the circular plate in Fig. 9-25. A synonym for *tooling* is *jigging*. There are instances in which tooling is not needed. It is possible, for example, to configure evaporation sources so that they evaporate downwards; the substrates sit on a plate at the bottom of the chamber. A large circular telescope mirror is often held in a strap around one-half of its circumference. Once it is inside the vacuum tank, the polished surface, which is vertical, is coated with aluminum that evaporates in a quasi-horizontal direction from the filaments. Films can be sputtered in any direction.

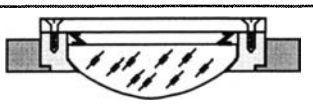


Fig. 9-4 A jig that holds a lens while multilayers are vacuum deposited on its bottom surface. After Matteucci [80].

The advantage of evaporating (or sputtering) upwards is that many fewer of the particles dislodged during the coating deposit on the surface. Fig. 9-4 depicts tooling that holds a lens during deposition [Matteucci 80]. There is a plate that prevents the rear of the lens from being coated and also a spring that prevents it from vibrating or shifting as it is rotated during the deposition. The lens is supported by the jig lip, whose width may be determined by any of several factors. First, the rim of the lens should be free of any coatings so that it mounts properly in the lens holder, as shown in Fig. 1-208. The thermal expansion of the tooling during the heat cycle in the chamber is also of concern. An ultra-narrow jig lip is dangerous because it concentrates a substantial amount of weight over a relatively small area. This highly stresses the substrate and can cause it to chip. This is especially true if there is excessive bevel, as shown in Fig. 1-205. A conservative guide for the jig lip width (see Fig. 1-207) is taken from a Mil. Spec. That is, if the optical system designer specifies this jig lip width, there is little danger that the periphery of the lens or planar substrate will chip in the tooling during deposition.

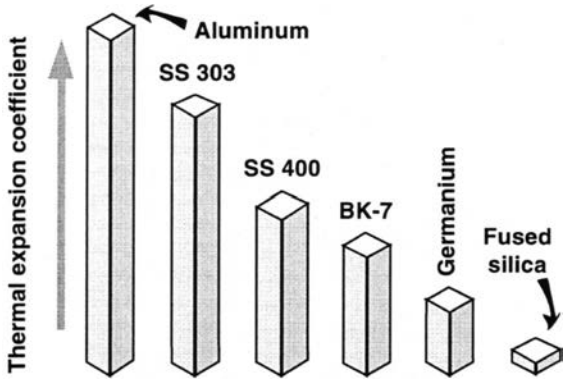


Fig. 9-5 Linear thermal expansion coefficients of metals used in tooling, as well as substrates upon which optical coatings are deposited. SS refers to stainless steels.

9.3.1.1 Choice of metals for tooling

The choice is based on factors such as the ease with which it is machined, the degree to which it outgases in a vacuum, hardness, its thermal expansion and its thermal conductivity. Both aluminum and stainless steel are widely used. Aluminum is easier to machine, although it outgases much more than stainless. Another advantage of stainless is that it may be cleaned chemically. The softness of aluminum is often a problem when it holds glass with an unbeveled edge. At elevated temperatures the aluminum expands more than the glass.

Fig. 9-5 graphs the relative linear expansion coefficients. When the tooling cools, the glass digs into the relatively soft aluminum lip and fractures. Still worse, glass sometimes welds to aluminum at elevated temperatures.

The thermal conductivity of the tooling is of concern because both it and the substrates are heated radiantly during deposition. The conductivity influences the substrate temperature. Both the thickness and index of the coating are temperature dependent. As shown in Fig. 9-4, a spring-loaded clip on the rear of the lens ensures good thermal contact with the tooling and that the thermal conduction from lens-to-tooling is consistent.

9.3.2 Pumpdown of the vacuum coating tank

The gas is evacuated from the coating chamber during the pumpdown. §9.4.8 discusses the pressure to which the chamber should be evacuated. The coating tank is initially at atmospheric pressure and contains air, with its usual fraction of water vapor. The "cloud chamber effect" occurs when the coating tank is evacuated too rapidly. The reduction of pressure causes the air to cool adiabatically, which in turn induces the water vapor to condense out as small droplets of liquid. These droplets eventually re-evaporate, of course. However, each droplet contains one or more dust particles, which remain on the surfaces inside the chamber (including the substrates) after the droplets evanesce. This "cloud chamber effect" is avoided by lowering the rate of pumpdown. This is often termed a "soft pumpdown" in the trade.

9.3.3 Exposure of the substrate to a gas discharge

The substrates are sometimes bombarded with ions after they have been cleaned, loaded into the tooling and mounted in the racks in the evaporation chamber. Figure 9-6 depicts two methods of bombardment: (1) The ion gun; and (2) a gas discharge.

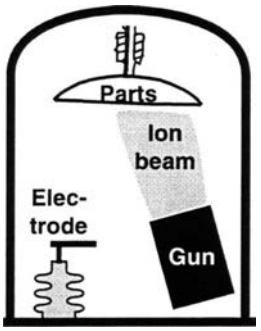


Fig. 9-6 Apparatus for the ion bombardment cleaning of parts to be coated, including an ion gun (ion beam is shaded) and a high voltage electrode for the introduction of a gas discharge.

9.3.3.1 Ion gun as source of ions

This device (see §9.4.5.3) directs a stream of ions—usually argon—at the substrates. The energy of these ions removes water vapor or organic materials from the surface, thereby enhancing the adhesion of the evaporant. It is always possible that the ion beam can impinge upon the tooling and sputter metal—which then deposits upon the substrates. Because the sputter yield is smaller at low voltages, a voltage of 100 V is typical.

9.3.3.2 Immersion of the substrates into a gas discharge

The gas discharge produces light—a *glow*—hence, this is often termed "glow discharge treatment." An electrode (see Fig. 9-6) is maintained at a potential of a few kV and a gas discharge ignites when the gas pressure is in the range of 5 mTorr (0.6 Pa) to 50 mTorr (6 Pa). The gas discharge enhances coating adhesion. Its detracting features are that it adds time to the coating cycle and also deposits particulates upon the substrates. For this reason, laser mirrors are infrequently subjected to a gas discharge.

Two theories have been advanced to explain the improvement of adhesion: (1) the ionic bombardment removes the layers of water vapor or organic materials from the substrates; or (2) the gas discharge sputters onto the substrates a layer of *something*, which serves as an adhesor (see §3.5.7). The thin film deposited later bonds well to this adhesor. The *something* could be the garbage residing on the surface of the glow discharge electrode, such as a medley of materials deposited recently.

9.3.3.2.1 Polarity and type of gas used in the gas discharge

The remaining question: What discharge polarity and gas should be used? Holland [58] ran the glow discharge electrode positive and thus directed ions at the substrates. Using either air or argon as the residual atmosphere, he found that hydrocarbons — even silicones — were removed effectively from glass. The world's two largest producers of coatings use the opposite polarity, thus bombarding the substrates with negative ions and electrons. This undeniably heats them, thereby enhancing adhesion. Holland [58] claims that the electron bombardment is undesirable because it decomposes and partially carbonizes any organic materials on the substrates. Colbert [43], on the other hand, deliberately coated glass with an oily substance that he later removed by a 3 kVA ac glow discharge. His coatings adhered well. Regardless of which method is used, the electrode should be composed of a metal with a low sputter yield such as pure aluminum [Benson 50] or a magnesium alloy [Turner 45]. The electrode should be cleaned chemically prior to use. A sparking along the aluminum electrode during the glow means that an insulating oxide layer is forming on its surface. Each spark is a rupture of that layer — and a reminder that the rod needs cleaning.

9.3.4 Heating of the chamber and substrates

The chamber should be heated to outgas the walls. There are several reasons that the substrates are heated. First, the film-substrate adhesion and the mechanical hardness of the layers are improved. Second, the deposition of an oxide layer at higher temperatures reduces its optical absorption.

9.3.4.1 Heating of the chamber

The principal reason to heat the "walls" of the chamber is to bake the gas — mainly water vapor — out of them. Why the quotation marks on "walls"? There are "ultra high vacuum" coating chambers in which it is sealed with metal gaskets. The entire enclosure may be baked at an elevated temperature. Here, the "walls" may truly be the walls of the vacuum tank.

In most of the coating tanks, however, the sides, top and bottom of the chamber are water cooled. There are shields interposed between those cooled surfaces and the interior of the chamber. Their purpose is to prevent the interior of the chamber itself from being coated by the evaporants. It is important that these shields be thermally insulated from water-cooled outer walls. This thermal isolation enables them to be baked to high temperatures — thus outgasing them thoroughly. These shields are often sheets of stainless steel or composed of aluminum foil.

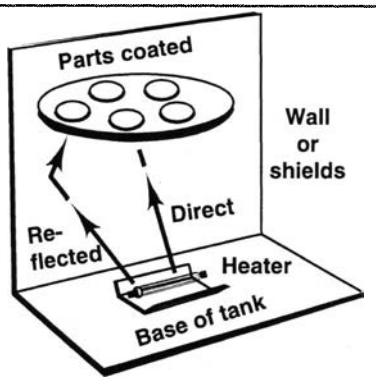


Fig. 9-7 Radiant energy emerging from a heater on the base plate of a coating chamber reaches the parts being coated both directly and by reflecting off the shields near the chamber walls.

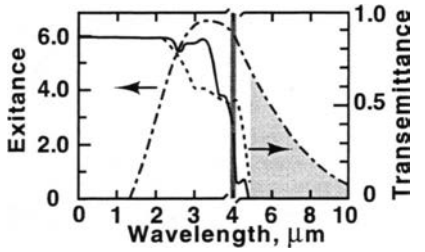


Fig. 9-8 Exitance (*long-short dash*, left ordinate) per unit wavelength (mW/mm^2 μm) of a Planckian source at 600 °C. Transmittance (right ordinate) of fused silica (*solid*) and soda-lime glass (*short-dash*). The shaded area is the onset of absorption. The abscissa changes scale at 4 μm.

9.3.4.2 Transfer of radiant heat from the heaters to the chamber

Heat is transferred to the substrates or shields either conductively or radiatively. It is possible to heat by thermal conduction if a gas at a pressure of 0.1 Torr or higher is admitted to the tank—but no evaporation can take place. Radiant heat is more widely used. At one time it was widespread practice to heat radiantly from the rear of the substrates. While this has a few advantages, it is more common to heat the side to be coated. Fig. 9-1 shows, for example, the heater assembly *H* at the base plate consisting of tubular incandescent sources backed with a radiation shield.

The design parameters considered in the installation of a heater are: (1) Its physical location relative to the substrates; (2) the heater area and temperature; and (3) the fraction of the radiant heat absorbed by the substrates. Each is discussed separately.

The heat transfer is determined by the heater-substrate *view factor*. This term is discussed in texts on radiant heat transfer. The larger this view factor, the more efficient the heat transfer. This is optimized by placing the heaters in proximity to the substrates, but not too close lest the heaters shadow the stream of evaporant or heat the substrates nonuniformily. Heat reflected from the walls of the chamber may contribute to the *view factor*. Figure 9-7 shows, for example, that radiant energy travels directly from the heaters in the bottom of the chamber to the parts being coated. A portion of the energy, however, may reach the parts by reflecting from the shields near the chamber wall. Suppose that the shields become coated with thick oxide layers. Because such oxides are usually absorbing in the IR, the net effect is to reduce the reflectance of the shields, thus reducing the view factor.

9.3.4.2.1 Efficiency of transfer of radiant heat

The efficiency of heat transfer may be optimized by adjusting the heater temperature so that the substrates absorb the largest fraction of its output.

The total radiant flux (i.e., radiant power) Φ a heater emits is

$$\Phi = \varepsilon S \mathcal{A} \mathcal{T}^4 \qquad (9\text{-}1a)$$

where ε is its hemispherical emissivity, S a constant, $\mathcal{A}$ its area and $\mathcal{T}$ its temperature in K. A trade-off between area and temperature is evident. Large-area low-temperature heaters can match the flux output of small-area high-temperature sources.

Glass is heated most efficiently by low-temperature sources. The reason this is true is related to the relative amount of radiation the heater emits at different wavelengths and the fraction of that radiation actually absorbed by the glass. The distribution of the radiation at various wavelengths depends upon the temperature of the heater and can be calculated from the Planck Radiation Law. This is illustrated by Fig. 9-8, which shows the flux per unit wavelength from a Planckian radiator at 600 °C. Superimposed is the transmittance of a glass and a fused silica substrate. Only the radiation in the absorbing region (*shaded*) is truly effective in heating. At this temperature, only 20% of the total radiant flux is at wavelengths shorter than 3 μm where the substrates are nonabsorbing. A higher temperature heater dumps more radiation into the nonabsorbing short wavelength part of the spectrum at the expense of reduced efficiency.

The foregoing arguments are reversed when germanium is heated. This material is absorbing at wavelengths less than 1.8 μm, as shown by the shaded area in Fig. 9-9. A relatively hot Planckian source at 2500 K radiates 58% of its energy in the spectral region where the germanium absorbs it. A cooler source radiates even less in that region; hence, a high temperature source is more efficient.

9.3.4.3 Heaters and the measurement of the temperature

A wide variety of heaters is available. Nichrome heats glass effectively to 150 °C. A ceramic heater consists of a tungsten wire embedded in a ceramic plate [Sylvania 85]. It functions at approximately 600 °C and is ideal for heating glass. A tungsten wire encapsulated in a quartz tube (hence the name *quartz heater*) operates at much higher temperatures.

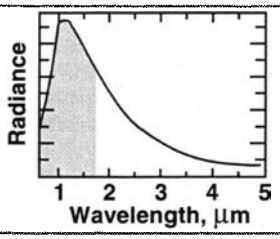


Fig. 9-9 Radiance per unit wavelength interval (arbitrary units) of a Planckian source at 2500 K. The shaded area is the onset of absorption for bulk germanium.

The longevity of any heater is increased by protecting it from burn out and contamination. Its power rating is quoted for ambient air, where typically 10% to 15% of its heat is dissipated nonradiatively. This nonradiative loss is smaller in a vacuum; hence, a heater tends to burn out unless operated at reduced power.

Heaters are often contaminated by evaporant vapor or spatter of molten evaporant. The latter burns a hole instantly in a ceramic heater. Heaters can be protected with a plate of high-silica glass or strips of sapphire. The latter is transparent to wavelengths near 7 μm and thus minimally reduces the efficiency.

Thermocouples, thermistors or radiometers are used to measure the temperature and regulate the heaters. A sensor is often encapsulated in glass so that its temperature follows more closely that of the glass substrates being coated. A more sophisticated approach is to deposit a thin film thermocouple on the surface of the substrate. Its actual surface temperature is thereby measured, which might be 100 K higher than the bulk substrate during deposition of the film, due to the heat of condensation surrendered by the impinging evaporant. A radiometer is an effective way to measure surface temperature, provided the substrate is opaque in the spectral region of its probe beam. The installation of a radiometer is expensive because it requires that a calcium fluoride window be installed in the wall of

the coating machine. A simple method of measuring the temperature of the substrates is a temperature gage that contains a bimetallic strip.

9.4 Thin film deposition

9.4.1 Introduction

§3.1 cites at least five methods of thin film deposition. Our attention is confined to two: (1) vacuum evaporation; and (2) ion beam sputtering. The first is cited because of its widespread use and the second because coatings of such unique qualities have been deposited. Basic principles are reviewed in §9.4 before the equipment and deposition processes are described.

9.4.1.1 Mean free path of a gas

The mean free path m_f is the mean distance an atom or molecule travels before colliding with another. It depends upon the collision cross section of the molecule and the gas density, which, at a given temperature, is inversely proportional to the pressure P. For air at room temperature (21 °C), an approximation is

$$m_f \approx 50/P \approx 0.6/P' \qquad (9\text{-}1b)$$

where m_f is in mm, P in mTorr and P' in Pa.

9.4.1.2 Effusion of gas from a Knudsen cell

This cell is an enclosure at a uniform temperature $\mathcal{T}$ containing gaseous evaporant whose vapor pressure is P. Molecules of this vapor effuse into the surrounding vacuum through an orifice of area $\mathcal{A}$ in the wall of the cell. The model assumes that pressure exterior to the cell is sufficiently low that the mean free path exceeds the dimensions of the cell. There are other restrictions discussed in Dushman [63] or Glang [70]. Several properties of this cell are of interest to a thin film technologist: (1) the molecular flux that emanates from it; and (2) the angular dependence of its molecular intensity.

9.4.1.2.1 Molecular flux Φ

This term is used in the same sense as in radiometry. The molecular flux Φ is the number of molecules effusing from the orifice per second, regardless of the direction at which they emerge. From kinetic theory, this is

$$\Phi = \frac{3.5\times10^{26}\, P\, \mathcal{A}}{\sqrt{M\,\mathcal{T}}} \qquad (9\text{-}1c)$$

where $\mathcal{A}$ is in m^2, $\mathcal{T}$ is its temperature in kelvin and M is the molecular weight of the evaporant (dimensionless). When P is in Pa, 3.5×10^{26} is replaced by 2.63×10^{24}. The flux in Eq. (9-1c) is proportional to the probability that a molecule escapes through the orifice. The larger its area, the higher the escape probability; hence, $\Phi \propto \mathcal{A}$. As P increases, the physical density of the evaporant vapor inside the cell increases and more molecules escape; hence, $\Phi \propto P$.

9.4.1.2.2 Molecular intensity

"Molecular intensity" is a term borrowed from photometry and is the number of molecules per second per solid angle effusing from the source. The intensity $I(\theta)$ is sampled at a distance d and angle θ (measured to the normal of the aperture). From kinetic theory it follows that a property of the Knudsen cell is

$$I(\theta) = I_0 \cos\theta \qquad (9\text{-}1d)$$

The total flux Φ is obtained by integrating the intensity over the solid angle contained in a hemisphere surrounding the source

$$\Phi = 2\pi \int_0^{\pi/2} I_0 \cos\theta \sin\theta \, d\theta = \pi I_0 \ . \tag{9-1e}$$

9.4.1.3 Application of the molecular effusion equation

The foregoing equations are used to estimate the deposition rate upon a small area dS on a substrate that is directly above the source of area $\mathcal{A}$. Both $\mathcal{A}$ and dS are perpendicular to a line joining them of length d. The solid angle of dS is approximated by dS/d^2. Multiplying I_0 in Eq. (9-1d) by the aforementioned solid angle provides the number of molecules per second $\mathcal{N}$ impinging upon dS

$$\mathcal{N} \approx \frac{I_0 \, dS}{d^2} \approx \frac{\Phi \, dS}{\pi \, d^2} \ . \tag{9-2a}$$

From the atomic density of the film and its physical density ρ (in g/cm^3), the thickness deposited per unit time $\dot{h}$'(in nm/s) is

$$\dot{h}' \approx \frac{1.84\times10^5 \, \mathcal{A} \, P \sqrt{M}}{\rho \, d^2 \sqrt{\mathcal{T}}} \approx \frac{2.45\times10^7 \, \mathcal{A} \, P' \sqrt{M}}{\rho \, d^2 \sqrt{\mathcal{T}}} \tag{9-2b}$$

where P is in Torr, P' is in Pa, M is dimensionless and $\mathcal{T}$ is in kelvin. The physical dimensions of $\mathcal{A}$ and d^2 are m^2

The deposition of magnesium fluoride (molecular weight M = 62) serves as an example. Hacman [70] has shown that its decomposition is minimal when it is thermally (not E-beam!) evaporated. The physical density of the film is about 90% of bulk—2.8 g/cm^3. The pressure within the cell is approximated by the equilibrium vapor pressure (see §9.4.1.4) of magnesium fluoride at a temperature $\mathcal{T}$. Table 9-1 lists the deposition rate as computed from Eq. (9-2b) at temperatures at which the vapor pressure differs by a decade. The substrate-source separation of d = 0.5 m in Table 9-1 is typical of many vacuum systems. At the higher rate, a layer of 500 nm *quarterwave optical thickness* (abbreviated QWOT—see §2.7.2.1) is deposited in about three minutes, which is a comfortable span.

§9.4.5.2 describes ion-aided deposition in which both ions and evaporant molecules simultaneously impinge upon the substrate. Assuming that each molecule carries one electron charge, the ion/molecule ratio may be computed, which leads to an equation for the equivalent current density J (A/m^2) of an evaporant

$$J \approx \frac{1.78\times10^7 \, P \, \mathcal{A}}{\rho \, d^2 \sqrt{M \, \mathcal{T}}} \tag{9-2c}$$

Table 9-1 Deposition rate vs temperature

Temperature, K	Evaporant	*P*, Torr	*P*, Pa	Rate, nm / s	Current density
1428	MgF_2	0.01	1.3	0.06	0.2
1535	MgF_2	0.1 (13)	13	0.53	2.3
2293	TiO	0.01 (1.3)	1.3	0.03	0.2
2520	TiO	0.1 (13)	13	0.24	1.8

Deposition rate [using Eq. (9-2b)] and *equivalent current density* (units of A m^{-2}) [from Eq. (9-2c)] on a substrate at a distance of 0.5 m from a source of area 10^{-4} m^2 Source and substrate are perpendicular to a line joining them. The evaporants emerge from a planar Knudsen source at the indicated temperature. Vapor pressures (listed in *P* column) are from Chandrasekharaiah [67]. The physical densities of the layers of MgF_2 and TiO are 2.8 g/cm^3 and 4.8 g/cm^3, respectively.

where the symbols have the same meaning as in Eq. (9-2b). Table 9-1 lists the equivalent current density for TiO deposited at several temperatures. For example, suppose that TiO is deposited at 2520 K, which Table 9-1 shows produces an equivalent current density of 1.8 A/m^2 The conclusion is that for the ion/molecule ratio to be 1.0, an ion source must shower the substrate with a current density of 1.8 A/m^2

9.4.1.4 Equilibrium vapor pressure and its applications

Table 9-1 makes use of equilibrium vapor pressure data. This concept is now discussed. Suppose a solid or liquid element such as zinc is placed in an evacuated flask and allowed to reach an equilibrium temperature $\mathcal{T}$. Depending upon the temperature, the flask contains either: (1) liquid zinc and its vapor; or (2) solid zinc and its vapor. In either instance, however, its vapor pressure *P* is determined solely by its temperature. The situation is more complicated when a compound is vaporized. An illustration is that TiO_2 dissociates. A consequence is that the vapor over it consists mainly of TiO. Although a few materials such as B_2O_3 vaporize without dissociation, this is quite rare.

The question arises as to whether it is valid to apply this concept to an evaporant being heated inside a vacuum system. Strictly speaking, it is incorrect to infer the pressure of an evaporant from equilibrium vapor pressure, for two reasons: (1) the system is *not* at a uniform temperature; and (2) equilibrium is *not* attained because the pumps are continually removing gas from the system. As an illustration, suppose that ceria is heated in a vacuum tank similar to that shown in Fig. 9-1. The oxygen evolving from the ceria or its suboxides such as Ce_2O_3 or CeO is removed continuously by the pumps. If heated at a few thousand °C, the ceria is eventually reduced to metallic cerium because the oxygen is scavenged by the pumps. Notwithstanding such limitations, order-of-magnitude estimates of evaporation rates can still be obtained from the vapor pressure data.

From calculations such as those in Table 9-1, a rule of thumb is that a reasonable film formation rate in a typical vacuum deposition system is attained if

> the vapor pressure of the evaporant is in the range 0.1 Torr (13 Pa) to 0.01 Torr (1.3 Pa).

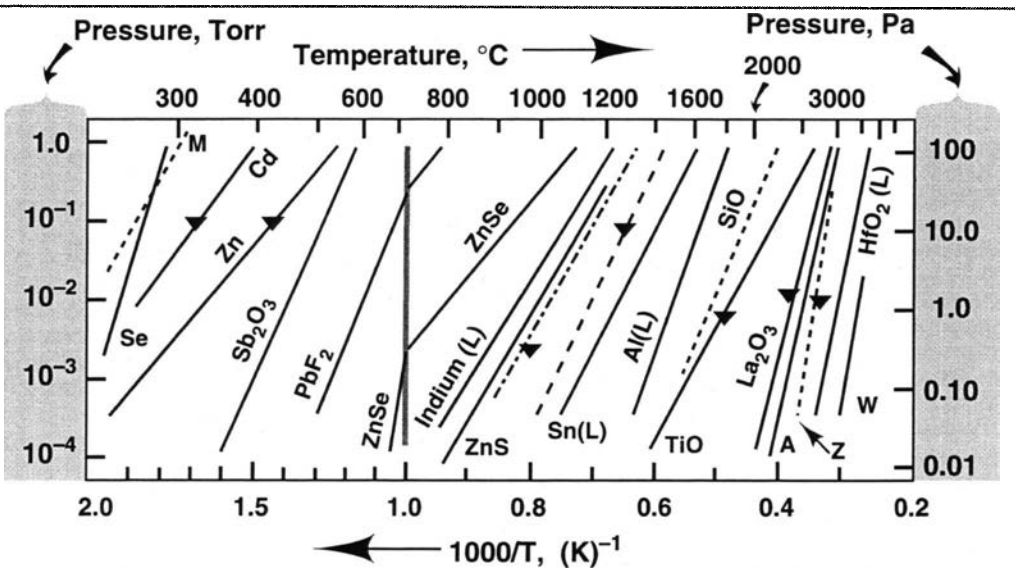


Fig. 9-10 *Ordinate*: equilibrium vapor pressure in Torr (*left*) and Pa (*right*) of elements and compounds from the data of Honig [69], Kazenas [76] and Chandrasekharaiah [67]. *Upper abscissa*: temperature in °C (nonlinear scale). *Lower abscissa*: Reciprocal temperature times 1000 on versed scale, with scale a change at 1000 K. Materials are solids except for those appended with (L). Triangles indicate melting points. Symbols on curves identify the elements or compounds, with the exceptions: *M* is polyphenylether diffusion pump oil produced by Monsanto; *long-short* is silver; *A* is alumina; and *Z* is zirconia.

A layer of quarterwave optical thickness (in the visible portion of the spectrum) is deposited in five minutes or less if the aforementioned pressure is attained. More rapid deposition is often avoided for two reasons: (1) control of the thickness is more difficult; and (2) rapidly deposited oxide layers (with oxygen bled into the chamber during deposition) are usually optically absorbing because of the formation of suboxides.

The vapor pressures of the elements are summarized by Honig [69]. Glang [70] and Balzers [84] show both tables and graphs of the vapor pressure of many materials used in optical coatings. Chandrasekharaiah [67] summarizes the vapor pressures of both fluorides and oxides. Kazenas [76] contains a plethora of information on the oxides. The aforementioned show that the log of the vapor pressure vs inverse temperature is approximated with a straight line. Fig. 9-10 shows such a plot for a few evaporants and contaminants that might appear in a coating tank. At a given temperature, the vapor pressures are the lowest for refractory materials such as HfO (i.e., partially decomposed hafnia). A metal such as cadmium has a relatively high pressure.

9.4.1.5 Gas bombardment of a surface during deposition

Another application of Eq. (9-1c) is to compute the number of molecules/second impinging upon a surface, provided the ambient pressure is known. Such a calculation shows that at room temperature and at a pressure of 10^{-5} Torr (1.3 mPa), a monomolecular layer of O_2, N_2 or H_2O impinges upon the surface in one second. The state of equilibrium determines whether these molecules remain. No net number of

molecules collect on a surface in equilibrium. That is, a molecule re-evaporates from the surface for each one that arrives. But if the surface is cryogenically cooled or chemically active (such as a freshly deposited titanium layer), each monomolecular layer remains because it either freezes out or reacts chemically.

Consider the residual gas bombardment when magnesium fluoride is deposited at a rate of 0.3 nm/s at an ambient pressure of 10^{-5} Torr (1.3 mPa). In one second, approximately one monomolecular layer of the evaporant deposits on the substrate and an *equal number* of ambient gas molecules also arrive. The predominant constituent of the residual gas is usually water vapor. A vacuum chamber at a pressure of 10^{-8} Torr (1.3 μPa) is needed to deposit a magnesium fluoride layer relatively free of water vapor.

9.4.2 Thermal evaporation

9.4.2.1 Introduction

The discussion of §9.4.1.4 establishes that the evaporant should be heated until its vapor pressure is close to 0.1 Torr in order to deposit layers of that material at a reasonable rate. To accomplish this, Fig. 9-10 shows that there is a wide range of temperatures to which the evaporant must be heated. They vary widely, from 800 °C to nearly 3000 °C for refractory materials.

An evaporation source must heat the evaporant to a temperature appropriate to that material. It is usually selected because it manifests one or more of the following attributes: (1) an unvarying molecular intensity (see §9.4.1.2)—constant not only during repeated use, but also independent of the evaporant depletion level; (2) a reasonable lifetime, especially if it is costly; and (3) it should not alloy or react chemically with the evaporant.

It is tacitly understood that, prior to their insertion into the vacuum chamber, the evaporants have been prepared by sintering and outgasing them. It is usually a disaster to heat finely ground powder in a vacuum system. The water vapor adsorbed in it is released quite rapidly by the heat, ejecting the powder violently from its container. The slang term is *popcorning.*

9.4.2.2 Resistively heated evaporation sources

Figure 9-1 shows a resistively heated source *R* at the periphery of the baseplate. A strip of metal holding a charge of evaporant is heated electrically. The shutter *S* is opened and deposition begins when a sufficient evaporation rate is attained. Three widely used resistance-heated sources are: (1) oven or other indirect heating devices; (2) boat or canoe that heats the evaporant directly; and (3) coils.

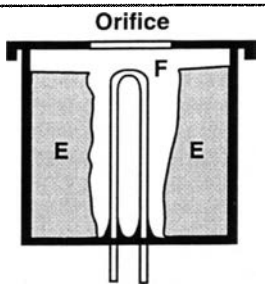


Fig. 9-11 Howitzer evaporation source in which the subliming evaporant *E* is heated by the filament *F*. The vapor emerges into the vacuum chamber through the orifice.

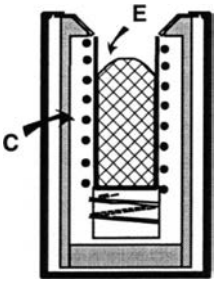


Fig. 9-12 A Luxel [83] oven evaporation source in which the crucible containing the evaporant *E* is heated by the coil *C*. The gray shaded region is radiation shielding.

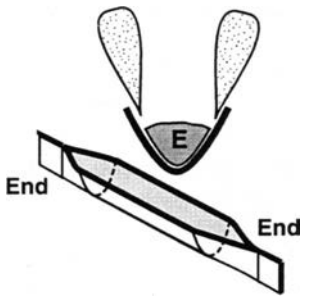


Fig. 9-13 (*Lower panel*) A canoe evaporation source is electrically heated by clamping copper electrodes onto its ends. (*Upper panel*) Cross section of canoe loaded with a subliming evaporant (*E*, *shaded*) and its molecular intensity distribution (*dotted region*) during deposition.

Holland [56], Macleod [86] and Pulker [84] discuss sources in more detail. A few are cited here. The howitzer in Fig. 9-11 indirectly heats zinc sulfide and other materials that sublime at relatively low temperatures. The oven source (see Fig. 9-12) contains a crucible with an evaporant. The crucible is heated radiantly by a coil. Radiation shields that surround it reduce its heat loss. The thermal load into the chamber is minimal with both of these sources.

The converse is true of the canoe type of source, as shown in Fig. 9-13. This source is widely used because of its low cost. It is usually composed of molybdenum, tantalum, tungsten or platinum. The platinum is essential when depositing a chemically reactive material such as lead fluoride. It is preferable that the evaporant wet the boat. The reason is that a vapor barrier forms between a nonwetting evaporant and a heated boat, similar to a drop of water on a hot stove. This inhibits heat transfer and sometimes produces an erratic molecular intensity, such as the "batwing" in the upper panel of Fig. 9-13.

9.4.3 Electron beam heating of evaporant

The chamber in Fig. 9-1 contains an electron gun in its lower right corner. A stream of electrons at voltages in the range 3 kV to 9 kV is bent 270° by a magnetic field before it impacts and heats the evaporant (see Fig. 9-14). Airco [76] discusses gun design and utilization.

The evaporant is usually contained in water-cooled copper cups or crucibles of carbon, TiB_2 (which holds molten aluminum) or a refractory metal such as molybdenum. The evaporant itself sometimes melts—as, for example, aluminum or TiO—or sublimes, as do zinc sulfide and chromium.

Electron beam deposition has these advantages:

1. The heat is concentrated in a thin surface layer of evaporant and hence it is much more efficient than resistance heating. This is illustrated by the fact that about 1200 W of E-beam power is required to deposit a dielectric film such as lanthanum fluoride or titania in a chamber in which the substrate-source separation is 0.5 m.
2. Refractory materials such as thoria or rhodium are deposited easily.
3. It offers an effective method of evaporating mixtures of dielectrics to obtain intermediate refractive index films (see §3.6.3). Fugiwara [63a], for example, used electron beam evaporation to deposit a mixture of ceric fluoride and ceric oxide. The mixture was sintered and outgased before it was evaporated. Such mixtures are prone to fractionate and change composition if the vapor pressures of the components differ by more than a decade.
4. Contamination from the filament and/or boat, so often prevalent in resistance-heated sources, is avoided.

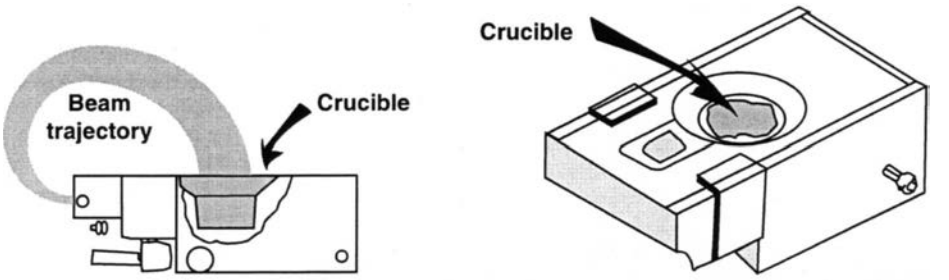


Fig. 9-14 Views of an electron gun that heats and evaporates materials in the crucible. The shaded portion of the elevation view represents possible electron trajectories.

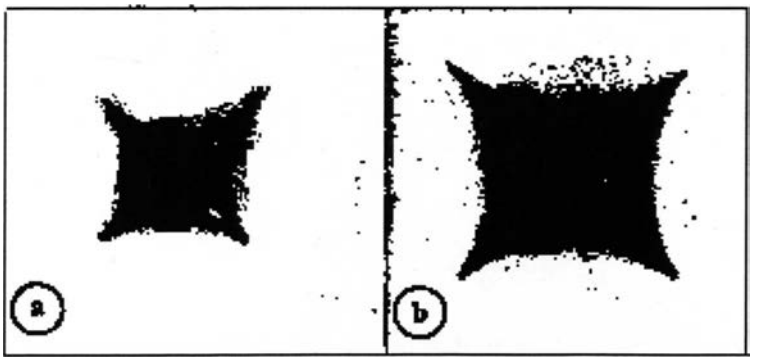


Fig. 9-15 Patterns in films of tungsten (*a— left panel*) and zirconia (*b — right panel*) produced by the electrostatic charging induced by E-beam evaporation of these materials onto rectangular glass substrates. After Zinsmeister [64].

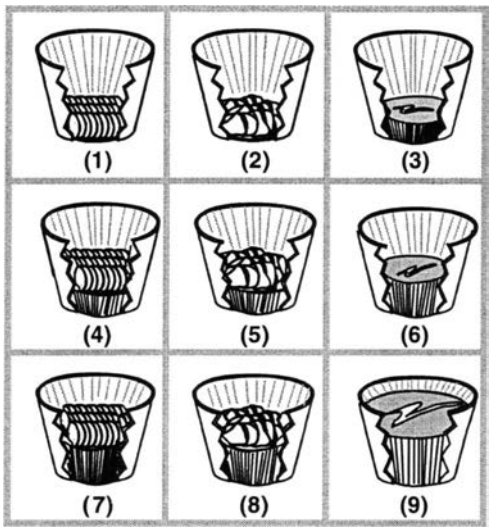


Fig. 9-16 In electron beam evaporation, steps in the conversion of tablets of the evaporant into a solid "plug" of material. The steps are enumerated in the text.

Disadvantages of E-beam deposition are:

1. Stray electrons bombard the substrates and chamber walls. The electrons are due both to the imperfect electron optics of the gun and also electron backscatter from the relatively high pressure cloud of evaporant near the source. Little is known about the manner in which these stray electrons affect the composition, structure and adhesion of the film. Hoffman [76] noted that electron bombardment induced finer grain size in rhodium films. Zinsmeister [64] shows photographs (see Fig. 9-15) of a pattern in films of tungsten and zirconia created by electrostatic charge buildup on an insulating substrate.
2. An E-beam deposition system is more costly than a resistance-heated source. The E-beam deposition system requires a vacuum tank with larger pumping capacity.
3. Films of refractory oxides such as titania, hafnia and tantala are deposited with repeatable optical properties only after the operator of the E-beam equipment is thoroughly trained.

9.4.3.1 Preparation of the evaporant prior to electron beam evaporation

In preparation for the deposition, the crucible containing granules (usually in the form of tablets) of a metallic oxide is preheated with the E-gun so that the material melts and forms a *plug*. This process is termed *burn down of the plug* in the trade and requires the skill of a trained operator to be done reproducibly.

This premelting also fills in any voids between the granules/tablets. If there are voids remaining in the plug, the gas in that void could be subsequently heated by the electron beam and it literally explodes. The molten evaporant ejected by this explosion is termed *spatter* in the coating industry.

The nine panels in Fig. 9-16 show the progressive steps in a *plug burn down*. In panel (1), the bottom of the crucible is packed with tablets of evaporant. If the evaporant is refractory, such as yttria, fragments of the tablets should be added to cover the tablets. The vacuum tank is evacuated and the E-beam melts the tablets until a solid material is formed, as shown in panels (2) and (3). The power needed to do this is usually about 20% higher than is used to evaporate the material. During the *burn down*, the operator uses the magnetic deflection to move the electron beam so that all areas of the *plug* are thoroughly melted.

The vacuum tank is opened and another layer of tablets is added, as depicted in panel (4). Then, under vacuum, these tablets are melted, as shown in panels (5) and (6).

The tank is again opened and the final layer of tablets is added, as in panel (7) of Fig. 9-16. The tank is evacuated and the final *burn down* occurs to form the final *plug*, as illustrated in panels (8) and (9). That plug is subsequently evaporated to produce multilayer coatings.

A final comment is that materials that sublime do not form a *plug*. Examples are silica, zinc sulfide and zinc selenide.

9.4.4 Ion beam sputtering as method of depositing coatings

9.4.4.1 Introduction

Sputter deposition of films in a plasma is treated superficially, yet this section discusses ion beam sputter deposition. Why? The latter merits our attention because of its proven success in depositing extremely high quality laser mirrors—mirrors with low optical scatter. Another reason is the similarity between the ion beam and thermal sources. In both cases, the evaporant molecules make few collisions—if any—with the residual gas in the chamber before they condense on the substrate. Plasma sputtered molecules, on the other hand, collide many times with the ambient gases. Both sputtering methods have proven successful. Pawlewicz [79] has reactively sputtered (in a plasma) coatings that have withstood the high irradiance of laser beams.

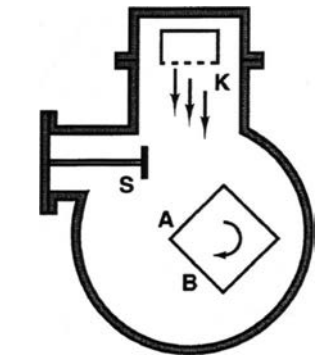


Fig. 9-17 A Kaufman ion gun *K* in a vacuum chamber directs an ion beam at a target, which it sputters. The sputtered materials on surfaces *A* and *B* of the target deposit on the substrate *S*. After Wei [79, 95].

9.4.4.2 Apparatus for ion beam sputtering

Fig. 9-17 depicts a rudimentary *ion beam sputtering* (acronym: IBS) apparatus [Wei 79]. A collimated beam of noble gas ions emerges from an ion gun—typical voltages range from 500 V to 1500 V. Electrons are simultaneously injected into the beam to maintain its electrical neutrality. The ions impinge on a target such as titania or silica. The target is sputtered and at the relatively low pressures of the vacuum of 10^{-5} Torr (1.3 mPa) or less, the evaporant vapor makes few collisions with the ambient gas molecules before it condenses upon the substrate. The titania deposits on the substrate just as though it had evolved from a thermal source, with the exception that the ion beam source is much larger in area.

A silica target is positioned in the beam when that material is deposited. The impingement of the ions on the target at 45° (rather than 0°) in Fig. 9-17 enhances the sputter rate. At the time of this writing (in the year 2000), the chief disadvantage of ion beam sputtering is its relatively low deposition rate, which is typically two to three times slower than that of vacuum evaporation. One advantage is the fact that the energy of the sputtered atoms is typically ten times that of the atoms emerging from a thermal source. The higher energy produces films with high packing density and better adhesion to the substrate. Films deposited by IBS are used in ring laser gyros because their the low optical scatter and plasma durability in the laser.

9.4.5 Introduction of gas during the deposition of a coating

Gaseous species refers to either neutral or ionized gases intentionally introduced during the deposition to modify film properties or the spatial distribution of the evaporant. Because the total pressure is the sum of the partial pressures of the introduced gases and the ambient gas, it is important that the latter—termed in coating technology the *base vacuum pressure*—be small enough so that the residual gases do not adversely affect the deposition. In most instances, 1 μTorr (0.13 mPa) is an adequate base vacuum.

There are two broad categories of the effects that occur when a deposition is made with a poor base vacuum. First, as discussed in §9.4.5.1, the evaporant molecules make many collisions with the molecules of the ambient gas. In general, films of poor mechanical structure are the result. Second, the ambient gas molecules contaminate the film. As an example, Hacman [73] observed that magnesium films deposited at 10 nTorr (1.3 μPa) manifested a substantially larger UV reflectance than those laid down at higher pressures.

9.4.5.1 The introduction of non-ionized gases

The usual procedure in vacuum deposition is to evaporate at a pressure sufficiently low that the molecules of the evaporants make few collisions, if any, with the ambient gas before they condense on the substrate as a solid layer. The thickness uniformity of the coating is usually achieved by masking. There is another way to improve thickness uniformity. Argon is bled into the chamber at pressures in the range of 0.1 mTorr to 1 mTorr (13 mPa to 0.13 Pa). At such relatively high pressures, the mean free path (see §9.4.1.1) is as short as 5 mm. The evaporant molecules collide many times with the ambient gas and a diffusion of the evaporant results. This technique, which was pioneered by Bausch and Lomb in the 1960s, provides a method of coating, with reasonable thickness uniformity, concave surfaces such as the cold mirror in Fig. 1-60. The pressure is critical. If too low, insufficient collisions occur and the uniformity suffers. The evaporant molecules collide many times at higher pressures. As a result, they lose kinetic energy and are enervated when they impinge upon the substrate. A mechanically soft film forms. At still higher pressures in the range of 0.01 Torr to 0.1 Torr (1.3 Pa to 13 Pa), evaporated metals condense in a spongy low-density structure. The resultant coating is termed a *black* [Harris 67] because it absorbs as much as 90% of the radiant flux at wavelengths from the far IR to the visible.

The oxygen bleed (also called *reactive evaporation*) process was pioneered by [de Rudnay 51] and [Auwärter 60]. A metallic oxide is more fully oxidized by this process. As an example, substrates are heated to well over 200 °C and the evaporation of either TiO or Ti_2O_3 [Pulker 76] with the residual oxygen pressure in excess of 0.1 mTorr (13 mPa) produces a titania film. The evaporation of SiO under the same conditions produces a film of Si_2O_3 that has an index of 1.55 in the visible portion of the spectrum. Its index is 1.85 or higher without the oxygen bleed.

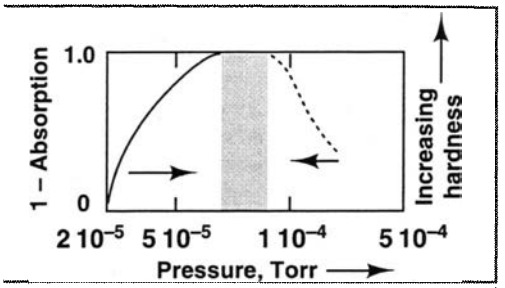


Fig. 9-18 The absorption (left ordinate, *dashed* curve) and relative hardness (right ordinate, *solid* curve) of an evaporated titania layer vs oxygen pressure during deposition. After Auwärter [60].

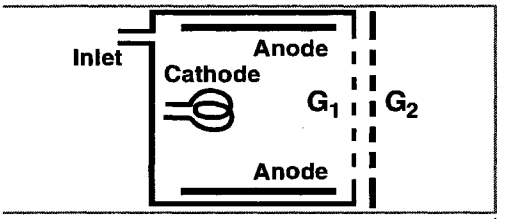


Fig. 9-19 Cross section of a dual grid ion gun, showing the cathode, anode, gas inlet and grids *G*. After Kaufman [82].

Figure 9-18 illustrates that the optical absorption and degree of mechanical hardness of titania layers change with oxygen pressure. If the pressure is large, the layers manifest low optical absorption, but are relatively soft. A lower pressure of oxygen produces mechanically hard, but absorbing, layers. In the shaded region in Fig. 9-18, both low absorption and hardness are achieved.

9.4.5.2 Ion-aided-deposition and hardware used to produce ions

Another technique is to bombard the substrate with ions while the film is being deposited. This is termed the *ion-aided deposition* (IAD) process. This is easy to accomplish when films are deposited by rf sputtering in a plasma (in contrast to ion beam sputtering mentioned in §9.4.4). The ion bombardment is easily controlled by adjusting the dc electrical potential of the substrate. As mentioned in the next paragraph, IAD in vacuum evaporation is more complicated.

Oxygen, argon and xenon are frequently used gases in IAD. There are several methods of producing the ions. Ionized oxygen can be produced by directing the stream of oxygen directly into the beam of the electron gun [Louderback 72]. Ion guns are more efficient sources. Van Cakenberghe [75] constructed a robust oxygen ion source in which an rf coil inductively coupled its electrical energy into an oxygen plasma. This source is bulky and complex to operate.

The dual grid ion gun in Fig. 9-19 is more widely used [Kaufman 82]. A heated filament (cathode) emits electrons in a cavity. The cathode is usually at a potential of approximately 50 V negative with respect to the outer cylindrical anode. A magnetic field surrounding the cavity increases the electron-ion collision probability. A plasma is established when gas is bled into the enclosure. The grid—a perforated plate close to cathode potential—extracts the ions. The grid is usually several hundred volts above ground and hence the ions are "work accelerated" as they emerge into the vacuum chamber. In the double-grid configuration (see Fig. 9-19), the outer grid imparts still more energy to the emerging ions.

9.4.5.3 Effects of bombarding a film with ion during its deposition

The modification of the physical and optical properties of films by IAD is reviewed by Martin [85]. Many process parameters can be varied, such as the ion current density, voltage and mix of ionic species. Potentials up to 1500 V optimize the rate when the beam sputters a target. Much lower voltages—less than a few hundred —are usually used with IAD to avoid sputtering the layer being bombarded. Ions can also induce

structural damage by implanting themselves into the film. Effects of IAD bombardment are mentioned below.

9.4.5.3.1 Enhancement of the adhesion of the film to the substrate

Hirsch [78] bombarded a substrate with argon ions while germanium was deposited from a thermal source. Although the ion density was small—the evaporant/ion ratio was 28:1—a notable increase in the adhesion of the film and apparent reduction in its mechanical stress occurred.

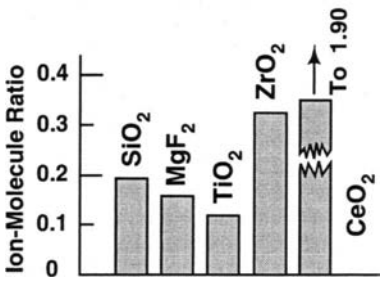


Fig. 9-20 Optimum ion-to-molecule ratio for IAD deposited optical films. "Optimum" means that there is no detectable change in the radiant transmittance in the visible spectrum when air is re-admitted to the chamber after the film was deposited. Data from Martin [86] and Netterfield [85].

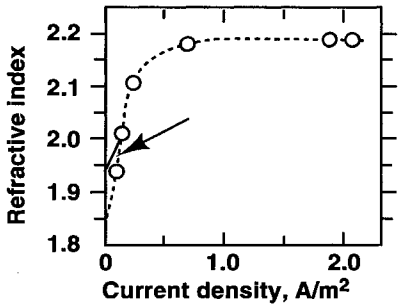


Fig. 9-21 As a function of the ion current density, refractive index of zirconia films (at 550 nm) measured in vacuum and after exposure to air. The films were deposited by vacuum evaporation with oxygen ions directed at the substrate during deposition. The arrowed dashed line is the change in index that occurs with the introduction of air into the chamber. After Martin [84].

Kennemore [84] also noted the same effects during the deposition of magnesium fluoride, although at the expense of a slight increase in its absorption in the UV part of the spectrum. Martin [85], using an ion density of 0.35 A/m^2, found that the IAD magnesium fluoride films deposited at a rate of 0.5 nm/s on substrates at ambient temperature had substantially the same characteristics as non-IAD films deposited at 300 °C. Figure 9-20 shows that the ion-evaporant ratio was 0.16.

9.4.5.3.2 Modification of the refractive index of a film

Figure 3-33 shows the dramatic effects of IAD on a zirconia film. The lower transmittance of the layer deposited with IAD indicates that the ion bombardment has increased considerably its refractive index. Figure 9-21 illustrates that the refractive index increases with ion current density—until a plateau is reached. This establishes the current density saturation level—the optimum. Martin [85] has found the optimum ion-to-molecule ratio for a few materials, as shown in Fig. 9-20. Equation (9-2c) was used in presenting these data.

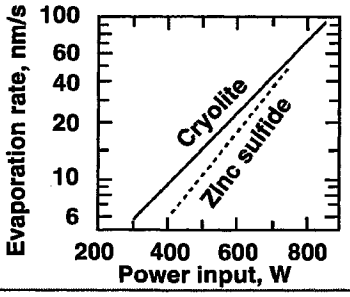


Fig. 9-22 On a substrate 250 mm above the source, deposition rate vs input power of an oven evaporating either cryolite or zinc sulfide. After Dobrowolski [82].

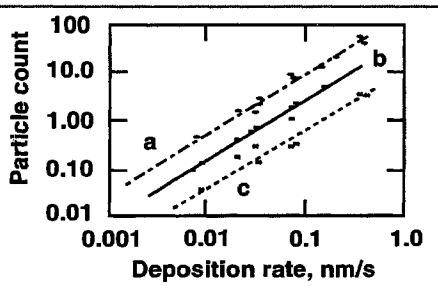


Fig. 9-23 Number of pinholes in an evaporated ceric oxide layer 100 nm thick deposited on a substrate of 150 square mm area. Pinhole size is: *a* 1 to 5 μm; *b* 5 to 10 μm; *c* > 10 μm. After Watanabe [80].

9.4.6 Rate control of evaporation sources

It is usually advantageous to control the rate of deposition, regardless of the evaporation method. The rate varies quite strongly with temperature near the threshold of evaporation. Suppose, for example, that a magnesium fluoride layer is being deposited from a thermal boat. Table 9-1 shows that a 107 °C increase of the magnesium fluoride temperature boosts the evaporation rate by at least a decade. Figure 9-22 shows the deposition rate vs input power for an efficient oven source [Dobrowolski 82]. A doubling of the power increases the rate tenfold.

There are several physical parameters that can be sensed to regulate the power and, hence, the rate: (1) Sensing the source temperature. Both contact devices, such as thermocouples, and radiometers have been used. (2) The thickness of the layer being deposited is measured by a quartz crystal (see §9.6.1.2) or an optical monitor. Differentiation provides the rate, provided the signal-to-noise ratio is acceptably low. (3) When the deposition is with an electron gun, the current of the stray electrons scattered by the evaporant cloud can be measured. The optical radiation emitted by the plasma of evaporant ions (just above the source) can also be measured. (4) The physical density of the vapor stream is measured by directing a low power electron beam probe into it. The fluorescence of the evaporant vapor is measured with a photodetector.

Once the rate is measured, feedback circuits are designed to regulate the power. The extreme nonlinearity of the rate vs power (shown in Fig. 9-22) makes the design of an optimal control circuit difficult.

9.4.7 Pinholes in coatings

Pinholes are objectionable in coatings such as opaque metallic reflectors, bandpass and neutral density filters and electroluminescent displays. There are at least three reasons that pinholes form: (1) Inadequate substrate cleaning; (2) deposition of dust on the substrates due to turbulence during the initial pump down, when the pressure exceeds 1 Torr (133 Pa)—see §9.3.2; and (3) spitting and crepetation from the evaporation sources.

Watanabe [80] demonstrated that a slow rate minimized the pinholes in a 100 nm thick ceric oxide barrier layer used in an image tube (see Fig. 9-23). This might not be practical in production, however, because the deposition took 2.7 h at the slowest rate.

Vergara [63] also noted a decrease in pinhole density with rate, but eliminated the particulates by reflecting the SiO vapor from a heated plate.

A method of disguising the pinholes in metallic mirrors is to deposit 50% of the coating thickness, remove the substrate from the chamber, clean it thoroughly and then finish the deposition. There are two sets of pinholes—one in each half of the layer. They are not seen because they are not aligned.

9.4.8 Procedures and hardware for deposition of layers

Practical details of coating are considered. The first question: To what pressure should the vacuum tank be exhausted before the deposition begins? If the pressure is too high, then either of two effects can occur: (1) the molecules of the evaporant scatter from ambient gas, as mentioned in §9.4.5.1; or (2) the residual gases in the chamber contaminate the film.

Scatter is present, but tolerable, if the source-substrate distance is comparable to a mean free path (see §9.4.1.1). A pressure (*base vacuum*) of 0.1 mTorr (13 mPa) is a typical upper limit. The contamination of the film by the residual gases in the chamber (see §9.4.5) depends upon the evaporant. As an illustration, if the intent is to deposit chromium and not a mixture of the metal and sundry oxides, the water vapor pressure in the chamber should be as low as possible, because chrome has an insatiable affinity for it.

The evaporants are usually preheated to outgas them after base vacuum (or lower) is attained. This dumps gas into the system and the pressure rises. Eventually the pressure drops again to below base vacuum as the pumps exhaust the chamber. The preheating is usually done with a mechanical shutter interposed between the source and substrate to avoid any premature coating of the substrates. The shutters *S* in Fig. 9-1 are open over both the E-beam and thermal sources.

The shutter is opened when the source spews out evaporant vapor at a sufficient rate. A typical value is 0.5 nm/s for dielectric films that function in the visible portion of the spectrum. How does the operator know when the rate is adequate? A few methods are: (1) The electrical power input to the evaporation source is preset; (2) the visual appearance of the evaporant surface is noted—an example is that magnesium fluoride melts just prior to evaporation; and (3) the rate is measured with a quartz crystal rate monitor (see §9.6.1.2) that "looks under" the shutter. The shutter is opened and closed with an air cylinder in an automated system.

A difficulty is that the bottom surface of the shutter is close to the source and is quickly coated with a thick layer. Quite often this layer manifests a substantial amount of mechanical stress, which causes it to peel off the shutter and fall like snow into the evaporant container below it. This contaminates it, if several different coating materials have been deposited upon the shutter before it peels. This can be minimized by positioning the shutter farther away, installing multiple shutters or using separate sources (and shutters) for each evaporant.

Rotating sources are mentioned, although they could appear in §9.5 because of their relation to the coating thickness uniformity. The usual approach is to rotate the substrates, rather than the sources, in order to achieve this uniformity. This produces vibrations that rattle the substrates in their tooling. This is minimized by stationary substrates and rotating sources. A system [Thelen 68] delivered up to 1500 A at 20 V to resistance-heated sources rotating at 0.6 Hz. At the time of this writing, the author is

aware of only two or three chambers that use moving sources. A rotating E-beam source is more difficult to design because of the high voltage. It is possible that coating production managers echo Hamlet's sentiment: *Rather bear those ills we have than fly to others that we know not of.*

9.5 Collection of the evaporant upon the substrates

9.5.1 Introduction — fundamental equations

This section focuses on the variation of the thicknesses of the layers at lateral positions across the substrates. The location of the substrates relative to the sources may be dictated by any of several requirements: (1) there are restrictions on the angles at which the evaporant molecules impinge upon the substrates; (2) the goal is to enable the substrates to collect the maximum amount of the evaporant; or (3) the thickness of the coating across the substrate is prescribed — a uniform distribution is common.

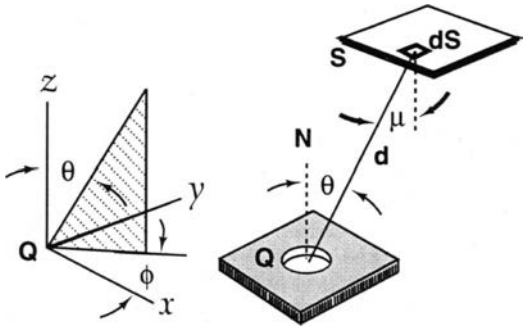


Fig. 9-24 An evaporation source Q deposits evaporant upon an infinitesimal area dS of the substrate. N is a normal to the source. The angles θ and μ are measured from a line of length d joining Q and dS to the normal to Q and the normal to dS, respectively. (*Left*) The molecular intensity of Q is measured in a spherical coordinate system with an azimuthal angle ϕ and polar angle θ.

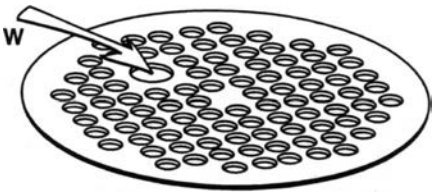


Fig. 9-25 Tooling that holds circular substrates as films are deposited on them by vacuum evaporation. The witness piece W is inserted in the larger hole.

The engineering trade-offs can be made once the requirements have been established. For example, it might be found that in order to achieve a thickness uniformity of a few percent across a 250 mm ϕ substrate, only 7% of the evaporant is deposited on the substrate — the remainder is collected on the coating chamber walls and tooling. In other situations, the uniformity might be of little importance where the goal is to optimize the collection efficiency.

The rate of film formation at a specified area dS at a given position on the substrate depends upon: (1) the molecular intensity of the evaporant; (2) the solid angle subtended by dS at the source; and (3) the sticking coefficient η of the evaporant.

The geometry appears in Fig. 9-24. In a system in which the deposition rate is constant, Eq. (9-2b) may be used to compute the relative thickness fi' deposited on the infinitesimal area dS of the substrate positioned at a distance d from the source whose molecular intensity (see §9.4.1.2) is $I(\theta)$

$$fi' = \frac{I(\theta)\,\eta\,\cos\mu\ const}{d^2} \qquad (9\text{-}3a)$$

where the *const*, which depends upon the deposition rate, converts from molecular density into film thickness. The angle μ is measured from a normal of dS to d.

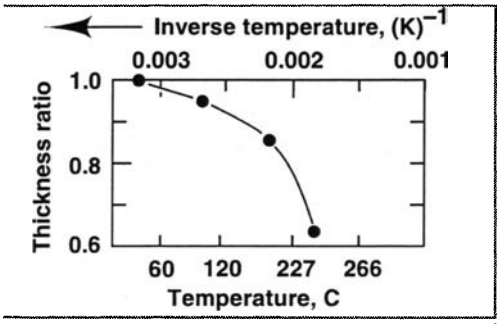


Fig. 9-26 Relative film thickness of a vacuum evaporated ZnS film vs temperature (*lower abscissa*) and inverse temperature (*upper abscissa*). The thickness is normalized to the thickness at 40 °C. After Ritter [69].

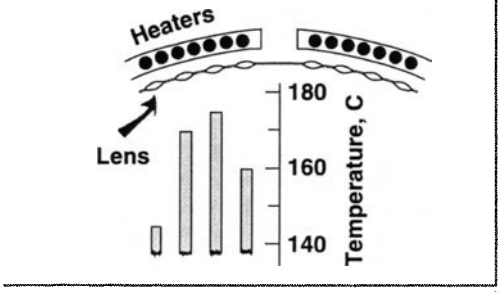


Fig. 9-27 In a coating tank four rows of lenses in a rack are rotated about the tank center and are heated from the rear. The temperatures of the rows are shown.

9.5.1.1 Influence of the sticking coefficient

Equation (9-3a) is straightforward to manipulate mathematically, once $I(\theta)$ is known. But what about the sticking coefficient η? It depends upon the temperature, impingement angle μ, as well as the texture and chemical composition of the surface the molecules are impinging upon—the undercoat. Suppose, for example, that a layer is deposited simultaneously upon an uncoated glass slab and a sheet precoated with a film of different composition. The undercoat offers many nucleation sites (see §3.3.2.2) and therefore the sticking coefficient is usually larger than that of the uncoated glass.

The surface temperature of the substrate is also important. Figure 9-25 depicts, for example, tooling that holds both small substrates and a larger witness piece. There are thermal gradients that could easily produce temperature differences. Hence, thickness variations can occur due to changes in the sticking coefficient at various lateral positions across a substrate. Figure 9-26 illustrates that the sticking coefficient of zinc sulfide changes with temperature. Figure 9-27 shows four rows of substrates in a circular rack as they were coated under single rotation and heated from the rear. The temperature measured in adjacent rows varies by nearly 20 °C.

The use of Eq. (9-3a) to determine film thickness seems formidable because it appears that detailed knowledge of both the sticking coefficient η and molecular intensity $I(\theta)$ are requisite. In practice, coating engineers often assume that η is constant and, from an educated guess of $I(\theta)$, obtain an approximate thickness at several positions across the substrates. A film is deposited and its thickness measured at those positions. The masking and geometry are altered until the desired thickness distribution is obtained. Masking of the substrates during deposition can alter the thickness distribution in an oblique manner if the masking occults the radiant heaters. This would alter the substrate temperatures and, hence, the sticking coefficients. An IR bandpass filter was deposited with a uniformity of 0.25% over a 300 mm ϕ substrate [Shimshock 85].

9.5.1.2 Assumptions about the configuration of the chamber

There are several tacit assumptions in the use of Eq. (9-3a) to find the thickness distribution. Neglected factors are: (1) Reflection of the evaporant from the chamber walls; (2) ionized species in the evaporant; (3) pressure differentials in the ambient gas; and (4) viscous flow near the source.

Evaporant vapor reflects from the chamber walls when the temperature is high enough that the sticking coefficient is less than one. The classic example is the deposition of zinc sulfide. It diffuses throughout the chamber and is found on the reverse sides of substrates, where numerous bounces are needed for it to arrive there. Vapor reflection is avoided by keeping the chamber walls cool and using a large chamber in which the substrates and sources are distant from the walls.

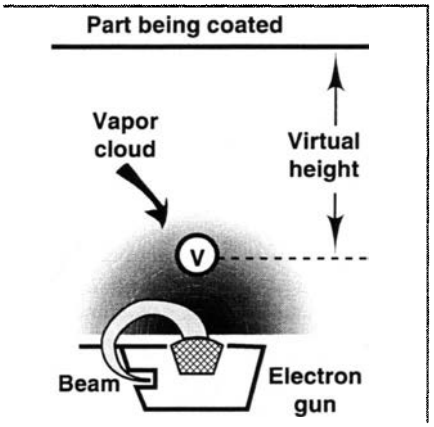


Fig. 9-28 For the evaporant vapor, showing the regions of viscous flow (*shaded*) and molecular flow (*nonshaded*) above an E-beam heated source. *V* is the location of the virtual source. Adapted from Airco [76].

The partial ionization of an E-beam-deposited evaporant is noted in §9.4.3. Any magnetic or electric fields may physically move these ionized species. Figure 9-15 demonstrates that the surface electrical resistivity of the substrate can also be important when ionized evaporant molecules impinge upon it. The pattern in that Fig. could be minimized by heating the glass during deposition, thereby increasing its electrical conductivity.

There are collisions of the evaporant vapor with residual gases in the chamber. This alters the distribution of the evaporant and makes it pressure dependent. There is a differential pressure in the ambient gas in the chamber because of the pumps. Suppose, for example, a Knudsen cell is centered on the baseplate of the coating tank in Fig. 9-1. The pressure is lower near the pump *E* and, even though the cell emits evaporant with the same molecular intensity at all azimuthal angles, the planet at *P* is coated differently than one nearer the pump *E* because of the pressure differential.

A complex subject is the relatively high pressure of the vapor near the source. The "local" pressure at the surface of the evaporant (see §9.4.1.3) is as large as 0.1 Torr (13 Pa). The mean free path (see §9.4.1.1) is 0.5 mm at that pressure. This clearly violates the assumption inherent in deriving Eq. (9-1c) that molecular flow occurs. One approach [Graper 73] is to assume that there is a virtual source *V* in the midst of the quasi-hemispherical evaporant cloud, as depicted in Fig. 9-28. In one study [Airco 76], the system was modeled by assuming that *V* was 20 mm above the E-beam hearth.

9.5.1.3 Effects of area of the evaporation source

It is emphasized that Eq. (9-3a) applies to a *small area* source. If it has any reasonable spread, then the total coating thickness *h* at a point on the substrate is obtained by integrating Eq. (9-3a) over the surface of the sources. Examples of this integration are:

an array of parallel wire sources [Schneider 60]; ring cavity source [Feisel 70, Glang 70]; and a sputter target [Gnaedinger 69].

9.5.2 Overview of methods of controlling layer thickness

There are many methods of obtaining a prescribed—if not a uniform—thickness distribution of the coating:

1. Gas scatter (see §9.4.5.1).
2. Positioning a small substrate at a large distance from the evaporant source.
3. Rotation of the substrate about a fixed axis (single rotation, see §9.5.5).
4. Single rotation with either stationary or moving masks (see §9.5.7).
5. Planetary rotation, as shown in Figs. 9-1 or 9-41 (see §9.5.6).
6. Planetary rotation with either fixed or moving masks (see §9.5.8).
7. Multiple evaporation sources whose rates are adjusted to produce uniformity.
8. Rotation of the evaporation sources (see §9.4.7).

Regardless of which method is used, the goal is to achieve a thickness distribution that is relatively insensitive to changes in the system parameters. Suppose, for example, the rate of evaporation decreases and the position of the virtual source (see Fig. 9-28) moves downwards, closer to the E-gun. The change in thickness induced by this rate change should be negligible in a well-engineered system.

There is an analogy between thickness uniformity in an evaporator and illumination uniformity in an optical projection system. If the molecular intensity in Eq. (9-3a) is replaced with luminous intensity and h' with $d\Phi$ (an infinitesimal unit of luminous flux), this equation transforms into the fundamental law of illumination. The $\cos^4\theta$ fall-off of illuminance on the flat screen of a projector is widely known and is analogous to the $\cos^4\theta$ decrease of the thickness on a flat plate (see §9.5.3).

This section on the thickness distribution of the coatings is brief. Holland [56], Behrndt [66], Glang [70], Pulker [84] and Frey [87] probe this topic in greater depth.

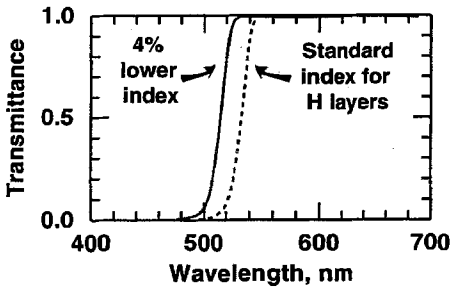


Fig. 9-29 $T(\lambda)$ of the edge filter whose design appears in §9.8.2 (*dashed* curve). For the *solid* curve, the refractive indices of the high index layers are 4% smaller.

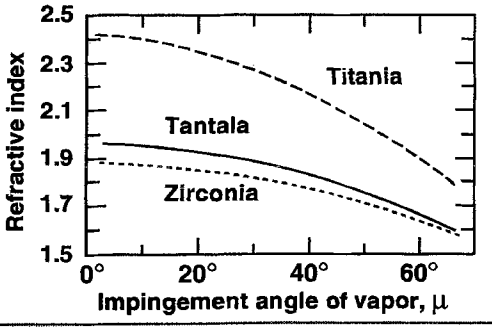


Fig. 9-30 Refractive index vs the impingement angle μ (of the vapor) for layers of titania, tantala and zirconia. Hodgkinson [97] also shows the birefringence in Fig. 16.4.

9.5.2.1 Control of the optical thickness

Let us assume that an edge filter is produced and that the wavelength of its 50% transmission point is to be controlled within ±1%. The coating is designed and is

deposited in the coating chamber in which there is single rotation of the substrate holder. The dashed curve in Fig. 9-29 is the transmittance of a substrate close to the center of rotation of that holder. The solid curve is measured at its periphery. Because the transmittance at that point is shifted 3.4% to shorter wavelengths, is it correct to assume that the layer thicknesses are 3.4% thinner? Not at all. In fact, the *metric* thicknesses of the layers in both coatings are precisely identical. For the solid curve, however, the refractive index of the high index layers are 4% smaller, which in turn lowers the optical thickness.

What are the likely causes of the lower refractive index? Possibilities are:

- There is a difference of temperature on the substrate holder. If oxide materials are deposited, it is generally true that a lower substrate temperature induces a lower refractive index.
- It is possible that the gas pressure is different at various physical positions of the substrate holder, which in turn alters the chemical properties of the film.
- Figure 9-30 illustrates that an increase in the impingement angle of the evaporant vapor hitting the substrate lowers the refractive index of oxide coating materials.

9.5.3 Molecular intensity distribution

9.5.3.1 General properties

Fig. 9-24 shows an evaporation source Q and substrate S. The insert shows a spherical coordinate system with Q at the origin. The z axis is perpendicular to the surface of Q. There is a polar coordinate system in which θ is the polar angle and ϕ is the azimuthal angle.

In the most general case, the molecular intensity depends upon both θ and ϕ, as shown in Fig. 9-24. As an example of the ϕ dependence of $I(\theta,\phi)$, Fig. 9-31 shows the azimuthal variation of the thickness from an E-beam evaporator that spewed out aluminum copiously onto a flat plate. A 2:1 azimuthal asymmetry occurred. Although it could be argued that the rapid deposition of aluminum is not apropos to optical coating, it reminds us that the assumption of an axially (about the z axis) symmetrical $I(\theta,\phi)$ is, in many instances, quite dubious.

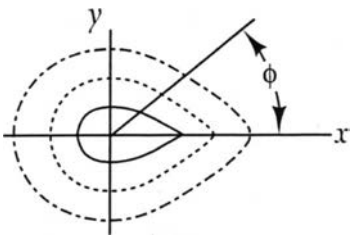


Fig. 9-31 Contours of rate of evaporation of E-beam-deposited aluminum as measured across a plate parallel to, and at a distance of 250 mm above, the molten pool. (*Solid*) 85 nm/s; (*short dash*) 67 nm/s; and (*long-short dash*) 50 nm/s. The E-beam direction is along the x axis. After Airco [76].

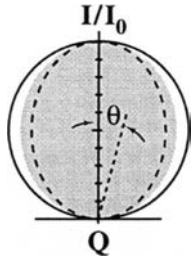


Fig. 9-32 Polar plot of the relative molecular intensity of a Knudsen source Q (*solid curve*), cryolite from an oven (*shaded*) [Dobrowolski 82] and SiO from a "Drumheller" source (*dashed*). After Anastasio [69].

Axial symmetry is, however, a valid approximation for many thermal sources and is widely used. This is illustrated by Fig. 9-32, which shows the $I(\theta)$ from oven type sources, similar to that shown in Fig. 9-12. Neither approximates the cosine distribution from the Knudsen source. Fig. 9-33 depicts the $I(\theta)$ of silver evaporated from a strip-type boat and also the crenelated $I(\theta)$ of a conical tungsten basket.

9.5.3.1.1 Normalization of the molecular intensity

There are at least two methods of normalizing the molecular intensity curves. In Figs. 9-32 and 9-33, they are normalized so that the maximum of the curve is 1.0. In the alternative method, the molecular intensity is normalized so that $\mathcal{W}$ is the total mass per unit time emerging from the source. This latter procedure is followed below.

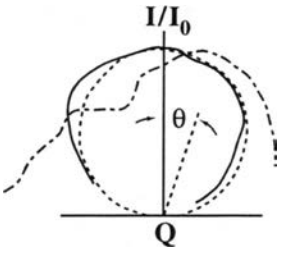


Fig. 9-33 Polar plot of the relative molecular intensity of a Knudsen source (*dashed*) and silver evaporated from a strip boat (*solid*) and conical basket (*long-short dash*).
After Anastasio [63].

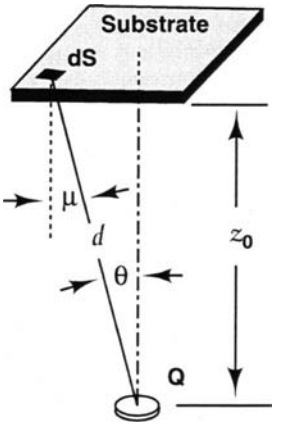


Fig. 9-34 Geometry of a planar substrate S at a distance z_0 above the source Q. The angles μ and θ are measured to normals to the plate and source to a line of length d joining the two.

9.5.3.2 Modeling of the molecular intensity

There are numerous models for the molecular intensity $I(\theta)$ curve, depending upon the material being evaporated, the source that evaporates it and the evaporation rate. The first is

$$d\mathcal{W} = (m + 1)\,\mathcal{W}\,(2\pi)^{-1}\,(\cos\theta)^m\,d\omega \qquad (9\text{-}3b)$$

where $d\mathcal{W}$ is the mass collected by a receiver that subtends a solid angle $d\omega$ at the source. m is typically in the range 1.5 to 2.5 for the vapors of dielectric materials. Assuming that the evaporation source is axial symmetric, the total mass $\mathcal{W}$ evaporated into a hemisphere is

$$\mathcal{W} = \int_{\textit{solid angle}} d\mathcal{W} \qquad (9\text{-}3c)$$

where $d\mathcal{W}$ is obtained from Eq. (9-3b) and the *solid angle* is the solid angle in a hemisphere. Evaluation of the foregoing leads to

$$2\pi \int_0^{\pi/2} (m+1)\, \mathcal{W}\, (2\pi)^{-1} (\cos\theta)^m \sin\theta\, d\theta = \mathcal{W} \ . \qquad (9\text{-}3d)$$

Dobrowolski [82] found that the molecular intensity of a cryolite source (see Fig. 9-32) is approximated by

$$d\mathcal{W} = (2\pi)^{-1}\ (2.633)^{-1}\ \mathcal{W}\,(\cos\theta + 6.4\cos^2\theta)\ d\omega \qquad (9\text{-}3e)$$

which shows that, although it contains a Knudsen component (i.e., the cosine), the $\cos^2\theta$ term is dominant. Pulker [84] noted that the molecular intensity of an electron beam aluminum source could be fitted as

$$d\mathcal{W} = (2\pi)^{-1}\ (0.312)^{-1}\,\mathcal{W}\,[0.86\cos^4\theta + 0.14]\ d\omega \qquad (9\text{-}3f)$$

The constant 0.14 in the foregoing equation represents the contribution of what is termed in radiometry a *point source* which emits vapor equally in all directions. For this planar E-beam source, however, the integration to collect the total evaporated mass $\mathcal{W}$ is still in a hemisphere, rather than a sphere.

9.5.3.3 Deviations from the model

It is possible that a coating merits the effort to measure the $d\mathcal{W}$ in §9.5.3.2 versus θ, which is proportional to the molecular intensity. One way of doing this is to position substrates, such as glass plates, equidistant from the source, deposit films on them and measure the thicknesses. Because the angle θ at each glass plate is known, a relative $I(\theta)$ can be computed.

Computer codes can then utilize this measured $I(\theta)$ to design the rotation and masking system needed to achieve uniformity. The problem is that $I(\theta)$ changes during the coating deposition for many reasons, such as: (1) evaporation rate; and (2) depletion of the evaporant or other changes in the source. Changes in the substrate temperature during deposition alter the sticking coefficient η in Eq. (9-3a).

§9.5.1.2 alludes to the *virtual source*, whose height is rate dependent, which, in turn, makes $I(\theta)$ rate dependent. Depletion of the evaporant in a source has several effects. First, the physical height of the evaporant [d in Eq. (9-3a)] increases. This depletion also alters $I(\theta)$ if the upper portions of the container holding the evaporant shadow the substrates. The effect of substrate temperature on the sticking coefficient is mentioned in §9.5.1.1. Thus what appears to be a change in $I(\theta)$ could, in actuality, be a manifestation of the sticking coefficient.

9.5.4 Layer thickness distribution on a flat nonrotating plate

Suppose the source Q in Fig. 9-34 is a Knudsen emitter (see §9.4.1.2) and a flat plate is suspended at a distance z_0 above with its surface parallel to the surface of the source. Figure 9-34 shows that the solid angle $d\omega$ [in Eq. (9-3b)] subtended by an infinitesimal planar area dS on the plate is

$$d\omega \approx d^{-2} \cos\mu\ dS \ . \qquad (9\text{-}4a)$$

Equation (9-3b) (with m equals 1) provides the molecular intensity of the Knudsen source. From geometry it follows that

$$\cos\theta = \cos\mu = z_0 / d \ . \qquad (9\text{-}4b)$$

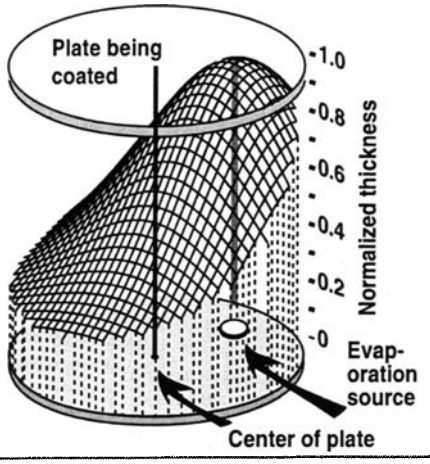


Fig. 9-35 A nonrotating plane plate is coated by a Knudsen evaporation source. The ordinate is the relative thickness of the coating.

Substitution of the above into Eq. (9-3b) results in the thickness per unit area h' of

$$h' \equiv \frac{d\mathcal{W}}{dS}\frac{1}{\rho} = \frac{\mathcal{W}\cos^4\theta}{\rho\,\pi\,z_0^2} \qquad (9\text{-}5a)$$

where ρ is the physical density of the film. Equation (9-5a) is proportional to

$h' \propto$ \F(the total mass of the material being evaporated = $\mathcal{W}$, square of the height of the plate above the source = z_0^2) .

(9-5b) If the plate is raised in height, then a "reciprocal-of-distance-squared" (i.e., z_0^{-2}) decrease in coating rate is observed, as long as θ remains constant.

The change in thickness at a given height is proportional to $\cos^4\theta$. From $\theta = 0°$ to $\theta = 10°$ the thickness drops by 6%. The decrease in coating thickness is illustrated in Fig. 9-35, which shows the relative thickness on a stationary plate. The precipitous decrease in thickness at larger θ values is evident. This can be lessened either by rotation of the substrate (see §9.5.5) and masking (see §9.5.7).

9.5.5 Single rotation of the tooling containing the substrates

9.5.5.1 Single rotation is equivalence to a ring evaporation source

Figure 9-36 depicts a rotating flat plate whose surface is parallel to the planar source Q. The displacement angle θ_0 is measured from the rotation center to the source. As θ_0 progresses from a few degrees to over 45°, the thickness distribution changes from "thick-at-the-center" to "thin-at-the-center," as shown in the relative thickness plots in the same Fig. This behavior is understood in terms of the model:

Single rotation is equivalent to a stationary plate and a ring source.

Imagine that a Lilliputian astronaut in a space suit is riding on the rotating substrate and looking down at the source. According to the perception of the astronaut, the source is rotating in a circle. Suppose that the deposition rate is constant and the plate makes an integral number of revolutions. The rotation can then cease, provided Q is replaced with a ring source. The radius of the ring is the offset of the source from the rotation center of the plate. The bottom of Fig. 9-36 shows a cross section of the hypothetical ring source along with the $I(\theta)$ of a Knudsen cell that emerges from each part of it [Dufour 50]. At an offset angle of 10°, it is not surprising that rotation does little to improve the uniformity and the distribution resembles a nonrotating flat plate. The center of the substrate is closer to the ring source than the edges, resulting in a "thick-at-the-center" distribution.

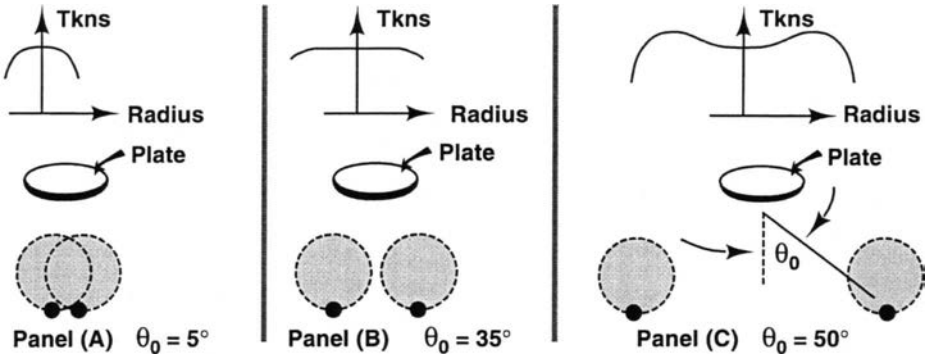


Fig. 9-36 (*Upper row*) Circular rotating plates are coated by Knudsen sources (*solid black circles*). (*Middle row*) Molecular intensity of a ring of Knudsen sources equidistant from the rotation center. θ_0 is the offset angle of the rotation center with respect to the source. (*Bottom row*) Relative film thickness vs radial distance on the substrate as θ_0 increases.

The opposite is true when the θ_0 approaches 45°, as in panel *C* of Fig. 9-36. The circumference of the substrate is closer to the ring than the center and "thin-at-the-center" distribution is produced. It is not surprising that the optimum flatness *at the center* is obtained when θ_0 is close to 35°, as shown in panel *B* [Ramprasad 72].

9.5.5.2 Thickness distribution of a layer on a rotating flat plate

Figure 9-37 depicts the polar coordinate system in which the disk rotates about the z axis at a height z_0 above the xy plane. The source Q is in the xy plane and is displaced from the origin by x_0. The variable s is the distance at an azimuthal angle ξ from the z axis to the position dS on the plate where the coating is collected. dS and Q are connected by a line of length d, which makes an angle μ to the normal of dS and an angle θ to the normal of Q. Table 9-2 summarizes the aforementioned variables, including the angles θ_0 and μ_0 which quantify the displacement of the source from the axis of rotation.

Given the physical density of the film ρ and the cosine of μ from Eq.(9-4b), the fundamental equation is obtained by combining Eqs. (9-3b) and (9-4a):

$$h' \equiv \frac{dW}{dS}\,\frac{1}{\rho} = \frac{(m+1)W\,(\cos\theta)^m\, z_0}{\rho\, 2\pi\, d^3} \quad . \tag{9-5c}$$

9.5.5.2.1 Normalization factor

The film thickness is normalized to the thickness that would be deposited on a small area at the center of rotation. From the geometry shown in Fig. 9-37, the slant distance d from the source to the area where the film is deposited is

$$d^2 = z_0^2 + x_0^2 + s^2 - 2\, s\, x_0 \cos\xi \tag{9-6a}$$

and the cosines of the angles are

$$\cos\theta = \cos\mu = z_0\, d^{-1} = (1 + (x_0/z_0)^2 + (s/z_0)^2 - 2\,(s/z_0)\,(x_0/z_0)\cos\xi\,)^{-1/2} \quad . \tag{9-6b}$$

The next step is to apply Eq. (9-6a) to the center of rotation (i.e., $s = 0$), which defines d_0 (in Fig. 9-38) as the slant distance from the source to the center of rotation

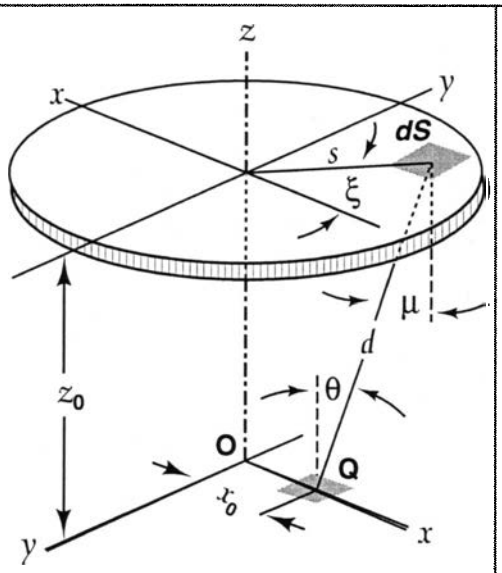


Fig. 9-37 A plate that is parallel to the xy plane is coated by the evaporation source at Q. The text explains the geometry.

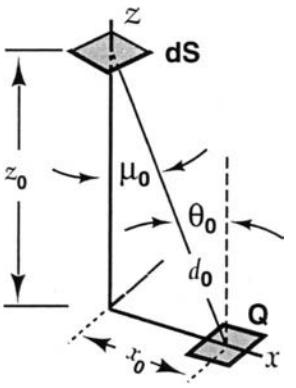


Fig. 9-38 Geometry related to the normalizing factor of the thickness at the plate center shown in Fig. 9-37.

$$d_0^2 \equiv z_0^2 + x_0^2 \tag{9-6c}$$

and Eq. (9-6a) is rewritten in dimensionless parameters

$$\frac{d_0^2}{d^2} = \frac{1 + (x_0/z_0)^2}{1 + (x_0/z_0)^2 + (s/z_0)^2 - 2\,(s/z_0)\,(x_0/z_0)\cos\xi} \,. \tag{9-6d}$$

The foregoing enables us to write equations for the film thickness, normalized to the thickness at the center of rotation.

9.5.5.2.2 Normalized film thickness

Assuming that the molecular intensity is given by Eq. (9-3b) as $I(\theta) = I_0\,(\cos\theta)^m$ and the sticking coefficient (see §9.5.1.1) is unity, it follows from Eq. (9-3a) that the normalized thickness h'/h_0 of a film on the substrate is

$$\frac{h'}{h_0} = \frac{\left[1 + (x_0/z_0)^2\right]^{(m+3)/2}}{\left[1 + (x_0/z_0)^2 + (s/z_0)^2 - 2\,(s/z_0)\,(x_0/z_0)\cos\xi\right]^{(m+3)/2}} \tag{9-7a}$$

where h_0 is the thickness at the center of rotation. The constant I_0 related to the molecular intensity [see Eq. (9-1d)] does not appear in the foregoing equation because of the normalization.

9.5.5.2.3 Rotation of the plate

The assumption is that the plate makes an integral number of rotations of 180°. The average thickness is then

$$\frac{h'}{h_0} = \kappa \pi^{-1} \int_0^{\pi} \frac{d\xi}{\left[1 + (x_0/z_0)^2 + (s/z_0)^2 - 2\,(s/z_0)\,(x_0/z_0)\cos\xi\right]^{(m+3)/2}} \tag{9-7b}$$

where the normalizing constant κ is

$$\kappa \equiv \left[1 + (x_0/z_0)^2\right]^{(m+3)/2} . \tag{9-7c}$$

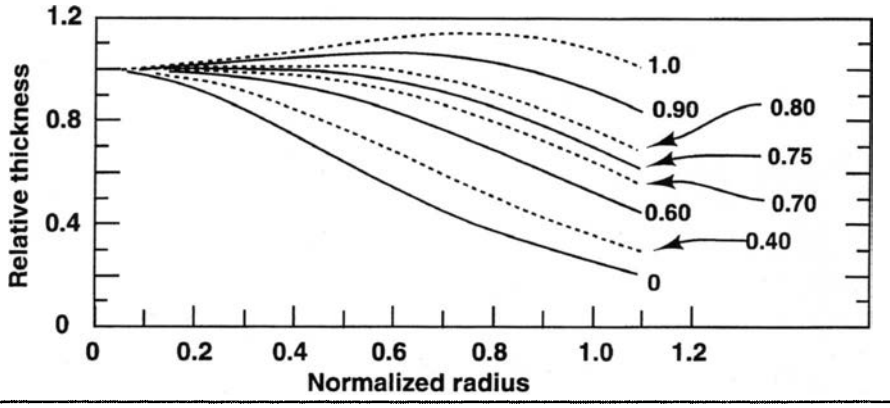


Fig. 9-39 For a Knudsen source, relative coating thickness (normalized to the thickness at the center of the plate) of a plate rotated continuously during deposition. The abscissa is f [see Eq. (9-9a)], which is the radial distance from the plate center of rotation. The numbers on the curves refer to g in Eq. (9-9b).

9.5.5.2.4 Knudsen evaporation source

This is the cosine distribution appearing in Eq. (9-1d). The integration in Eq. (9-7b) may be expressed in closed form when $m = 1$. Defining

$$a \equiv 1 + (x_0/z_0)^2 + (s/z_0)^2 \qquad \text{and} \tag{9-8a}$$

$$b \equiv 2\,(s/z_0)\,(x_0/z_0)\,, \tag{9-8b}$$

Eq. (9-7b) becomes

$$\frac{h'}{h_0} = \kappa \frac{a}{(a^2 - b^2)^{3/2}} . \tag{9-8c}$$

The presentation of Eq. (9-8b) in dimensionless parameters is helpful. Let f be the ratio of the radius s from the center of the rotating plate being coated to the height z_0 of the plate above the source

$$s \equiv f z_0 \tag{9-9a}$$

and g is the ratio (to z_0) of the offset of the source about the center of rotation

$$x_0 \equiv g z_0 . \tag{9-9b}$$

When g is zero, there is no offset of the source and the distribution is equivalent to the stationary flat plate in §9.5.4. Omitting the κ in Eq. (9-9b), it becomes the relative thickness when the factor in the "curly brackets" {} below is set equal to one.

$$h'_{\text{relative}} = \left[1 - \frac{(2fg)^2}{(1 + g^2 + f^2)^2}\right]^{-3/2} \left\{ (1 + g^2 + f^2)^{-2} \right\} \tag{9-10a}$$

Figure 9-39 shows the relative thickness vs f for several g values. When g is 0.4 the thickness peaks at the center and drops rapidly at the edges (see panel *A* in Fig. 9-36). At a greater offset (g is 0.90), the center is thin and the thickness peaks at the edges — this corresponds to panel C in Fig. 9-36. As observed in panel *B* of Fig. 9-36, at g approximately equal to 0.7 there is a flatness of the coating at the center of the plate. With reference to Fig. 9-38, this corresponds to an angle θ_0 of 35°.

Table 9-2 Variables associated with layer deposition

Variable	Definition
z_0	Vertical height of the plate above the plane containing the evaporant source.
x_0	Displacement of the evaporant source from the center of rotation.
d_0	A line (slant distance) to the source from the plate at its center of rotation.
μ_0	Angle between the z axis and the line d_0 .
θ_0	Angle between a normal to the source and the line d_0 .
d	Distance to the source from the "point being coated" on the plate.
ξ	Azimuth angle (about the z axis) of the "point being coated"
s	Radial distance from the z axis to the "point being coated."
μ	Angle between a normal at the "point being coated" on the plate and the line d.
θ	Angle between a normal to the source and the line d.
dS	Infinitesimal area of the "point being coated" on the plate

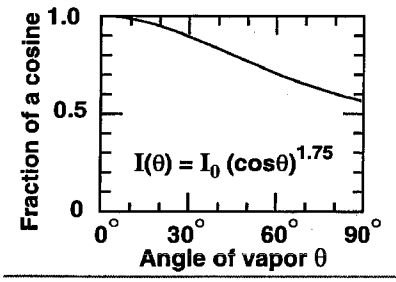


Fig. 9-40 Versus angle θ at which the vapor is emitted, molecular intensity of the evaporation source appearing in Fig. 9-12.

9.5.5.3 Molecular intensity of specific hardware

The fact is that few evaporation sources pattern the molecular intensity in Eq. (9-3b) when m is an integer. An example is shown in Fig. 9-40, which depicts the difference between the molecular intensity $I(\theta)$ of a Luxel source (see Fig. 9-12) and a Knudsen. The more rapid decrease of $I(\theta)$ at higher angles means that Eq. (9-3b) becomes

$$d\mathcal{W} = (1.75 + 1)\,\mathcal{W}\,(2\pi)^{-1}\,(\cos\theta)^{1.75}\,d\omega\,, \qquad (9\text{-}10b)$$

which is a more reasonable fit.

9.5.5.3.1 Use of numerical integration to find the thickness distribution

As depicted in Fig. 9-37, a position on the plate being coated is defined in terms of the radial distance s and the azimuthal angle ξ. The slant distance d in Eq. (9-6a) is found, as well as $\cos\theta$ in Eq. (9-4b). The thickness contribution in Eq. (9-5c) is then found. Texts on numerical analysis guide the programmer on the number of increments of ξ that should be used in the integration between 0° and 180°.

9.5.5.3.2 The angles at which the vapor impinges upon the substrate

The angle at which the coating vapor impinges on the substrate is often a matter of concern. This angle appears as μ in Figs. 9-37, 9-34 and 9-24. Consider the case in the previous paragraph in which $g = 0.7$. A quick calculation shows that μ is 55°. The arrival of the vapor at such large angles may produce films that are of lower refractive index and physically less durable. The influence of the vapor incidence angle upon the refractive index is mentioned in §9.5.2.1.

9.5.5.4 Geometries other than a flat plate

A closed-form solution for the thickness exists when the evaporation source is Knudsen and the plate is spun about an axis tilted at an angle γ to the z axis [Behrndt 72]. Behrndt [72] and Pulker [84] show that the thickness on a rotating segment of a sphere may also be expressed in an equation similar to Eq. (9-16d), provided $I(\theta)$ varies as $\cos\theta$. Frey [87] evaluates it for the more general molecular intensity appearing in Eq. (9-5b).

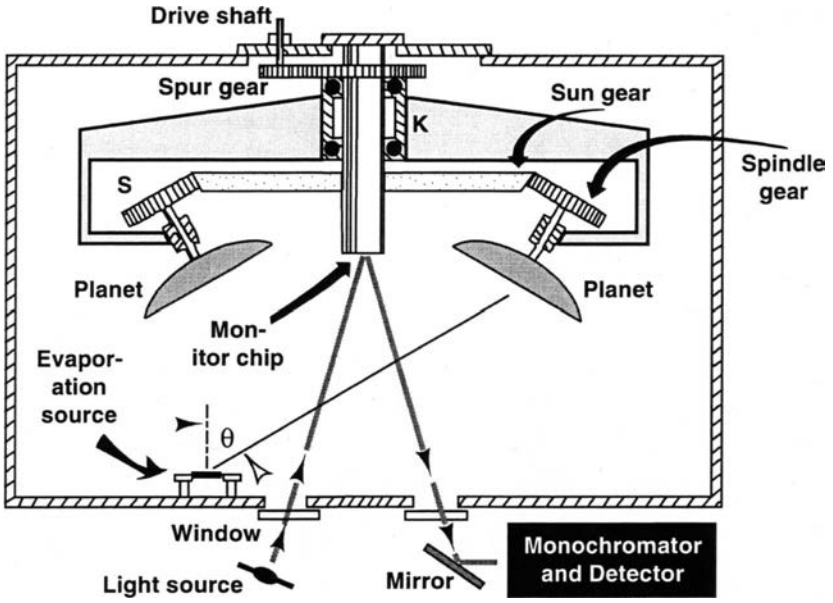


Fig. 9-41 Planetary rotation of the substrates in the planets is achieved by the relative motion of the spindle gears S and the stationary sun gear (dotted). The bearings K support the rotating spindle assembly. A light beam is directed through a vacuum tank window and reflects from the monitor chip. After Bergfelt [75].

9.5.6 Planetary rotation of the plano tooling

9.5.6.1 Introduction

Figures 9-1 and 9-41 depict substrates in rotating planets. The large sun gear and mechanical bearings are expensive—both to design, construct and maintain. For a deposition system in which there is planetary rotation of the substrates, several of its advantages are:

- The thickness distribution becomes less sensitive to changes in the molecular intensity of the source, as compared to single rotation. This is due to the fact that the substrate collects the evaporant at a greater variety of positions above the source, as compared with single rotation.
- As compared to single rotation, it is easier to install masking in a planetary (see §9.5.7).
- As compared with single rotation, there is much better averaging of the impingement angle of incidence of the evaporant vapor.

The disadvantages of planetary rotation are:

- As compared to single rotation, its mechanical construction is more complex and more prone to failure. For example, the mechanical bearings that hold the spindles may be heated relentlessly, which causes them to fail.
- As compared to single rotation, the area that may be coated is smaller.
- The parts in the planets are subjected to high acceleration and thus the parts must be secured mechanically. When prisms are coated, for example, the tooling that holds them must be able to withstand these high "g" factors.

9.5.6.2 Design parameters for planetary rotation

Figure 9-42 traces a point on the substrate as it revolves in a planet. The shape of this path changes if the design parameters are varied, such as:

(1) Radius of the planet, which is the distance from the spindle center to the edge of the planet holding the substrates;
(2) offset of the source from the sun gear center, as shown in Fig. 9-41;
(3) the gear ratio, which determines the number of spindle rotations for each complete revolution about the sun gear;
(4) tilt of the planet, as shown in Fig. 9-41; and
(5) an eccentric mounting of the sun gear so that spindle tilt changes during its rotation around the sun gear.

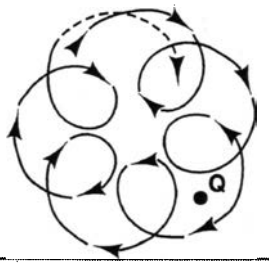


Fig. 9-42 The path of a point on a planet as the spindle makes a complete circuit around the sun gear. The gear ratio is 47:9. The source *Q* is also shown.

9.5.6.3 Methods of computing the thickness distribution

With the exception of Bloom [79], studies of thickness uniformity have used numerical integration of various forms [Behrndt 63, Hopkins 64, Zhiglinskiy 71, Behrndt 72, Ramprasad 73]. This procedure is usually followed: (1) The molecular intensity $I(\theta)$ (see §9.5.3.2) is either estimated or measured; (2) an array of positions on the substrate is chosen; (3) published equations [Behrndt 72, Pulker 84] are used to compute μ and d^{-2} in Eq. (9-3a), thereby finding the relative thickness at each substrate position; and (4) the rotation of the planet is advanced and step (3) is repeated, thereby accreting the thicknesses. Steps (3) and (4) are repeated until the planet spindle has made at least one complete rotation about the sun gear.

9.5.6.3.1 Detailed equations

The objective is to program on a digital computer the aforementioned numerical integration. An appendix (§9.8.1) outlines a method of finding the Cartesian coordinates x, y of a point P on the planet. Given $\cos\theta$ and d in Eq. (9-5c), it follows that

$$d = \sqrt{(x_0 - x)^2 + y^2 + z_0^2} \tag{9-11a}$$

$$\cos\theta = z_0 / d \tag{9-12a}$$

where z_0 and x_0 appear in Fig. 9-37. As the angle of rotation of the planet is advanced, the numerical integration mentioned in §9.5.6.3 is performed.

9.5.6.3.2 Design constraints

The planet diameter determines the area that can be coated. It is constrained for several reasons. First, if it is made too large, the vapor may impinge on the substrates at angles as large as 60°. As mentioned in §3.3.1.1, this may produce a mechanically soft film. It can also alter its refractive index. Second, it may produce an unacceptably large nonuniformity in the thickness of the layer if the vapor distribution of the source changes. Suppose, for example, that at the start of an evaporation, the $I(\theta)$ of a source is approximated by $(\cos\theta)^{2.5}$ and masking is devised which, along with planetary rotation, produces coating thickness uniformity of less than 1%. But what happens later in the deposition when the evaporant is depleted and the $I(\theta)$ changes to $(\cos\theta)^{2.8}$? The thickness uniformity suffers if a physical size of the planet is too large relative to its height above the source.

The same general comments apply to the selection of the offset of the source from the center of rotation. It is advisable to position the source not far from the center of rotation, to avoid large vapor impingement angles.

The gear ratio should be selected so that the number of teeth on the sun gear and the number on the planet spindle gears are noncommensurate. Suppose, for example, that a ratio of 45:9 is chosen. That is, there are 45 teeth on the sun gear and 9 on the spindle gear. The gear ratio is 5:1. The planet then rotates six times for each full rotation of the sun gear. Why six? The spindle gear rotates the planet 5 times and the sun gear rotates it 1 time: 5 + 1 = 6. This is undesirable because a point on the planet retraces the same path during each trip about the sun gear. If a ratio 47:9 is used, however, then a point on the planet does not retrace its path but shifts each revolution, as shown by the dashed line in Fig. 9-42. It is tempting to use friction, rather than gears, to drive the planets (an example is [Illsley 64]). It is perhaps because of the uncertain effects of slippage that its use has been limited.

The advantage of tilting the axis of rotation of the planets, as shown in Fig. 9-41, is that the angle at which the vapor impinges upon the substrates is much closer to normal than when the planet spindles are vertical. One problem this introduces is that the evaporant is collected from the source at a relatively large angle θ. This is illustrated by the θ in Fig. 9-41, which is as large as 60°. Although the thickness uniformity across any given planet may be relatively insensitive to variations in $I(\theta)$, the ratio of the thickness on the monitor plate (see Fig. 9-41) to the thickness on the substrate changes if $I(\theta)$ changes. The result is that the substrates are coated with films of incorrect thicknesses, yet with superb uniformity.

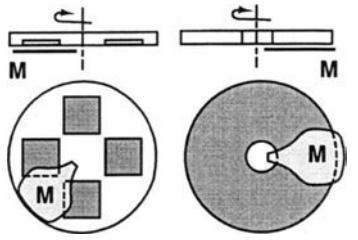


Fig. 9-43 Configurations in which the center of rotation does not intersect any part of the substrates being coated. (*Left*) Separate substrates are nested about rotation center; and (*right*) the mirror has a hole in its center. The mask *M* (*lightly shaded*) produces thickness uniformity.

9.5.6.4 Complete and partial rotation of the planets

The integral in Eq. (9-7b) is between 0 and π. This is a reminder that in computing the thickness of both single rotation (see §9.5.5) and planetary rotation (see §9.5.6) it has been assumed tacitly that the substrates make a complete half revolution about a line that contains the rotation center and the evaporation source. But coating machine operators will attest that this occurs seldom, if ever. What errors are introduced? Hopkins [64] computes the thickness variation across the substrate for two cases: (1) the planet makes one or more full revolutions about the sun gear; and (2) 40% of a full rotation is made. In the latter case, the thickness distribution across the substrate resembles the former but is displaced from the center. That is, the distribution is no longer rotationally symmetrical about the spindle of the planet, as it is in the case when the number of rotations is integral.

An example is a production coating machine [Coherent 92] in which each planet rotates completely about the sun gear at a frequency of 0.15 Hz, which implies that 45 rotations are completed during a 300 s deposition of a layer. It may be concluded that the "worst case" thickness nonuniformity could be 1/45 or 2.2%. In practice, however, it is not nearly as large as this.

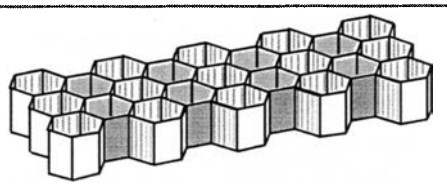


Fig. 9-44 The hexagonal "honeycomb" array is a grille-collimator mask for producing films of uniform thickness.

9.5.7 The use of masking to achieve thickness uniformity, single rotation

9.5.7.1 Undesirable effects of masking

The most obvious effect of masking the substrate during part of the deposition is that the amount of material deposited upon the substrates is reduced, thereby increasing the coating time. Because evaporant collection efficiency and thickness uniformity are at odds with each other, it is not possible to "have it both ways." Pulker [84] (p. 191), for example, describes a coating geometry in which an excess of 50% of the evaporant is collected. The layer thickness on the substrates varies enormously, of course.

If the substrates are heated (see §9.3.4) from the bottom, then the masking reduces the amount of heat reaching the substrates and alters the temperature distribution.

The final observation that the masks are usually in close proximity to the substrate. A consequence of this propinquity is that the mask is a quite effective emitter. Emitter of what? Absorbed water vapor, for one. Still worse, if that mask was previously cleaned by glass bead bombardment, minute fragments of that glass are embedded into the surface of the mask. The heating of the mask loosens these fragments and enables them to migrate to the substrate.

9.5.7.2 Overview of masking

9.5.7.2.1 Types of masks

There are at least two types of masks that may be used: (1) Sheet masks; and (2) the grille/collimator mask.

The sheet mask simply consists of a sheet of material, usually a metal such as stainless steel. This is widely used in the industry and is, for example, depicted in books such as Willey [96].

Figure 9-44 depicts one form of the grille/collimator mask—apertures in a hexagonal grille. This apparatus has been used in the semiconductor industry for decades and is mentioned in the review article of Rossnagel [98]. It offers the advantage that it controls not only the relative area of the substrate that is allowed to be coated, but also the impingement angle of the vapor (see §3.4.2). The control of the vapor impingement angle avoids problems such as films with poor mechanical properties which occur when that angle is large.

9.5.7.2.2 Arrangement of the substrates on the tooling

The masking is treated differently, depending upon the location of the rotation center of the spindle relative to the substrate(s) mounted in the planet. Either of two cases obtains: (1) the projection of the center of rotation (i.e., the axis of the spindle shaft) is *on* one of the substrates; or (2) the projection of the center of rotation intersects the tooling, but *not* any of the substrate(s).

Figure 9-43 shows two configurations of the hardware in which the latter occurs: (1) the parts are relatively small and are nested about the rotation center—there is no part at the center; or (2) the substrate is relatively large, but its center is uncoated, as the Cassegrainian telescope mirror in Fig. 9-43.

9.5.7.3 The center of rotation is on a substrate, fixed masking

For reasons described below, it is highly desirable that the source be positioned well off center, as shown in panel *C* of Fig. 9-36. The source in Fig. 9-45 is directly below the circumference of the circular substrate and hence θ_0 (see Fig. 9-36) is 45°. The upper curve in Fig. 9-46 shows that the film is deposited 14% thicker (at point *A*) near its periphery than at its center. The mask in Fig. 9-45 corrects, at least partially, for this thickness bulge. It is cut so that its angular width is $(y-1)\times 360°$, where y is the normalized thickness plotted in Fig. 9-46. At point *A*, for example, the film is 14% thicker than at its center; hence, it must obscure 14% of the total rotation (i.e., 360°) or $(1.14-1.00)\times 360° = 50°$.

The foregoing analysis is incorrect because it assumes a uniform distribution of the evaporant over the bottom of the mask. In fact, the evaporant density is decreasing steadily along the *OP* axis as $\cos^4\theta$ (see §9.5.4 and Fig. 9-35). This is shown by the dashed curve in Fig. 9-46. For example, at points *B* and *C*, the evaporant density has decreased to 50% or less and hence the mask is only *half* as effective at the center of rotation. Its angular spread must be approximately doubled at these points to compensate for this decrease. When this is done, the correct shape of the mask is shown by the dashed line in Fig. 9-45. Figure 9-47 depicts a mask that corrected such 10% thickness nonuniformity on a 1.8 m ϕ telescope mirror [Itek 76].

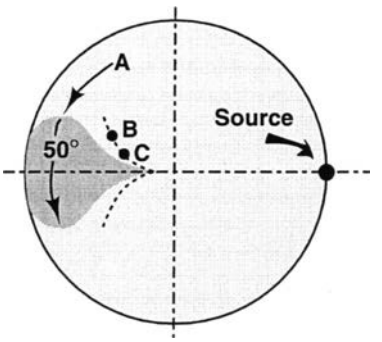


Fig. 9-45 Construction of a mask with source Q located under circular a substrate that rotates about its center. The mask (*darker shaded*) corrects for a 14% thickness bulge shown in Fig. 9-46. Points B and C on the dashed line represent the correct mask shape, as explained in the text.

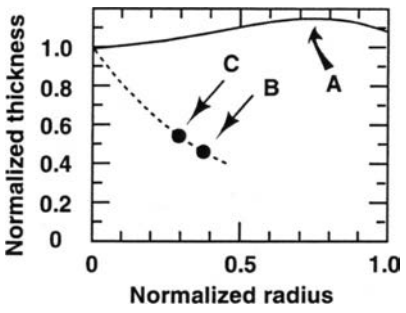


Fig. 9-46 For a source Q located under a circular disk rotating about its center, thickness (normalized to center) without masking (*solid*). (*Dashed*) decrease of thickness (without rotation) along a diameter that intersects source and heads towards the center of rotation in Fig. 9-45.

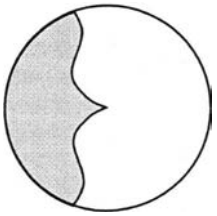


Fig. 9-47 Mask (shaded) that produced a coating of relatively uniform thickness on a 1.8 m ϕ telescope mirror [Itek 76]. The mirror was rotated about its center during deposition.

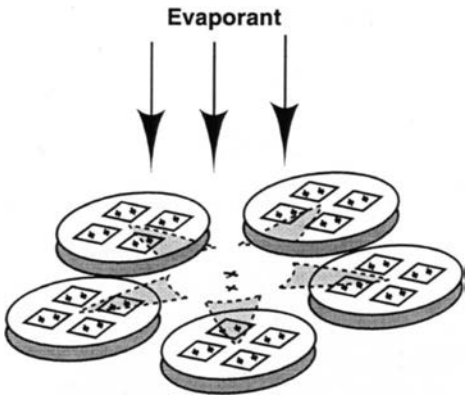


Fig. 9-48 Five planets are rotated synchronously with the petal masks (*shaded*) to achieve thickness uniformity. For purposes of illustration, the system is inverted—the evaporant flows from the top downwards. After Drumheller [69].

Stationary masking usually provides an adequate thickness uniformity. There are at least two instances in which it does not: (1) the molecular intensity of the source drifts during deposition; or (2) the evaporant impinges upon the substrates at too steep an angle. The maximum angle is 63° in Fig. 9-45, which may produce films with poor mechanical properties. The advantage of this system, however, is that the mask need not actually approach the center of rotation and hence no "center spot" is produced.

Suppose the source is moved close to the center and a *thick center* occurs, as in panel *A* of Fig. 9-36. Can a mask be designed to produce thickness uniformity? In

theory, perhaps, but not in practice. The problem is that it must be positioned quite precisely with respect to the center of rotation. A center spot of zero thickness occurs if any part of the mask obscures the center, even one iota. The effects of the lateral extent of the evaporation source (see §9.5.1.3) and outgasing of the mask (which diffuses the evaporant), make the elimination of the center spot even more difficult. The recourse is to use rotating masks (see §9.5.8).

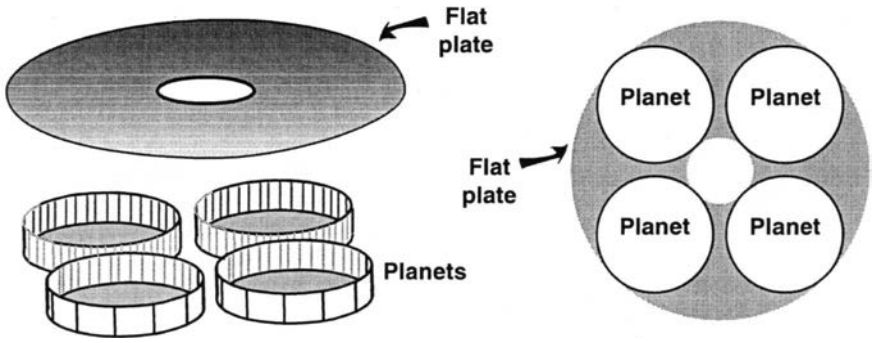


Fig. 9-49 (*Lower left panel*) A perspective view of the four substrate holders rotated in planetary motion. (*Upper left panel*) The coating is deposited upon a rotating flat plate. (*Right panel*) A plan view that superimposes the flat plate and planets.

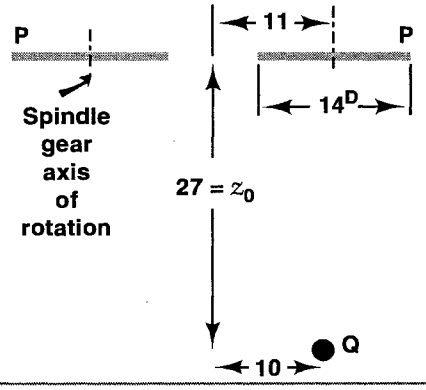


Fig. 9-50 For an evaporation source Q displaced from the axis of rotation by a distance of $x_0 = 10$ *units*, at a height of $z_0 = 27$ *units* there four planets P whose spindle gear centers are 11 *units* from the sun gear center.

9.5.7.4 Center of rotation off substrate, fixed masking

In this case, the center spot (see §9.5.7.3) is of no concern. Hence, there is wide latitude as to where to position the sources and the position of the masks. An example is that Netterfield [81] achieved uniformity with both fixed and rotating masks. Figure 9-48 shows petal masks that rotate in synchronism with the planets.

9.5.7.5 Rotating masks

Its disadvantage is an increased cost and complexity. The advantage is an improved averaging over the changes in the azimuthal distribution (see §9.5.3.1) of the molecular intensity of the source. Behrndt [63], for example, rotated a mask with an external sun gear and deposited silver and aluminum with a uniformity of 0.2% over a 90 mm ϕ substrate. The mask of Bergfelt [80] is highly recommended for producing thickness uniformity on substrates in a planetary system.

9.5.8 A comparison of single rotation and planetary rotation

The object is to compare the following coating systems:

- There are four planar substrate holders in planetary rotation.

• There is a planar substrate holder with single rotation.

In each configuration, the substrate holders are parallel to the base plate of the coating chamber. The geometry is shown in Figs. 9-49 and 9-50. There are four planets, each of which is 14 *units* diameter. The center of rotation of each planet is 11 *units* from the center of the sun gear and at a height of 27 *units* above the evaporation source. The molecular intensity of the source is given by Eq. (9-10b). Each *unit* may be regarded as 25 mm and in the nonmetric world would translate into an inch.

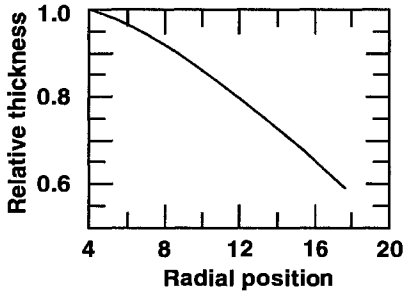


Fig. 9-51 For the flat plate coated with single rotation shown in Fig. 9-49, relative thickness of the coating vs radial distance from the center of rotation.

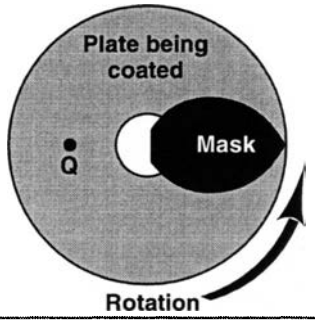


Fig. 9-52 The flat plate depicted in Fig. 9-49 is masked (*solid black area*) to improve the thickness uniformity. Q is the evaporation source.

9.5.8.1 Single rotation

Figure 9-51 depicts the relative thickness distribution on the substrate. Its thickness decreases toward the outer edges. At its periphery it decreases to 60% of the thickness near the center — at a distance of 4 units (a *unit* is 25 mm) from the rotation center.

It is quite evident that masking is needed. Figure 9-52 shows the mask that produces a uniform coating. Because the coating is "thick at the center and thin on the outside," the mask width is essentially zero at the periphery and becomes thicker towards smaller radii.

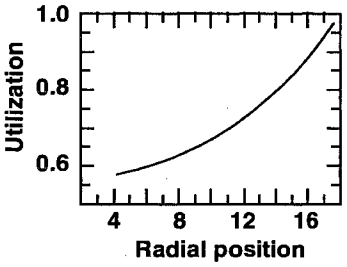


Fig. 9-53 For the flat plate depicted in Fig. 9-51, fractional utilization [see Eq. (9-13a)] of the coating material vs radial position from the center.

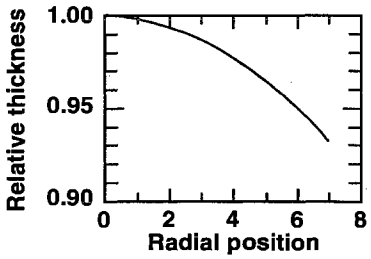


Fig. 9-54 For the planetary system depicted in Fig. 9-50, relative thickness of the coating vs radial position from the planet rotation center.

As mentioned above, the mask reduces the efficiency of the coating deposition because it intercepts a portion of the coating material. The utilization factor is defined as

$$\text{utilization factor} = \frac{\text{Thickness deposited on plate with masking}}{\text{Thickness deposited on plate without masking}} \qquad (9\text{-}13a)$$

Figure 9-53 graphs this utilization factor. It is, by definition 1.00 at the periphery, where the mask width is vanishingly small. It decreases to below 60% close to the center where the masking is the heaviest. The average utilization factor is 73%.

9.5.8.2 Planetary rotation

Figure 9-54 shows the relative thickness on each of the planets. Similar to the single rotation above, the coating is "thick at the center and thin at the periphery." In this case, however, the *center* is the center of rotation of the planet about the spindle gear axis. The decrease in thickness at the edge is only 7% below the thickness at the center.

The aforementioned center bulge in the thickness is corrected by moving each planet under a mask during part of its rotation cycle. The shape of the mask is similar to that of the football used in the United States. Bloom [79] published details on the design of this type of mask. It is apparent, *ab intuito*, that this shape shades the center of the planet more than the periphery, thereby reducing the center bulge. If no portion of the substrates covers the spindle center (see Fig. 9-43), then each planet may be masked separately, similar to the arrangement in Fig. 9-48.

As compared to the single rotation, the utilization factor is much higher for the planetary rotation — close to 96%.

9.5.8.3 Comparison of the results

Table 9-3 summarizes the results of the single rotation and planetary systems.

Table 9-3 Summary of single and planetary rotation

Parameter	Single	Planetary
Minimum and maximum radius coated	4 to 18	0 to 7
Area coated	977	615
Utilization factor	73%	96%
Product of (Area coated) (Utilization factor)	713	590
Relative effective area coated	100%	82%

Summary of the performance of the single rotation system and planetary designed in §9.5.8. The area coated is in *square units*, where each *unit* is 25 mm.

Improved averaging over changes in the azimuthal distribution of the evaporant cloud is achieved with petal masks that cover part of the planets, as shown in Fig. 9-48 [Drumheller 69, Drumheller 85, Bergfelt 80]. In this approach, the petal masks rotate at the same speed as the spindles rotate about the sun gear.

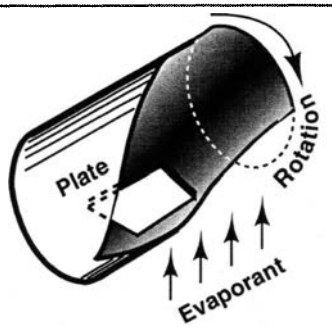


Fig. 9-55 During vacuum evaporation, a linearly wedged thickness film is deposited on a plate inside a rotating cylinder and near its periphery. After Fabre [61].

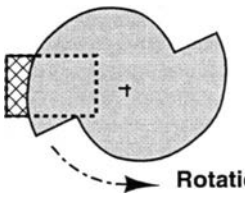


Fig. 9-56 A rotating mask (*shaded*) produces a linearly wedged thickness film on a substrate (*cross hatched*) during vacuum evaporation. After O'Brien [39].

Another approach is to have the petal masks rotate separately and in opposite directions. That is, if the spindles rotate about the sun gear clockwise, the masks should rotate counterclockwise. A stack containing 30 alternating layers of dielectric material was deposited using such a system. A thickness uniformity of ±1/2% across a 900 mm ϕ planet was achieved.

9.5.9 Use of masking to produce a nonuniform thickness distribution

9.5.9.1 Introduction

§9.5.7 and §9.5.8 focus on thickness uniformity. Films of nonuniform thickness have also been used in various applications. They are used to modify the optical figure of a lens or transmissive element or mirror [Kurdock 78]. Nonrotationally symmetrical thickness distributions are achieved by masking with a large photo-etched metal sheet. The coatings in §1.8.5.3.1 and §1.9.8 require a wedged thickness. The performance of each of these devices is improved if nonlinear thicknesses are produced, with a different thickness profile for each coating material. Flip masks — different masking for different layers — are designed if the performance of the optical system merits the expense.

9.5.9.2 Production a coating with a wedged thickness

Figures 9-55 to 9-58 depict the hardware that produces a coating in which the layer thicknesses are wedged. In Fig. 9-55, a helically shaped aperture is cut into the wall of a rotating tube. The occultation of the substrate, which is inside the tube, enabled Fabre [61] to deposit a magnesium fluoride layer along a 150 mm slab with a 6:1 thickness variation. O'Brien [39] used a rotating cam shaped mask shown in Fig. 9-56 to produce wedged neutral density coatings. Figure 9-57 depicts one of the many methods of oscillating a substrate. It is attached to a dual helix lead screw so that different portions are obscured from 0% to 100% during the deposition. In the United States, such dual helix lead screws are found in "Yankee" brand screwdrivers. The use of stepping motors in any of the aforementioned devices can introduce a nonlinear thickness variation.

Wedges have also been deposited on circular substrates where the thickness varies with azimuthal angle at a constant radius. As depicted in Fig. 9-58, one approach is to interpose two masks in front of a rotating substrate. One mask is nonrotating and the other mask rotates in a direction opposite to that of the substrate. The products of such masking are termed *circular variable filters* [Apfel 65, Thelen 65b], *universal filters* [Holm 68] and *variable wavelength filters* [Ward 71].

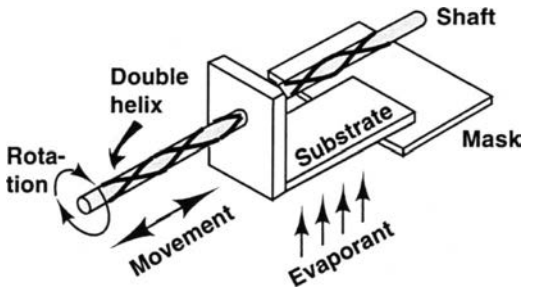


Fig. 9-57 During vacuum evaporation, a linearly wedged thickness film is produced on a substrate by oscillating it over a mask that shields it from the evaporant. The double helix thread produces the oscillatory motion.

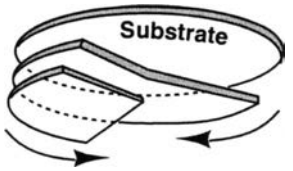


Fig. 9-58 During vacuum evaporation onto a circular substrate, contra-rotating masks produce a variable thickness film — wedged with increasing azimuthal angle. After Apfel [65].

9.6 The control of layer thickness during deposition

9.6.1 Non-optical methods

There are several ways to control film thicknesses. The first is to use a fixed process, where the film is deposited at a predetermined rate for a specified time. The second is to sample the thickness during deposition, which is termed *monitoring* in the optical coating industry. Radiometric and quartz crystal monitors are widely used. In addition, the thickness of an electrically conductive film can be controlled by measuring its sheet resistance (see §1.9.7) *in situ* during deposition. Schmidt [66], for example, established that molybdenum layers with a specified sheet resistance functioned optimally as photothermal solar absorbers and thus controlled the metal layer thicknesses in that manner.

9.6.1.1 Control by a fixed process

One approach is to control the deposition so precisely that a monitor is unnecessary. Suppose, for example, that multilayers are deposited by sputtering. The sputtering rate is controlled by adjusting the gas pressure and power. The layer thickness is determined by its time duration in the plasma.

The *weighed charge* method in vacuum evaporation also avoids monitoring. A predetermined mass of coating material is completely evaporated from one or more boats. It is desirable that the evaporant not wet or alloy with the boat. It has the advantage that it is relatively simple to use. Large areas are coated easily. Many metallic optical coatings are deposited by this method.

9.6.1.2 Quartz crystal thickness measurement

The sensor consists of a thin wafer of crystalline quartz metalized on both sides. It is connected to an electrical circuit that causes it to oscillate at a frequency of a few MHz. The crystal is positioned inside the vacuum tank. One side of the wafer is exposed to the evaporant stream, which adds mass as it condenses on its surface. This change in mass in turn produces a shift in frequency Δf, which is related to the thickness change Δh of the film by

$$\Delta h \approx 3\,\rho\,\Delta f \qquad (9\text{-}17a)$$

where ρ is the physical density of the film in g/cm^3 [Glang 70]. The crystal holder is water cooled to minimize frequency shift due to temperature variation.

The precision of the crystal in any given sampling time can be estimated from Eq. (9-17a). Substituting the ρ of aluminum (2.7 g/cm^3), a rule of thumb is that each 0.1 nm of thickness shifts the frequency by 1 Hz. Commercial instruments usually measure the period digitally over a 1 second averaging time. These are unusable for rapid evaporation because the thickness is sampled only five times during a 5 second deposition. This can be circumvented by faster circuitry.

The quartz crystal is indispensable for measuring the evaporation rate. For thickness measurements, its advantages are: (1) It is relatively easy to install, compared to an optical monitor; (2) several sensors may be active simultaneously in the vacuum chamber and their locations adjusted with relative ease; and (3) it can measure films that are thin—of the order of 10 nm. Such a feat is difficult with an optical monitor.

Its disadvantages are: (1) It measures the mass of the film, which translates into thickness only if its physical density remains constant; (2) the deposition of relatively thick layers causes them to flake and delaminate from the surface of the crystal, thus rendering the sensor useless; and (3) there may be a considerable temperature difference between the crystal whose maximum operating temperature is typically 70 °C and substrates that may be as warm as 300 °C. This produces problems because the sticking coefficient (see §9.5.1) may be quite different at these two temperatures.

Notwithstanding the aforementioned strictures, the author is aware of several optical coating facilities in which *all* coatings (including antireflection coatings at 10 μm) are crystal monitored.

9.6.2 Overview of "optical" monitoring

§9.6.2.1 mentions the reason for the quotation marks. The layer thicknesses are measured—"monitored"—during deposition by an optical probe beam that impinges upon a surface inside the chamber. That surface either reflects or transmits the probe beam. Two possible operation modes are: (1) radiometric monitoring—the radiant reflectance or radiant transmittance of that surface is measured; and (2) the state of polarization is determined. The former is most widely used.

Lostis [59] monitored polarimetrically by comparing the differential phase shift upon reflection of the coated and uncoated portions of a glass plate (or hypotenuse of a prism) as a nonabsorbing layer was deposited. He concluded that this apparatus could measure the optical thickness with the same precision as a "classical" method.

9.6.2.1 Advantages and disadvantages of radiometric monitoring

The term *radiometric monitoring* is more general than *optical monitoring* because the radiant transmittance or reflectance is measured at a broad range of wavelengths extending from the UV to the IR. Strictly speaking, *optical* refers to the visible portion of the spectrum. As an example, flux from the light source in Fig. 9-41 reflects from the mirror and passes through the window in the vacuum tank wall. It reflects from the monitor plate and reverses its path until it is collected by the detector. The electrical signal from the detector, which is amplified and presented in an analog or digital display, is proportional to the reflectance of the monitor plate.

An advantage of this type of monitoring is that the optical thickness *nh* [in Eq. (2-5b) or (2-37a)] is controlled at a particular wavelength. Thus, even if the index of a layer were to drift during deposition, at least the correct *nh* is obtained. The same is not true of the crystal monitor. Information is also provided about the refractive index of the film, its absorption and its optical homogeneity.

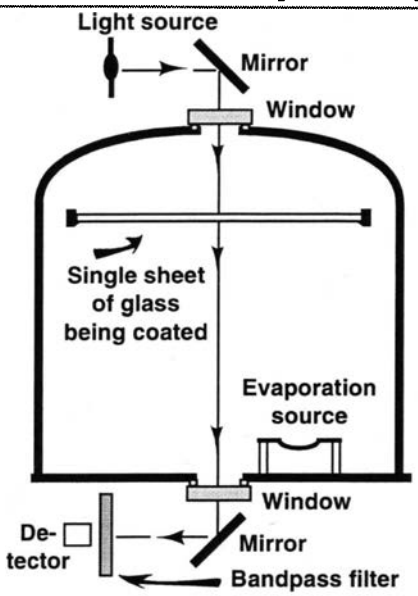


Fig. 9-59 A single sheet of glass is coated in a vacuum tank. The layer thicknesses are controlled by direct composite optical monitoring in transmission.

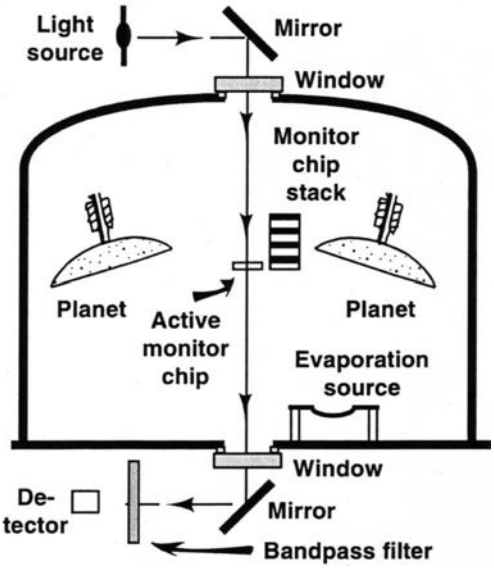


Fig. 9-60 A vacuum tank in which indirect optical monitoring in transmission is used to control the thicknesses of the layers deposited on the parts in the planets. The "coin changer" type of monitor chip apparatus is from Schwiecker [80].

The disadvantages of radiometric monitoring are: (1) The cost and complexity of its installation; (2) its lack of precision in controlling the thickness of relatively thin layers; and (3) the output is a quite nonlinear function of the film thickness, which makes it difficult to use for rate control. Radiometric monitors are expensive because one or more windows must be installed in the vacuum system. The optical train must be maintained in precise mechanical alignment and be free of drift and vibration.

9.6.2.2 Terminology used in radiometric monitoring

Words such as *direct, indirect* and *composite* describe optical monitoring systems and should be defined.

9.6.2.2.1 Monitoring — direct, indirect and composite

The monitoring probe beam is either reflected or transmitted by: (1) The substrate on which the multilayer itself is being deposited, as in Fig. 9-59; or (2) a separate surface (see Figs. 9-41 and 9-60). The latter is usually composed of glass and is termed the *monitor plate, monitor chip, monitor glass* or *test glass.* Monitoring through the multilayer itself is termed *direct monitoring*, whereas monitoring from the test glass is called *indirect monitoring*. *Composite* means that more than one layer is deposited on the

monitor chip. An illustration is that monitoring with a quartz crystal is, by its very nature, indirect and composite.

Table 9-4 Summary of monitoring schemes

	In reflection		In transmission	
	Indirect	Direct	Indirect	Direct
Recommended for...	Highly recommended as the optimum way to monitor most coatings. See §9.6.4.	Sometimes used for antireflection coatings. See §9.6.5.2.	Recommended for metallic films, if the stray light is suppressed.	Highly recommended for all-dielectric bandpasses. See §9.6.6.
Disadvantages	Not recommended for nearly opaque metallic films	Change in *signal* is small if the reflectance exceeds 95%. See Fig. 9-76.	Change in *signal* is quite small when the index of the film is close to that of the substrate.	Not recommended for blocking filters containing many layers. See §9.6.3.2.2.

Advantages and disadvantages of optical monitoring systems. *Signal* is the electrical output of the detector.

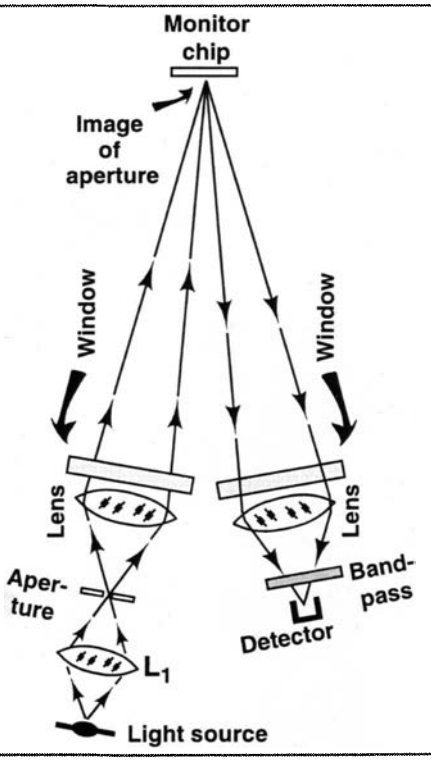


Fig. 9-61 The optical layout of a monitoring system, showing the light source, lens L_1 that images it upon the aperture and the lens that images the aperture (through the window in the wall of the vacuum tank) upon the monitor chip. The lens L_2 directs the flux reflected from the chip through a bandpass filter and thence upon the detector.

Table 9-4 lists four types of monitoring: (1) direct in reflection; (2) direct in transmission; (3) indirect in reflection; and (4) indirect in transmission. Which is best to use? The answer is dictated by the type of coating, the cost of installing and maintaining the system, the required precision, as well as the experience and level of confidence the operators have gained in its use. An example is a complex UV bandpass filter that I deposited when I was at the University of Rochester. It was necessary to switch between reflectance and transmittance — and also between direct and indirect — during the same coating run.

9.6.3 The hardware of an optical monitor

Thickness control seems easy when viewed superficially, with the mandate: "Terminate the deposition when the reflectance or transmittance attains a prescribed value." This ignores many important factors that relate to the precision and stability of the system, such as the relative change in signal during the deposition of a layer, the signal-to-noise ratio in the electronics, stray light and spectral bandwidth of the monochromator and constancy of the light source.

This section also discusses the monitoring of single layers.

9.6.3.1 Optical system of a monitor

The optical system in Fig. 9-61 is designed to maximize the flux transferred from the lamp to the detector. Other considerations in its design are: (1) The monitor probe beam should reflect from the glass monitor chip and nothing else; and (2) the detector should be shielded from stray light emanating from high temperature heaters or incandescent evaporants.

The flux collected by the detector depends upon both the field stop and aperture stop of the optical train. A lens L_1 images the light source—usually a tungsten filament—onto an aperture, which is the field stop. Another lens images the aperture onto the monitor chip. The flux reflects from the monitor chip and passes through another window before it is re-imaged through a bandpass filter upon the detector.

The size of the aperture mentioned above is important. Its diameter is adjusted so that the physical size of its image of the filament on the monitor chip is comfortably smaller than the chip itself. That is, the chip is usually held in a metal holder. Any flux spilling off the chip reflects from this holder and produces an erroneous monitor signal. Thus the field stop is determined by the physical dimensions of the chip and the linear magnification.

The aperture stop determines the spread in the cone of flux converging upon the chip. For a tank of predetermined size, it is usually dictated by the diameter of the window in the wall of the vacuum tank. That is, since the distance between the lens and the monitor chip is constant, any increase in the diameter of the lens decreases the f number and, therefore, transports more flux through the system and to the detector. The window size must also be increased if a larger lens is used. In more complex systems, the size of the field and aperture stops could be dictated by the monochromator slit width and, if used, the NA of a fiber bundle.

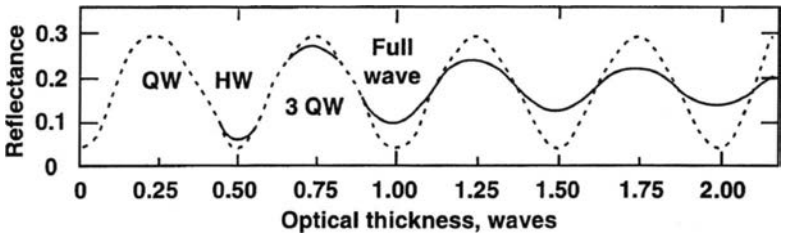


Fig. 9-62 Reflectance vs optical thickness as a layer of index 2.25 is deposited on a substrate of index 1.517. (*Dashed curve*): The light source is quasi-monochromatic. (*Solid*): The spectral shape of the light source is Lorentzian, with a 10% bandwidth. QW, HW, 3QW and Full wave refer to the reflectance extrema when the optical thickness attains those values, respectively.

9.6.3.2 Production of monochromatic light

A bandpass filter (see §1.6) or a monochromator is essential to establish the monitoring wavelength. It is unusual that the field and aperture stops are unaffected by insertion of a bandpass into the optical train. Hence the only flux loss is in the passband transmittance (T_{max} in Fig. 1-83), which is usually 50% or better. Most monochromators severely limit the flux. The loss can be minimized if the monochromator incorporates a diffraction grating with as large an area as possible [Jacquinot 54]. Hiraga [74], Bousquet [81] and Van Milligen [85] describe gratings with on-line, real time scanning of wavelength.

9.6.3.2.1 Spectral bandwidth of the monochromator

Suppose that a nonabsorbing layer of index 2.25 is deposited on a BK-7 substrate and its reflectance vs thickness is recorded. The dashed reflectance in Fig. 9-62 is relatively constant and occurs when the light source is quasi-monochromatic, as would be produced by a frequency stabilized laser. The damping of the reflectance peaks in the solid curve occurs because of excessive spectral bandwidth of the monochromator or bandpass filter.

9.6.3.2.2 Effects of stray light in the monochromator

Stray light means that a fraction of the light reaching the detector is from spectral regions other than the wavelength of the monitor. Offband leakage in bandpasses is discussed in §1.6.5. Monochromators that utilize diffraction gratings are notorious for their scattered light. The stray flux becomes a problem when the monitor signal is *self blocked.*

As an example of self blocking, suppose the dielectric mirror shown in Fig. 1-126 is monitored compositely in transmission at a tuned wavelength λ_0. After the deposition of 25 layers, the reflectance of the mirror is 99.9% or better and the flux reaching the detector at the monitor wavelength is decreased by as much as 1000. The problem is that although the mirror still transmits a little at λ_0 of 514 nm, it still has a relatively high transmittance in other broad spectral regions. An example is that the mirror in Fig. 1-126 transmits 60% or greater at wavelengths shorter than 460 nm and longer than 580 nm. It is the flux in those spectral regions that scatters in the monochromator. It produces an erroneous monitor signal if collected by the detector.

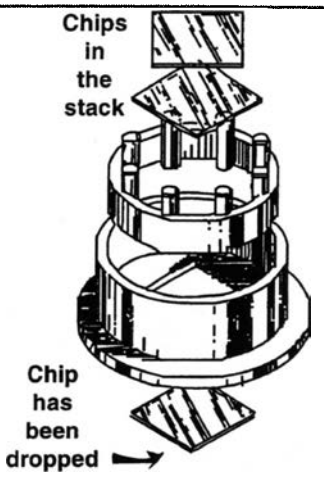


Fig. 9-63 In a chip changer for a reflection monitor, square chips are stacked, with alternate chips are rotated by 45°. Rotation of the square aperture below the bottom chip permits the chip to drop. After Bergfelt [68].

9.6.3.3 Misalignment of the optical monitor

One of its chief sources is mechanical distortion due to heat or mechanical vibration. Reflection monitoring is much more susceptible to this optical misalignment than is transmission monitoring. If possible, the detector should be larger than the image so that any image movement affects the signal minimally.

9.6.3.4 Light source stability

A separate detector that tracks the light source, along with an electronic ratiometer, can compensate for drifts in its output. A laser is not a particularly stable source. A fluctuating signal can occur when a laser beam passes through a window that changes its physical size due to thermal expansion. If the incidence angle is normal and the surfaces of the window are relatively parallel, the window acts like a Fabry-Perot interferometer that changes its transmittance with temperature, thereby modulating the optical probe beam. The probe beam is almost always modulated—usually by a mechanical chopper—to avoid dc drift and stray light pickup.

9.6.3.5 The chip changer and monitoring chips

There are at least two mechanisms that change the monitor chip. In the "coin changer" type in Fig. 9-60, a circular chip is slid from the bottom of the stack and moved laterally into position. The physical dimensions, especially the thickness, of the chip are critical.

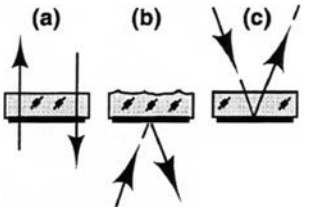


Fig. 9-64 A film (*solid black*) is deposited on the bottom side of the transparent monitor plate. It is measured:(*a*) In transmittance, with the light traveling in either direction; (*b*) In reflection, from its obverse side; or (*c*) In reflection, from its reverse side.

Figure 9-63 shows a chip changer that has been used at OCLI for many years. The square chips are stacked, with alternate chips rotated by 45°. A chip is dropped each time the square aperture in the bottom plate is rotated 45°. An advantage is that chips of *any* thickness can be intermixed. It is, however, strictly for monitoring in reflection.

As for the composition of the monitor chips, polished BK-7 is strongly recommended. I have also used chips of high index glass or sapphire. The latter were recycled—cleaned and stripped chemically and, because of their mechanical durability, they seemed to last forever.

9.6.3.6 Optical monitoring of a single layer

Both nonabsorbing and absorbing layers are considered.

9.6.3.6.1 Dielectric film

The lower side of the transparent monitor plate in Fig. 9-64 is coated with a nonabsorbing layer. The optical probe beam may be transmitted (panel *a*) or it can reflect from either its obverse side (panel *b*) or reverse side (panel *c*). When the obverse reflectance (see §1.2.2.1) is measured, the back surface of the plate is usually frosted to null the Fresnel reflectance at that interface.

The irradiances may be added (see §10.3.2.4), provided the coherence length of the light is shorter than the thickness of the plate. Using the terminology of §1.2.2.2, the monitor signal in transmission is approximately

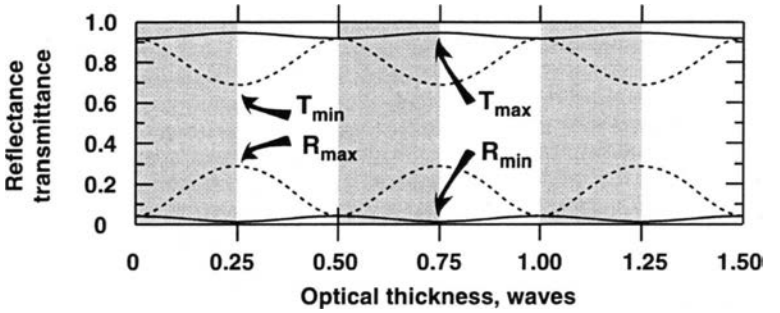


Fig. 9-65 Reflectance (two lower curves) and overall transmittance (two upper curves) vs optical thickness as a film of index 1.38 (*solid*) or 2.25 (*dashed*) is deposited on a bare glass slab of index 1.52. In computing the reflectance curves, reflection from the reverse side of the substrate is ignored.

$$T_{oa} \approx T_f T_e \ , \tag{9-17b}$$

in reflection from the reverse (i.e., uncoated) side of the plate

$$R_{oa} \approx R_e + T_e^2 R_f \tag{9-17c}$$

and from its coated side

$$R \approx R_f \ . \tag{9-17d}$$

Figure 9-65 plots idealized monitor signals (change in reflectance R and transmittance T_{oa}) vs optical thickness as films of relatively low index ($n = 1.38$) or high index ($n = 2.25$) are deposited on soda lime glass. The reverse side of the glass is frosted when obverse side reflection monitoring is used.

The key factor is the relative change in signal — the ratio of the R or T_{oa} of the uncoated glass to the maximum excursion in R or T_{oa} as the layer is deposited. Consider the system in which magnesium fluoride is monitored in transmission. T_{oa} increases from 91.7% (for the uncoated substrate) to a maximum of only 94.5% — an excursion of only 3%. Is this adequate? Yes, or no, depending upon the stability and the signal-to-noise ratio of the system.

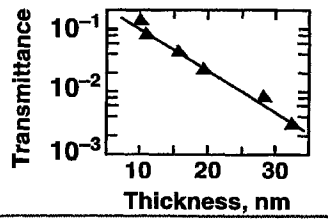


Fig. 9-66 Measured transmittance vs metric thickness as ruthenium films at $\lambda = 546$ nm were vacuum deposited on glass substrates at 300 °C. After Cox [74].

Figure 9-62 depicts the points where the optical thickness is a quarterwave, halfwave, three quarterwave and so on. Such higher order peaks occur when an IR coating, such as an antireflection coating containing QWOTs at 10 μm, is monitored at a wavelength of 1 μm.

9.6.3.6.2 Absorbing layers

This class of coatings includes metals and also semiconductors at wavelengths shorter than their cuton wavelengths. The R or T_{oa} (of the layer) vs thickness can be computed from equations in §2.4.2.9, provided the optical constants are known. Figure 2-49 shows that R reaches a plateau quite quickly when the thickness of the metal film exceeds 20 nm. The absorbance, however, continues to decrease linearly with metric thickness. This is illustrated by Fig. 9-66, which plots T_{oa} logarithmically (linear absorbance) vs thickness of a ruthenium layer [Cox 74]. The lower limit of the layer thickness that may be controlled in this manner depends only upon the stray light level in the optical monitor, the dynamic range of the detector and amplifier signal-to-noise ratio.

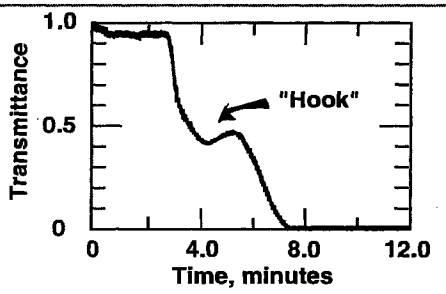


Fig. 9-67 Measured T as a silver film was deposited on glass. The "hook" is attributed to a change in film structure — from island to continuous. After Gur [77].

The computed simulation of the transmittance T vs thickness h of a metal film is difficult because of the island structure (see §3.3.2.3) that occurs in its early stage of formation. Its n and k are thickness dependent and usually known with dubious precision. The "hook" in the measured $T(h)$ for silver in Fig. 9-67 occurs at its transition from island to quasi-continuous structure.

A thin metallic film can be monitored in reflection by the *double pass method.* This provides more precise control in depositing relatively thin layers. The reverse (i.e., *back*) side of the monitor chip is pre-coated with a reflector such as opaque aluminum. The film is deposited on the obverse (i.e., *front*) side of the chip; hence, the monitor beam passes through it twice. Suppose a glass plate is monitored in transmission and the film deposition decreases its T_{oa} from 92% (net, through two air-glass surfaces) to 82%. Compare this with double-pass monitoring, where the signal "swing" is from 92% to 74%. The greater excursion of the signal means that the thickness can be controlled more precisely.

9.6.4 Indirect monitoring in reflection

9.6.4.1 Introduction

As noted in Table 9-4, composite indirect monitoring in reflection usually offers the most accurate thickness control — for nonabsorbing layers, at least. The attainment of any

"excellent" level of control, however, is achievable only if the refractive indices and control ratios are known — and known well.

9.6.4.2 Control ratio — its definition

Figures 9-1, 9-41 and 9-60 depict indirect monitoring systems. There is, for example, in Fig. 9-41 both a monitor plate and a planet holding the substrates — often called the *rack.* Suppose that a dielectric layer is deposited. A film of optical thickness $n_m\, h_m$ is deposited on the monitor and simultaneously a substrate in the rack is coated with a layer of optical thickness $n_r\, h_r$. Of paramount importance is the *control ratio* $\mathcal{M}$

$$\mathcal{M} \equiv n_m\, h_m / n_r h_r \ . \tag{9-17e}$$

The success or failure of the coating depends on the thoroughness with which $\mathcal{M}$ is known, its stability during the deposition and its constancy from run to run. Unlike the "configuration factor" of Van der Laan [85], which is defined in terms of *metric* thickness, $\mathcal{M}$ is the ratio of the *optical* thicknesses.

The coefficient $\mathcal{M}$ is greater than 1.00 in virtually all of the coating machines that the author has encountered in his experience as an optical coating engineer. This means that the monitor plate is coated *thicker* than the parts in the rack. As an illustration, suppose the design requires the deposition *on the rack* of a nondispersive film with a QWOT at 500 nm in a system with an $\mathcal{M}$ of 1.3. This is accomplished by adjusting the monitor wavelength to compensate for the monitor ratio. Solving Eq. (9-17e) for the optical thickness deposited on the monitor chip provides

$$n_m\, h_m = \mathcal{M}\, n_r\, h_r = 1.3 \times (500 \text{ nm}) = 650 \text{ nm} \ . \tag{9-17f}$$

In this instance, a QWOT of 650 nm is deposited on the monitor to obtain a 500 nm QWOT on the substrate held in the rack.

The next two sections present two methods of determining the control ratio. The procedure in §9.6.4.3 is quite general and involves only one film. The method of §9.6.4.4 is particularly useful if a high and low index are deposited — sequentially.

9.6.4.3 Determination of the control ratio for a single layer

The procedure to determine the control ratio is:

(1) A substrate consisting of the same material as the monitor chip is loaded into one or more places in the rack of the coating machine. These are the "parts."
(2) A relatively thick layer (such as two full waves) is coated on both the monitor chip and parts.
(3) The $T(\lambda)$ of each of the "parts" is measured on a spectrophotometer (see §10.4.3).
(4) The optical thickness ratio of the coating on the monitor to the coating on the part on the rack is determined from the spectrophotometer data.

In the implementation of the foregoing procedure, a few comments. A measurement of transmittance is preferred because it is, in general, more reliable than a reflection measurement. After its removal from the chamber, the rear of the monitor chip (which is usually frosted) is "oiled" with a matching fluid to a substrate of the same thickness and composition as the chip. The transmittances of both "chip" and "rack" samples are measured on a spectrophotometer.

The wavelengths λ_m at which the halfwave extrema occur (see §2.4.2.6) are measured. These wavelengths are converted into wave numbers σ_m and are plotted vs an integer m. According to Eq. (2-23g), the *slope* of the curve is the optical thickness. A

least squares software package may be used to enhance the precision of the determination of the slope. The ratio of the slopes is the control ratio.

A final comment is that the transmittance vs wave number is usually measured in the IR. This is quite desirable. To the extent that dispersion is minimal, the abscissa in Fig. 2-41 could be replaced with wave number. The transmittance maxima and minima are equally spaced in wave number.

9.6.4.4 Determination of the control ratio using the Grèzes-Besset method

The classical Fabry-Perot is described in §7.2. Grèzes-Besset [95] employed a nine-layer Fabry-Perot to determine the uniformity of coating thickness. Each reflector contains four layers. §7.2.2 shows a bandpass with an air spacer in which each reflector contains five layers. Since the bandpass of Grèzes-Besset contains a high index layer for the spacer, the surrounding reflectors contain four layers. The order number of the spacer is three (see §2.11.4.1.1) and the bandpass design is

air **H L H L** 6**H L H L H** *substrate*

where **L** and **H** represent layers of quarterwave optical thickness of low and high refractive index, respectively. The foregoing design is simplified by reducing the layer count to seven. A spacer of full wave optical thickness (i.e., second order) is employed

air **H L H** 4**L H L H** *substrate.*

The refractive index (which is dispersive) of each layer is given by Eq. (2-2c), where A and B appear in Table 9-5. In the example below, the refractive indices of the layers approximate those of silica (for **L**) and titania (for **H**).

The next question is: Given the foregoing design, what is the sensitivity of the transmittance to a thickness error in each layer? *Sensitivity*, as defined in §9.6.4.4.5, implies an error of ±2%. The sensitivity is computed at the peak of the passband transmittance. Figure 9-68 shows that the center spacer layer of full wave optical thickness is most sensitive and that the outer layers of the stack manifest a sensitivity that dwindles almost to zero.

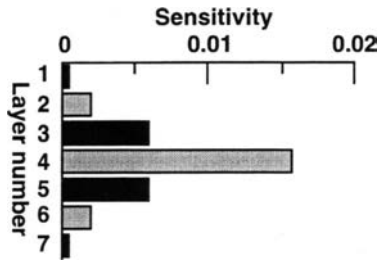


Fig. 9-68 At a wavelength λ_0 of 600 nm, *sensitivity* (defined in §9.6.4.4.5) of the transmittance to a thickness error in each of the layers in the design

air **H L H** 4**L H L H** *substrate*

where **H** and **L** represent layers of optical thickness $\lambda_0/4$. The substrate refractive index is 1.517. The refractive index data [using Eq. (2-2c)] for **H** and **L** are in Table 9-5.

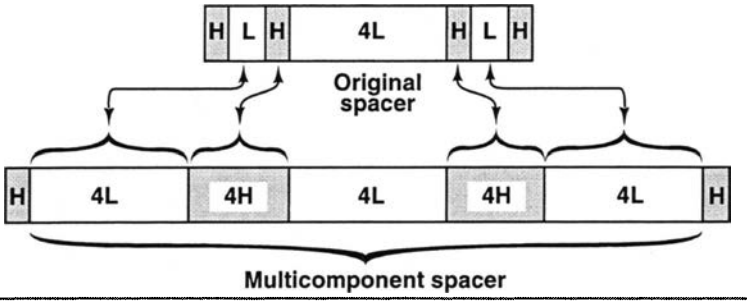


Fig. 9-69 The metric thicknesses of several layers of a classical Fabry-Perot are modified to form a Fabry-Perot with a multicomponent spacer.

9.6.4.4.1 Modification of the Grèzes-Besset stack

The quest is to devise a stack in which the sensitivities of the various layers are less disproportionate than those in the Fabry-Perot above. One approach is to use a multicomponent spacer.

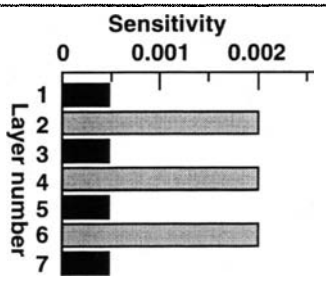


Fig. 9-70 The caption to Fig. 9-68 obtains, with the exception that the design is

air **H** 4**L** 4**H** 4**L** 4**H** 4**L** **H** *substrate.*

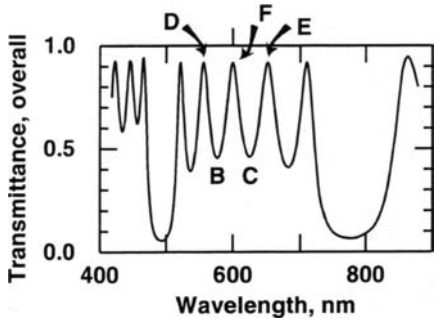


Fig. 9-71 Overall transmittance (see §1.2.2.2) of the coating captioned in Fig. 9-70.

Figure 9-69 shows that this spacer is fabricated by expanding the thicknesses of each of the four inner quarterwave thickness layers into a full wave. The design is still a Fabry-Perot — there are single quarterwave optical thickness layers at either end of the stack, which are the reflectors. The difference is that the spacer of this Fabry-Perot is *multicomponent*. That is, the spacer is composed of five contiguous layers, each of which is full wave optical thickness

air **H** 4**L** 4**H** 4**L** 4**H** 4**L** **H** *substrate* .

The symbols have the same meaning as previously. Figure 9-70 illustrates that the relative disparities between the sensitivities of the layers, while not reduced to zero, are still much less than in Fig. 9-68. This is an advantage of using this design.

9.6.4.4.2 Example — visible spectral region

Suppose that a Fabry-Perot with a multicomponent spacer is deposited in the visible portion of the spectrum. Each of the layers is a quarterwave (or multiple thereof) at 600 nm — which defines the center wavelength λ_F. Figure 9-71 depicts the spectral transmittance of this coating. There is a central maximum (designated as *F*) that

is reasonably close to the center wavelength. Other maxima adjacent to F are identified by D and E. They occur at wavelengths λ_D and λ_E. In addition, the transmittance dips to a minimum at points B and C adjacent to F. The transmittances are T_B and T_C at the minima at the short-wavelength position B and long-wavelength position C, respectively.

Two pieces of information are gleaned from the pseudo-spectrophotometric trace depicted in Fig. 9-71: (1) The *average* optical thickness of the two materials used in the stack — i.e., the *centering*; and (2) the *ratio* of the optical thicknesses of the two materials. Each is discussed separately.

9.6.4.4.3 Determination of the wavelength centering

The center wavelength λ_{actual} that is actually attained is the average—on a wave number scale—of the peak transmittances at E and D

$$\lambda_{actual} = 2\,(\,\lambda_E^{-1} + \lambda_D^{-1}\,)^{-1}\ . \tag{9-18a}$$

If the control ratio is correct and the refractive indices have been determined properly, λ_{actual} equals λ_F. The foregoing equation is essentially identical to Eq. (4-10).

9.6.4.4.4 Determination of the optical thickness ratio

The ratio of the transmittances at the minima B and C furnishes information about the ratio of the optical thicknesses. The curves in Fig. 9-72 present an enlarged view of the data in Fig. 9-71. For the solid curve, all low index films are 1% thick and the high index layers are 1% thin. The converse is true for the dashed curve. The net effect is that the 1% thickness change induces a transmittance difference between the shorter wavelength minimum and longer wavelength minimum of at least seven percent.

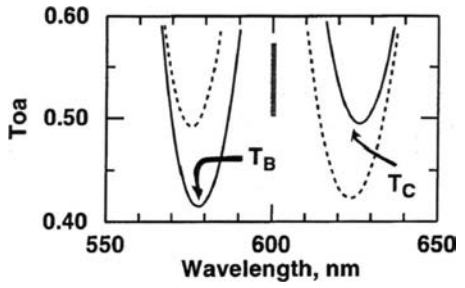


Fig. 9-72 The transmittance of the coating captioned in Fig. 9-70, with the exception that for the *solid* curve all low index layers are 1% thick and all high index layers are 1% thin. For the *dashed* curve, all low index layers are 1% thin and all high index layers are 1% thick.

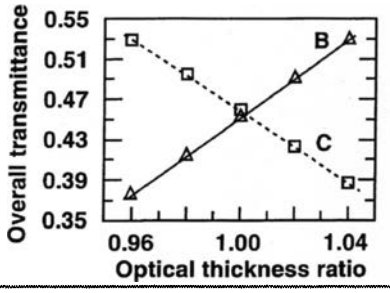


Fig. 9-73 Versus optical thickness ratio [see Eq. (9-18b)], transmittances T_B and T_C of the minima in Fig. 9-72.

The *optical thickness ratio* is defined as

$$\textit{optical thickness ratio} \equiv \frac{\text{Optical thickness of } \mathbf{H}}{\text{Optical thickness of } \mathbf{L}}\ . \tag{9-18b}$$

Figure 9-73 graphs the transmittances at minima B and C vs optical thickness ratio. The conclusion is that, for a few percent excursion of the thickness ratio, the variation is quite close to linear. The data from a spectrophotometer trace make it possible to determine the optical thickness ratio. This information is used, in turn, to adjust $\mathcal{M}$ in Eq. (9-17e).

A few comments are in order. The first is that if there were no dispersion and no thickness errors, the transmittance minima at *B* and *C* in Figs. 9-70 and 9-72 would be equal. The data in Fig. 9-73 are presented for materials that represent layers of silica and titania. Computer simulation may be used to generate similar curves for other pairs of coating materials. Also, information about the extinction coefficients of the films may be obtained from the peak transmittances—provided that the absorption is relatively small. It is also possible to use fewer than or more than five layers in the multicomponent spacer.

9.6.4.4.5 Definition of the sensitivity

The sensitivity is determined as follows. The metric thickness of a given layer in the stack is adjusted by a random *amount*. That *amount* is obtained from a table of 256 random numbers that approximate a Gaussian distribution. All 256 table entries are used to produce an array of transmittances (of the entire stack). For the peak transmittance closest to the center wavelength, the *sensitivity* is the lower quartile transmittance in the array. The aforementioned Gaussian distribution was inspected to ensure that it approximated the erf function.

In each case, the thickness error is ±2% for each *final* quarterwave thickness. Thus, if the layer is four quarterwaves (i.e., a full wave) in optical thickness, a thickness error of ±2%/4 = ±0.5% is used for that layer. This is realistic in terms of optical monitoring, since it is the uncertainty in the deposition of the *final* quarterwave thickness of the layer that determines the thickness error. The 2% or 0.5% "thickness error" cited above is the variance of the Gaussian distribution.

9.6.4.5 Composite monitoring of pairs of dielectric layers

Many coatings consist of alternate layers of high index **H** and low index **L**. A recommended procedure is to deposit first an **H** layer on an uncoated monitor chip and then an **L** layer on top of it. That monitor chip is then discarded and the procedure is repeated with a fresh chip.

9.6.4.5.1 Procedure

It is assumed that the thickness of each layer is specified in QWOTs at a reference wavelength λ_0. Given that refractive indices of the layers are dispersive, the following procedure is used to generate the reflectance curves for optical monitoring:

(1) The refractive index of each layer is found at λ_0.
(2) Equation (2-47i) is solved for h and the metric thickness of each layer is found.
(3) A monitor wavelength λ_w is selected. If possible, it should be not too distant from λ_0.
(4) The refractive index n_w at λ_w is found for each layer.
(5) The optical thickness OT (or QWOT) deposited on the monitor chip is

$$\text{OT} = n_w h \mathcal{M} \quad \text{or} \quad \text{QWOT} = 4 n_w h \mathcal{M} \tag{9-18c}$$

for each layer.
(6) Compute the reflectance vs optical thickness for the single layer **H**—its optical thickness is found in step (5).
(7) Compute the reflectance vs optical thickness of the **L** layer of the bilayer *air* **L H** *substrate*.

(8) If the reflectance curves for both the single **H** layer and bilayer terminate well away from either a maximum or minimum, λ_w is accepted as a monitoring wavelength for that pair.

(9) If the foregoing is not true, another monitor wavelength is selected and steps (4) to (8) are repeated.

The foregoing steps are repeated for each pair of layers in the stack.

Table 9-5 Optical monitoring of two layers

	A	B	$\mathcal{M}$	h	QWOTs¶	n, 1060 nm	n, 740 nm	QWOTs*
L	2.05	.05	1.25	187.9	1.026	1.447	1.4633	1.858
H	5.34	0.20	1.20	159.9	1.417	2.350	2.389	2.478

Optical properties of **L** and **H** layers, whose refractive indices are dispersive [A and B appear in Eq. (2-2c)] and whose optical thicknesses are defined as QWOTs at 1060 nm. $\mathcal{M}$ is the control ratio [see Eq. (9-17e)] and h is the layer's metric thickness in nm. ¶At 1060 nm. *The QWOTs (at 740 nm) deposited on the monitor chip to produce the required thickness on the part.

9.6.4.5.2 Example of monitoring

The design of a reflector is captioned in Fig. 5-54. It is centered at $\lambda_0 = 1060$ nm and the first two layers "1.026**L** 1.417**H**" are to be deposited by optical monitoring. The **H** layer is deposited first on the bare monitor chip. Using the refractive index values in Table 9-5, the metric thicknesses in the h column are generated. A monitor wavelength of 740 nm is chosen. The QWOTs deposited on the monitor chip are obtained from Eq. (9-18c) and appear in the right-most column of Table 9-5. A computer code using admittance recursion (see §2.4.5.1) generated the curves in Fig. 9-74.

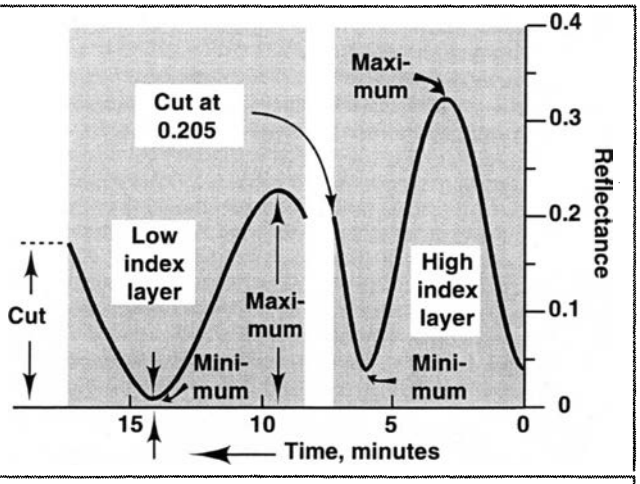


Fig. 9-74 At a constant layer deposition rate, reflectance vs time as the two layers in Table 9-5 are deposited sequentially upon a substrate of refractive index 1.517. The wavelength is 740 nm.

During the deposition of the **H** layer, the reflectance R starts at 4.22% (that of a bare BK-7 substrate), rises to a maximum of 33.6% and drops again to the R of the uncoated substrate. The deposition is terminated when the R finally is rising and attains 20.5%, as listed in Table 9-6.

The **L** layer is deposited directly on top of the **H** layer. Figure 9-74 shows that the reflectance of the bilayer rises to a maximum 2% higher than the starting value and then plunges to a minimum of

1.8%. The deposition is terminated when the reflectance is 17.3% and rising.

Table 9-6 Optical monitoring of two layers

Layer	Comment	R_{start}	R_{min}	R_{max}	R_{cut}	f_c
H	Deposited first	0.0422*	0.0422*	0.3363	0.2054	0.555
L	Deposited second	0.2054	0.0186	0.2268	0.1722	0.738

The **H** layer listed in Table 9-5 is deposited on a substrate of refractive index 1.517. At a wavelength of 740 nm, R_{max}, R_{min} and R_{cut} are, respectively, the maximum reflectance, minimum reflectance and reflectance at the termination of the deposition. The same terminology applies to the **L** layer, which is deposited on top of the **H** layer.
f_c appears in Eq. (9-18d). *R of the uncoated substrate.

9.6.4.5.3 Fractional cuts

The *fractional cuts* are listed in the *run sheet*—the document that controls the monitoring of the coating machine. This information makes it easy for the coating machine operator to compensate for changes in the refractive indices during the deposition. In other words, the reflectance values, including the "cut" value of reflectance, are predicated on an assumption that certain refractive indices are attained during the deposition of those layers. Any drift in the indices is manifested as a change in the reflectance values.

If R_{max}, R_{min} and R_{cut} are, respectively, the maximum reflectance, minimum reflectance and reflectance at the termination of the deposition, the definition of the fractional cut f_c depends upon whether the reflectance has encountered a maximum and is decreasing towards a minimum, or vice versa. If the reflectance has previously hit a minimum and is rising, f_c is

$$f_c \equiv \frac{R_{cut} - R_{min}}{R_{max} - R_{min}} , \tag{9-18d}$$

which may be inverted to

$$R_{cut} = R_{min} + f_c (R_{max} - R_{min}) . \tag{9-18e}$$

If the reflectance has previously hit a maximum and is decreasing, f_c is

$$f_c \equiv \frac{R_{max} - R_{cut}}{R_{max} - R_{min}} , \tag{9-18f}$$

which may be altered to

$$R_{cut} = R_{max} - f_c (R_{max} - R_{min}) . \tag{9-18g}$$

As an example, Table 9-6 lists the maximum, minimum and cut reflectance values as the **H** layer captioned in Table 9-5 is deposited. Table 9-6 shows that f_c is 0.5548 for that layer. These data, however, are predicated on an *assumed* refractive index of 2.389 (at λ = 740 nm) for that **H** layer. Suppose that index is actually lower—2.32. In this case, the maximum reflectance is not the 33.63% listed in Table 9-6, but rather a lower value of 31.39%. It is important, however, that the correct *optical thickness* be deposited. This is done by adjusting the reflectance of the "cut." Using the fractional cut of 0.5548 in the right-most column of Table 9-6, Eq. (9-18e) provides the new cut value of

$$R_{cut}]_{actual} = 0.0422 + 0.5548\ (0.3139 - 0.0422) = 0.1929 , \tag{9-18h}$$

which is slightly in excess of 0.01 units lower than the assumed cut of 0.2054 in the sixth column of Table 9-6.

Why do fractional cuts work and for what reasons are they used? Remember that the optical monitor system *automatically* compensates for the index drift—in terms of depositing the correct optical thickness, irrespective of any change of refractive index. In the above example, the *actually deposited* index of 2.32 is nearly 3% lower than the *assumed* index of 2.289. Consider the process as that layer is deposited and its reflectance (i.e., the monitor signal) rises to its first maximum—its "turning point." The fact that the optical thickness of this layer is a QWOT means that its physical thickness is automatically made 3% *greater*, to compensate for the 3% *smaller* refractive index. The use of fractional cuts extends this compensation to "cuts" of the monitor signal at reflectance values other than the "turning point."

9.6.4.6 Nucleation and drift of the control ratio

Both delay of the nucleation and temperature differentials are problems in optical monitoring. Also of concern is the change of refractive index as the coatings are transported from a vacuum environment to the humidity of ambient air (see §3.4.3.2.1).

9.6.4.6.1 Delay in nucleation

As discussed in §3.3.2.2, the evaporant molecules may impinge upon the substrate, yet no film growth takes place until a substantial number of nuclei form. The film begins to grow as these nuclei collect more molecules and grow larger. The nucleation is retarded if the substrate is extremely smooth, as, for example, a surface of fire-polished glass. The nucleation proceeds with minimal delay if there is already a film on the surface.

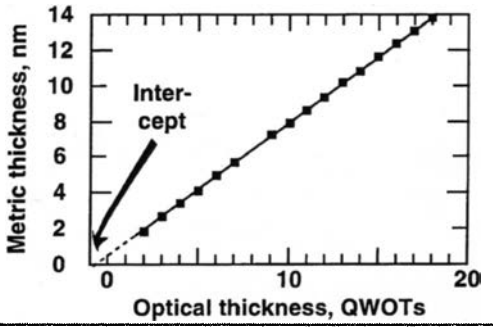


Fig. 9-75 Metric thickness (as recorded on a quartz crystal monitor) as a single zinc sulfide layer was deposited upon a glass substrate. The thicknesses were recorded at the extrema of reflectance during deposition. The wavelength is 950 nm. Courtesy of Coherent [92].

The aforementioned effects are illustrated by the deposition of zinc sulfide, which is used extensively for IR coatings. The worst case occurs when there is *already* a coating on the parts in the rack, which makes it easy for the zinc sulfide to nucleate—a layer starts forming immediately on the parts. The situation is much different on the glass monitor chip. Due to the fact that it is uncoated, the nucleation takes a much longer time on it than on the parts in the rack. As mentioned below, the chamber also contained a quartz crystal thickness measuring device (see §9.6.1.2). The crystal is also precoated before the zinc sulfide impinges upon it. Hence there is minimal time delay for nucleation to take place upon it.

This is illustrated in Fig. 9-75, which plots the metric thickness of the quartz crystal instrument vs the optical thickness (in QWOTs of approximately 950 nm) on the monitor chip. The data were obtained by noting the quartz crystal reading at each maximum and minimum of the optical monitor. There is a linear extrapolation to zero optical thickness. The negative intercept on the abscissa in Fig. 9-75 is clear evidence that there is approximately one-half of a QWOT on the crystal (and presumably on the substrates in the rack) when the film finally begins to nucleate and form on the monitor.

One solution is to pre-deposit on the monitor chip an ultra-thin (such as 2 nm thickness) layer of *something* that will assist the zinc sulfide in its nucleation.

9.6.4.6.2 Temperature differential

The problem of temperature differential is illustrated by Fig. 9-27, which shows there is a 30 °C difference in temperature of substrates in the rack. Such temperature differentials alter both the nucleation time and the sticking coefficients (see §9.5.1.1).

Suppose that noncomposite monitoring is used and a fresh chip is dropped into place for each layer (or pair of layers). What is the chip temperature relative to the substrates in the rack? The substrates are usually exposed directly to the radiant heaters (see §9.3.4) and therefore are "roasted and toasted" quite effectively. The monitor chips, on the other hand, are usually thermally isolated from the heaters and consequently are cooler than the substrates. This temperature differential affects both the refractive index and sticking coefficient—and, hence $\mathcal{M}$. The stack of monitor chips may warm up as the deposition progresses. Compensation for this temperature drift is made by adjusting $\mathcal{M}$. That is, a different $\mathcal{M}$ is used at the end of the deposition (when the monitor chips are relatively warm) than at the beginning (when the monitor chips are relatively cool). Changes in $\mathcal{M}$ also compensate for variations in the sticking coefficient, as, for example, its dependence upon the composition of the undercoats on either the chip or substrates.

9.6.4.6.3 Refractive index shift

More precise monitoring is achieved if the refractive index of the film is known while it is *in* the vacuum tank. Those indices are used to generate the optical monitoring data. In other words, the refractive indices used to design the coating are not those that occur in the vacuum tank as the layers are deposited. The *wet-dry* shift of index cited in Table 3-3 introduces uncertainties if the reflectance of a coating is measured after its removal from the vacuum.

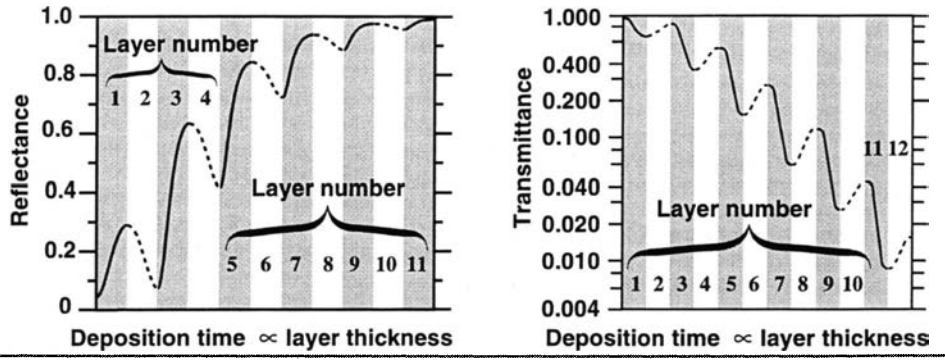


Fig. 9-76 Versus the optical thickness of each layer, reflectance (linear scale) and overall transmittance (log scale) as layers of alternating high refractive index (*solid*) and low index (*dashed*) of quarterwave optical thickness are deposited upon a glass substrate (see Table 9-7). The layers are numbered in the order in which they were deposited.

Table 9-7
Reflectance and transmittance during deposition of QWOTs

Layer	Design	R	T_{oa}
0	*glass*	0.043	0.916
1	H *glass*	0.289	0.680
2	L H *glass*	0.074	0.886
3	H L H *glass*	0.635	0.349
4	L H L H *glass*	0.417	0.558
5	H L H L H *glass*	0.844	0.150
6	L H L H L H *glass*	0.723	0.265
7	H L H L H L H *glass*	0.938	0.059
8	L H L H L H L H *glass*	0.885	0.110
9	H L H L H L H L H *glass*	0.976	0.022
10	L H L H L H L H L H *glass*	0.955	0.043
11	H L H L H L H L H L H *glass*	0.990	0.0086
12	L H L H L H L H L H L H *glass*	0.983	0.016
13	H L H L H L H L H L H L H *glass*	0.997	0.0033

At the turning point, reflectance R (from the air side) and overall transmittance T_{oa} as nonabsorbing layers (numbered as deposited) of quarterwave optical thickness are deposited on a glass substrate of index 1.52. Layer indices are $n_H = 2.25$ and $n_L = 1.38$.

9.6.4.6.4 Depletion of the coating material

During a long coating deposition of 120 layers, the author has adjusted $\mathcal{M}$ several times to compensate for the depletion of the coating material in the crucibles of the electron guns. The results were noticeable.

9.6.5 Composite direct monitoring

§9.6.5 discusses composite direct monitoring of a quarterwave stack and also an antireflection coating. Indeed, many layers are deposited on the same substrate. §9.6.6 reviews the problems encountered when a bandpass is monitored in this manner.

9.6.5.1 Deposition of a quarterwave stack

Table 9-7 and Fig. 9-76 display the R and T_{oa} (on an absorbance scale) of a compositely monitored stack of quarterwave layers

air **H L H L H L H L H L H L H** *glass* .

The properties of this type of stack are discussed in §5.2.5.1.

The R and T_{oa} oscillate from a maximum to a minimum (or vice versa), depending on whether the layer count is odd or even. The monitoring signal in reflectance is adequate for the first six layers but saturates when R exceeds 90%. This saturation does not occur when the absorbance is measured. As long as the gain of the amplifier can be increased, the monitoring of the eleventh layer is done with the same precision as the fifth layer because the "swings"—the changes between a minimum and a maximum —are essentially the same.

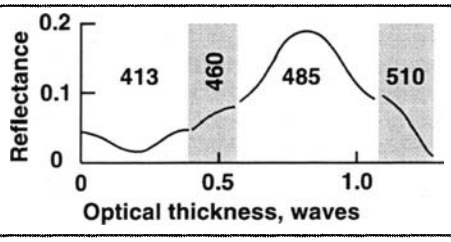


Fig. 9-77 Reflectance vs optical thickness as the antireflection coating is deposited

glass 0.385 **L** 0.204 **S** 0.5 **H** 0.25 **L** *air*

where the design is listed *in the order* in which the layers are deposited. The optical thickness of **L** and **H** is 510 nm. n_L, n_S and n_H are 1.38, 1.65 and 2.15, respectively. 413, 460, etc. are the wavelengths of the monitor (in nm), which is changed for each layer. After Uetsuki [75].

The foregoing is termed *turning point monitoring* because the deposition of each layer is terminated when its R or T_{oa} attains a maximum or minimum. Its advantages are: (1) It produces a layer with a QWOT at the wavelength used in the monitoring system, independent of the dispersion of the refractive index of the film; and (2) there is *error compensation* in its optical thickness for certain types of designs, as discussed subsequently. Its detractions are: (1) Skilled operators are needed to obtain precise thickness control; and (2) it is unsuited for the deposition of thin layers — films that are fractions of a QWOT.

9.6.5.2 Deposition of an antireflection coating

In general, it is much more difficult to monitor compositely if the "cuts" during the depositions of the layers are not made at turning points. This is illustrated in Fig. 9-77, which shows the composite reflectance monitoring of a four-layer antireflection coating [Uetsuki 75]. The monitor wavelength is changed for each layer. This type of monitoring functions superbly and with more precision than turning point monitoring, provided the refractive indices of the films *in the vacuum* are known with reasonable certainty and do not drift during the coating deposition.

9.6.6 Composite, direct monitoring of a bandpass filter

9.6.6.1 Introduction

As mentioned in Table 9-4, this method is widely used in the deposition of bandpasses. The monitor probe beam is directly transmitted through the substrate being coated, as in Fig. 9-59. One of its virtues is the absence of any compensation for the control ratio, because the actual part is measured.

9.6.6.2 Example — two-cavity bandpass

Composite turning point monitoring is used to deposit a two-cavity all-dielectric bandpass (see §7.7.2)

air **H L H L H 2L H L H L H L H L H 2L H L H L H L** *glass* .

§9.6.6.2.1 shows that this method of monitoring is advantageous because of compensation for thickness errors. §9.6.6.2.2 examines the problem of the null point that may occur due to drift in refractive index.

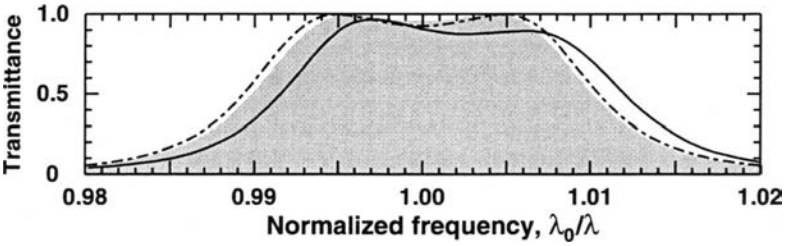


Fig. 9-78 $T(\lambda_0/\lambda)$ of a bandpass with no errors in any of the layer thicknesses (*shaded*)

air **(H L)^2 H 2L (H L)^5 H 2L (H L)^3** *glass* ,

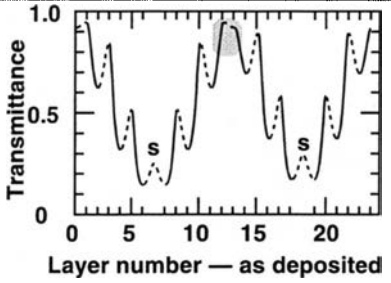
the **LL** spacer nearest the substrate is 10% thin — but the thickness of the adjacent **H** layer is adjusted by optical monitoring to compensate for that thinness (*long-short dash*)

air **(H L)^2 H 2L (H L)^5 1.308H 1.8L (H L)^3** *glass*

and the **LL** spacer nearest the substrate is deposited 1% thin (*solid*)

air **(H L)^2 H 2L (H L)^5 H 1.98L (H L)^3** *glass* .

The indices of **L**, **H** and *glass* are 1.38, 2.22 and 1.52, respectively.
The optical thickness of **H** and **L** is $\lambda_0/4$.



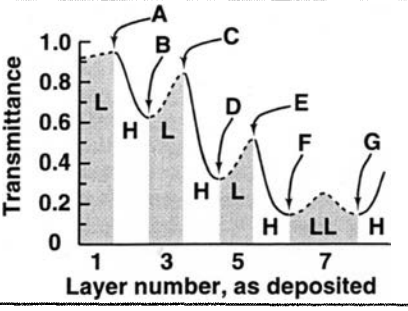


Fig. 9-79 Total overall transmittance of the substrate at λ_0 as each layer is deposited of the bandpass

vacuum $(\mathbf{H}\,\mathbf{L})^2\,\mathbf{H}\,2\mathbf{L}\,(\mathbf{H}\mathbf{L})^5$

$\mathbf{H}\,2\mathbf{L}\,(\mathbf{H}\,\mathbf{L})^3$ *glass*

where the indices of **L**, **H** and *glass* are 1.38, 2.22 and 1.52, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$. The shaded area is magnified in Fig. 9-82.

Fig. 9-80 The total overall transmittance of the substrate at λ_0 as each of the initial seven layers (and part of the eighth layer) is deposited of

vacuum $\mathbf{H}\,2\mathbf{L}\,(\mathbf{H}\,\mathbf{L})^3$ *glass* .

The indices of **L**, **H** and *glass* are 1.38, 2.22 and 1.52, respectively. The optical thickness of **H** and **L** is $\lambda_0/4$. The letters are identified in Table 9-8.

Table 9-8 Composite transmission monitoring of a bandpass

Layer* and symbol	R	T	T_{oa}	Comment
0	0.0426	0.9574	0.9166	Bare substrate
1 **L**	0.0126	0.9874	0.9453	A in Fig. 9-80
2 **H**	0.3536	0.6464	0.6189	B in Fig. 9-80
3 **L**	0.1208	0.8792	0.8417	C in Fig. 9-80
4 **H**	0.6742	0.3258	0.3119	D in Fig. 9-80
5 **L**	0.4690	0.5310	0.5084	E in Fig. 9-80
6 **H**	0.8591	0.1409	0.1349	F in Fig. 9-80
7 **L** At maximum	0.7485	0.2515	0.2408	Spacer
7 **L** At completion	0.8591	0.1409	0.1349	G in Fig. 9-80
8 **H**	0.4690	0.5310	0.5084	
9 **L**	0.6742	0.3258	0.3119	
10 **H**	0.1208	0.8792	0.8417	
11 **L**	0.3536	0.6464	0.6189	
12 **H**	0.0126	0.9874	0.9453	
13 **L**	0.0426	0.9574	0.9166	
14 **H**	0.2794	0.7206	0.6899	
15 **L**	0.0676	0.9324	0.8927	
16 **H**	0.6194	0.3806	0.3644	
17 **L**	0.3970	0.6030	0.5773	
18 **H**	0.8317	0.1683	0.1611	
19 **L** At maximum	0.7035	0.2965	0.2839	Spacer
19 **L** At completion	0.8317	0.1683	0.1611	
20 **H**	0.3970	0.6030	0.5773	
21 **L**	0.6194	0.3806	0.3644	
22 **H**	0.0676	0.9324	0.8927	
23 **L**	0.2794	0.7206	0.6899	
24 **H**	0.0426	0.9574	0.9166	Completion

*Layer number in order in which deposited.

At the termination of the deposition of a layer, reflectance R, transmittance T and total overall transmittance T_{oa} in vacuum at a wavelength λ_0 for a slab of refractive index 1.52 uncoated on one side. It is coated on its opposite side with the nonabsorbing layers listed in the first column. All layers are $\lambda_0/4$ optical thickness, with the exception of layers 7 and 19, which are $\lambda_0/2$ in thickness. The index of **H** is 2.22 and the index of **L** is 1.38.

Each layer of the foregoing design is a quarterwave or a multiple thereof. The spectral transmittance of the "perfect" coating appears as the *shaded* curve in Fig. 9-78. Listed in Table 9-8 and graphed in Fig. 9-79 is T_{oa} as the entire 24 layer stack is deposited. Figure 9-80 shows the T_{oa} during the deposition of the first seven layers (and part of the eighth), including the spacer **LL** of halfwave optical thickness.

The deposition of the first **L** layer on the naked substrate reduces the reflectance from 4.3% to 1.3%—the transmittance rises, ever so slightly, to point *A* in Fig. 9-80. Table 9-8 shows that the change in transmittance is from 91.7% (for a glass slab uncoated on both sides) to 94.5% at point *A*. The deposition of the high index **H** layer increases the single surface reflectance 35%. The transmittance drops to point *B* in Fig. 9-80. The addition of another **L** layer cancels most of the gains in reflectance of the previous **H** layer—at point *C* the transmittance rises to 88%.

The oscillation of the transmittance between high and low continues as more **L H** pairs are deposited. At point *F* in Fig. 9-80 six quarterwave layers have been deposited. The first halfwave thickness spacer **LL** is now deposited. Because it is absentee (see §2.6.3.1), the transmittance returns to its initial value (at point *G* in Fig. 9-80) at the completion of its deposition. The transmittance rises as the next **H** layer—the eighth to be deposited—is laid down. Figure 9-79 depicts the overall transmittance during the entire deposition.

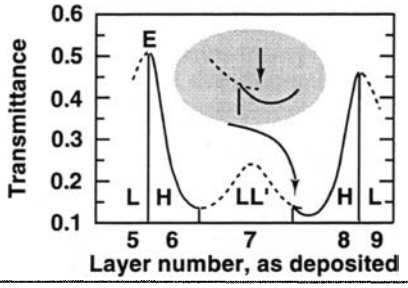


Fig. 9-81 The caption to Fig. 9-80 obtains, with the exception that layers 6, 7 and 8 are deposited and the design is

vacuum 1.308**H LL'** $(\mathbf{H\,L})^3$ *glass*

where the optical thickness of **LL'** is 1.8 λ_0/4. The arrow shows the termination of the transmittance of layer 7 that *would* occur *if* its optical thickness were a full halfwave.

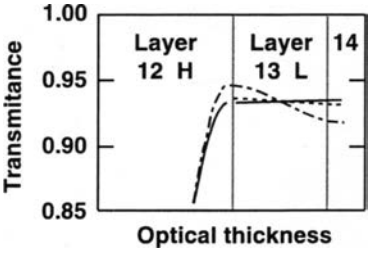


Fig. 9-82 (*Long-short dash*) Total overall transmittance at λ_0 as layers 12 and 13 are deposited of the bandpass in Table 9-9 with indices of 1.38 and 2.22 for **L** and **H**, respectively. (*Solid*) The indices of each of the previously deposited **H** layers drift upwards from layer to layer (see Table 9-9). (*Short dash*) The indices of the **H** layers are light shaded in Fig. 9-83.

9.6.6.2.1 Compensation of previous thickness errors

Error compensation is illustrated by making a deliberate error so that the optical thickness of the spacer layer is 1.8 QWOTs, rather than the required 2.0 QWOTs. Following Bousquet [72], compensation for this 10% thickness error is achieved by requiring that the **H** layer that follows the spacer is deposited so that the transmittance *first* achieves a minimum and then rises to a maximum, as shown in Fig. 9-81. The net result is that the **H** layer is deposited thicker to compensate for the thickness error in its wayward neighbor. The resulting design is

air **H L HL H 2L H L HL HL HL HL HL** 1.308**H** 1.8**L** **H L HL H L** *glass* .

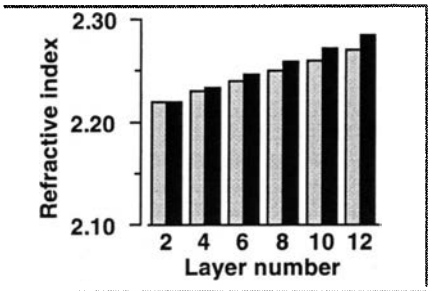


Fig. 9-83 *(Black)* Refractive indices of the high index layers in Table 9-9, showing the upward drift. *(Shaded)* The drift of refractive index is smaller.

Its $T(\lambda/\lambda_0)$ is shown in Fig. 9-78 as the *long-short dash*—it is virtually indistinguishable from the "perfect" coating (*shaded* in Fig. 9-78). For comparison, the *solid* curve in the same Fig. shows the transmittance with a 1% (not a 10%!) thickness error in that spacer—but with no compensation. The distortion of the passband shape is quite evident.

9.6.6.2.2 Occurrence of a null point

The *null point* occurs when a drift in the refractive index occurs. The net effect is that there is a minuscule change in transmittance when a particular layer is deposited. A consequence is that it is practically impossible to control the thickness of that layer by optical means. This is best illustrated by an example.

Assume that the first twelve layers of the bandpass in §9.6.6.2 have been deposited

air **H L H L H LL H L H L H L** *glass* .

The eleven layers next to *air* constitute a symmetrical stack with a halfwave in its center. Using the arguments of §2.6.3.1, at λ_0 those layers are *absentee* and therefore the reflectance of the stack at λ_0 is the same as

air **L** *glass* ,

namely, a single magnesium fluoride layer (which functions as an antireflection coating) on a glass substrate. The addition of the next layer (i.e., layer 13), which is also an **L** layer, converts the single quarterwave layer into a halfwave thickness layer, which is absentee. As shown in Fig. 9-82, the transmittance decreases from its initial value of 94.5% to that of an uncoated slab—91.7%. Although a 3% decrease in optical monitoring signal is not stupendous, it is adequate.

Table 9-9
Transmittance during optical monitoring of a bandpass

Layer*	Symbol	Index	T_{oa1}	T_{oa2}	Layer*	Symbol	Index	T_{oa1}	T_{oa2}
13	**L**	1.380	0.9162	0.9336	6	$\mathbf{H_3}$	2.246	0.1348	0.1305
12	$\mathbf{H_6}$	2.285	0.9449	0.9315	5	**L**	1.380	0.5082	0.5041
11	**L**	1.380	0.6186	0.6356	4	$\mathbf{H_2}$	2.233	0.3118	0.3088
10	$\mathbf{H_5}$	2.272	0.8414	0.8546	3	**L**	1.380	0.8414	0.8414
9	**L**	1.380	0.3118	0.3117	2	$\mathbf{H_1}$	2.220	0.6186	0.6186
8	$\mathbf{H_4}$	2.259	0.5082	0.5082	1	**L**	1.380	0.9449	0.9449
7	**LL**	1.380	0.1348	0.1305	0	*Glass*	1.520	0.9162	0.9162

Total overall transmittances at a wavelength λ_0 for a slab of refractive index 1.52 uncoated on one side. Its opposite side is coated with the nonabsorbing layers listed in the second and third columns. The incident medium is *vacuum*. All layers are $\lambda_0/4$ in optical thickness, with the exception of layer 7, which is $\lambda_0/2$ in thickness. T_{oa1} is the transmittance at the completion of a layer and is computed with a refractive index of 1.38

for L and 2.22 for all H layers. The indices for the T_{oa2} calculation appear in the *Index* column.

*Numbered in the order in which layers are deposited.

Figure 9-83 also shows that the aforementioned 3% change in signal may virtually disappear due to drift in the refractive index of the coating materials. This is illustrated by an example in which there is an upward drift of the refractive index of the H layers. As depicted in Fig. 9-83, it drifts upwards a mere 3% — from 2.22 to 2.285. This index is also listed in Table 9-9.

The strange behavior of the transmittance (i.e., the monitoring signal) in Fig. 9-82 may be explained in terms of the normalized admittance Y of the twelve-layer stack, which consists of 13 quarterwave layers. The design of the partially completed bandpass is

vacuum $\mathrm{H_6\,L\,H_5\,L\,H_4\,LL\,H_3\,L\,H_2\,L\,H_1\,L}$ *glass* .

Equation (2-48c) may be used to find the admittance of the foregoing design because *reduced* admittance and admittance are identical when the incident medium is vacuum. Table 9-9 identifies the symbols for the layers. Equation (2-48c) is factored into three parts. The first consists of the five layers deposited after the spacer. The second is the spacer itself consisting of two quarterwave layers of L—since n_L appears in both numerator and denominator, it is identically 1.00. The third group represents those six layers adjacent to the substrate—layers deposited prior to the spacer. The admittance is pure real and is

$$Y = \frac{(n_{H6}\, n_{H5}\, n_{H4}\,)^2}{(n_L\, n_L)^2\ n_s}\ \frac{n_L^2}{n_L^2}\ \frac{(n_L\, n_L\, n_L)^2}{(n_{H3}\, n_{H2}\, n_{H1}\,)^2} = \frac{n_L^2\,(n_{H6}\, n_{H5}\, n_{H4}\,)^2}{(n_{H3}\, n_{H2}\, n_{H1}\,)^2\, n_s} \ . \qquad (9\text{-}19a)$$

Lest there be confusion, the numbering of the layers in Eq. (9-19a) is the order in which they are *deposited* and conforms to Table 9-9. If there were no drift in the index, then $n_{H1} = n_{H2} = n_{H3} = n_{H4} = n_{H5} = n_{H6}$ and the normalized admittance Y_a is

$$Y_a = \frac{n_L^2}{n_s} = \frac{(1.38)^2}{1.52} = 1.253 \ . \qquad (9\text{-}19b)$$

Y_a may be viewed as the pseudo-substrate for layer number 13, which is next to be deposited.

Consider next the situation where the index drifts, as depicted in Fig. 9-83. The indices from Table 9-9 are substituted into Eq. (9-19a). The normalized admittance Y_b is

$$Y_b = \frac{(2.285\times 2.272\times 2.259)^2\ (1.38)^2}{(2.246\times 2.233\times 2.220)^2\ \ 1.52} = \frac{(11.73)^2}{(11.13)^2}\ \frac{(1.38)^2}{1.52} = 1.392 \qquad (9\text{-}19c)$$

and is exceedingly close to the 1.38 index of the layer that is to be deposited on top of it.

But what would happen if the drift were a bit smaller, as shown by the shaded bars in Fig. 9-83? In this instance, rather than 1.392, suppose the right of Eq. (9-19c) produced a normalized admittance Y_b of precisely 1.38. The deposition of a layer of index 1.38 on a substrate of the *same* index produces *no* change in reflectance or transmittance. Thus a *null point* has been reached in the optical monitoring. The signal is independent of the thickness of the layer. In the language of §2.6.3.2, the L layer is a *buffer layer.* In this case, however, the buffer layer is entirely unwanted.

The foregoing arguments explain the possible behavior of the monitoring signal for layer 13 (see Fig. 9-82): (1) It decreases approximately 3% when there is no drift in

the refractive index; (2) it *decreases* by 0.2% (rather than 3%) when there is a smaller upward drift of the index; and (3) it *increases* a few tenths of a percent when there is a slightly higher upward drift. Optical inhomogeneities in any of the layers can also produce effects similar to the index drift. The only recourse is to deposit layer 13 with a crystal monitor or to use a special optical monitor with a monitor chip of refractive index different than 1.52.

9.6.6.2.3 Optical monitoring—thin layers next to the substrate

§7.6.5.6.2 observes that equivalent layers are sometimes used at the beginning and at the end of the bandpass. They contain layers whose thicknesses are not quarterwaves and in fact they are often much thinner than a quarterwave. The purpose of this section is to show that it is often desirable to alter the multilayer design so that the thickness control of these thin layers is improved when optical monitoring is used.

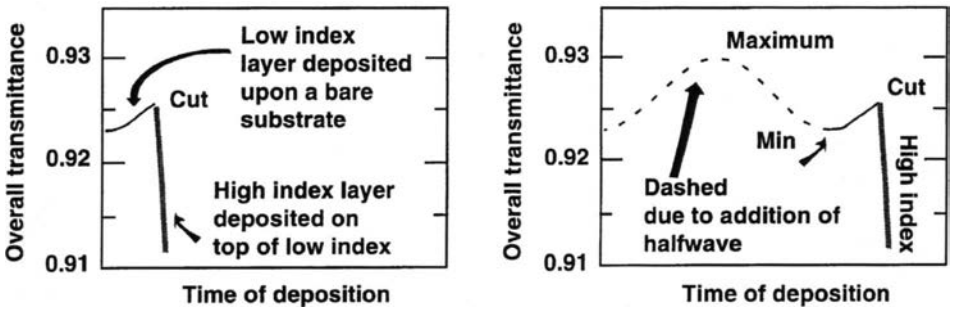


Fig. 9-84 (*Left panel*) Versus time, at a constant deposition rate and at a wavelength λ_0, overall transmittance of a glass plate of refractive index 1.50 as the layer 0.41**L** is deposited upon it (*solid*). An **H** layer is then deposited on top of 0.41**L** (*shaded* line). (*Right panel*) The layer 2.41**L** is deposited upon the glass (*solid* and *dashed*). An **H** layer is deposited on top of 2.41**L** (*shaded* line). The optical thickness of **H** and **L** is $\lambda_0/4$.

Consider, for example, the three cavity bandpass captioned in §7.10.3.2.1. At λ_0, the sequence 0.41**L** 0.17**H** 0.41**L** is an equivalent layer that manifests a equivalent index of 1.61 and 90° phase thickness. Yet inspection of §7.10.3.2.1 shows that the design 0.41**L** 0.17**H** 2.41**L** is used. In other words, a halfwave (i.e., two quarterwaves) is added to the layer next to the substrate. Why?

The answer is that the optical monitoring is facilitated by the addition of this halfwave. Consider the transmittance of the optical monitor as the first layer is deposited. Since the refractive index of that layer is less than that of the substrate, it functions as an antireflection coating as it is deposited. Its effect is to increase the transmittance. The problem is that the increase in transmittance is quite small. The left panel of Fig. 9-84 shows that the deposition of that 0.41**L** is to be terminated after the transmittance has increased by only a few tenths of a percent. Since the 100% transmittance level (i.e., with no glass in the system) is rarely known, it is difficult to convert the gain of the amplifier in the optical monitor system to an absolute percentage.

The solution is to add a halfwave optical thickness to the layer next to the substrate. The equivalent layer still functions with the addition of this halfwave. The dashed curve in left panel of Fig. 9-84 shows that, after start of the deposition, the

transmittance rises until it attains a maximum of 93% before it decreases again to a minimum, which is the same transmittance as the uncoated glass substrate. The *solid* portion of the curve is identical to the solid curve in the left panel of Fig. 9-84. The deposition of that layer is to be terminated after the second increase in transmittance after the minimum. The value of 93% as the maximum assists in the calibration of the optical monitor.

The right panel of Fig. 9-84 illustrates that the transmittance plunges after the high index is deposited on top of the low index layer — the 2.41**L** sequence.

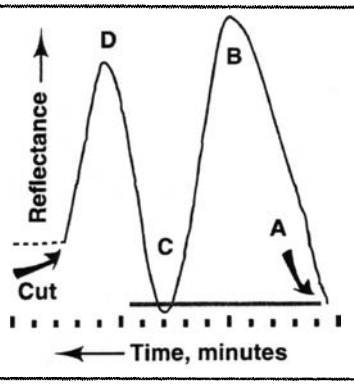


Fig. 9-85 Reflectance (arbitrary units) vs time as a layer of yttria was deposited upon a BK-7 glass substrate by electron beam evaporation. The horizontal line is the reflectance of the bare substrate. "cut" is the termination of the deposition. Courtesy of Coherent [92].

9.6.7 The optical monitor as a process control instrument

Previous parts of §9.6 have focused on the use of an optical monitor to control the thicknesses of layers — in real time — as the coating is deposited. The optical monitor is often useful in process control. An example is that the reflectance changes (as, for example, in Fig. 9-74) of the optical monitor provide information about the refractive index actually attained. The plot in Fig. 9-74 is, however, based upon the assumption that the layer is optically inhomogeneous.

The output (vs time) of a genuine optical monitor in Fig. 9-85 illustrates that the assumption of such homogeneity may be at variance with reality [Coherent 92]. Fig. 9-85 graphs the measured reflectance (at a single wavelength in the visible spectrum) as a single layer of yttria was deposited upon a bare BK-7 glass substrate. Because the yttria index is higher than the glass index, the reflectance rises from that of the uncoated substrate (at *A*) to the first maximum (at *B*), where its optical thickness is a quarterwave. The reflectance then decreases and at the minimum (at *C*) the optical thickness is a halfwave. The R_{min} that is below the uncoated substrate is clear evidence of a negative optical inhomogeneity (see Figs. 3-37 and 3-39). The reflectance then rises to a maximum at *D*, where the optical thickness is three quarterwaves. The fact that peak reflectance at *D* is below the reflectance at *B* is evidence of either optical absorption or a substantial amount of negative optical inhomogeneity. Indeed, if the layer were optically homogeneous and nonabsorbing, the reflectances at the quarterwave and three quarterwave thicknesses should be equal, as is illustrated in Fig. 2-41).

9.7 Mechanical stress in optical coatings

9.7.1 Overview

Mechanical stress causes the following problems in optical coatings:

(1) Cracking or delamination of the coating from the substrate, as shown in Fig. 3-44;

(2) Deformation of the substrate; and (3) stress birefringence. The last is particularly a problem in laser gyro mirrors. The author has witnessed the stress in a coating deform a 25 mm ϕ half millimeter thick substrate at least a few μm.

Stress in thin films is a vast subject, as is evident from the review articles of Hoffman [66] and Campbell [70]. In particular, Ennos [66], Clapham [68] and Kurdock [74] discuss the measurement of stress and methods of controlling it. §9.7.2 reviews stress measurement and §9.7.3 discusses ways to mitigate it.

9.7.2 Measurement of the mechanical stress

9.7.2.1 Introduction

Stress is rarely measured directly. Rather, strain is measured and stress is inferred from the mechanical properties of the material. An example is that strain in a crystalline film may be measured by X-ray diffractometry.

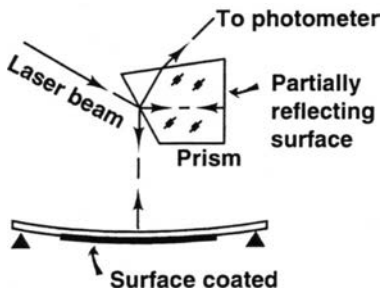


Fig. 9-86 The compressive mechanical stress of a coating on the bottom side of a lamina bends it. A laser beam measures interferometrically the physical displacement of the lamina.

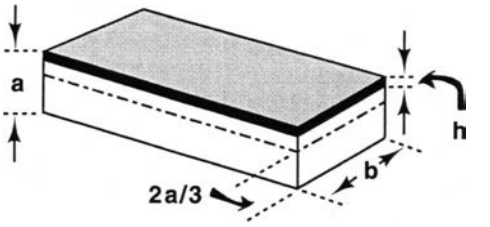


Fig. 9-87 A film (*shaded*) of metric thickness h and width b is deposited on a lamina of rectangular cross section. Its thickness is a. The bending moment (due to stress in the film) is computed about the neutral axis, which is identified by the *long-short* dash.

9.7.2.2 Bending of a rectangular lamina This stress measurement method is included for pedagogical reasons. It is preferable to measure stress with a disk (see §9.7.2.3). Figure 9-86 shows the thin lamina — the *beam*. It is held at each end such that it is statically determinate. An interferometer measures the deflection of the beam as a thin film deposits on its bottom side. The coating stress causes it to bend. Its curvature C is

$$C = \frac{M}{Y I} \tag{9-20a}$$

where M is the bending moment (units of N m), Y is Young's modulus (units of Pa) and I is its moment of inertia (units of m^4) about its neutral axis, which is one-third of the distance from the coated face. Figure 9-87 shows the dimensions of the beam.
The bending moment M is

$$M = (\text{force})\ (\text{distance from neutral axis}) = (S\ b\ h)\ a/3 \tag{9-20b}$$

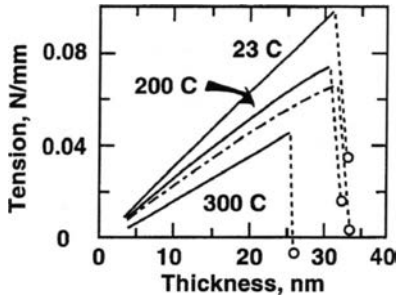


Fig. 9-88 Tension vs metric thickness during deposition of MgF_2 at approximately 10 nm/s at several substrate temperatures. The *long-short* dash curve is 235 °C. The *dashed* line and open circle is the change when air is admitted to the chamber. After Clapham [68].

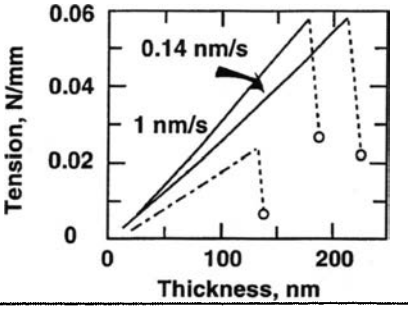


Fig. 9-89 Tension vs metric thickness during deposition of MgF_2 at two rates. The *long-short* dash curve is for ceric oxide. The *dashed* line and open circle is the change when air is admitted to the chamber. After Clapham [68].

where S is the average stress. The product bh in the foregoing equation is the cross-sectional area of the coating. The important parameter in Eq. (9-20b) is the *tension*

$$\text{tension} \equiv S\,h. \tag{9-20c}$$

This parameter increases linearly with the metric thickness h of the coating, provided the mean stress remains constant. It is the *tension* that is measured with the beam deflection as the film is deposited on it. Published stress data often vary considerably. Quite often this can be attributed to the uncertainty in measuring the thickness h in Eq. (9-20c).

As an example, Fig. 9-88 depicts the tension vs time (i.e., thickness, assuming a constant deposition rate) as magnesium fluoride is deposited at several temperatures. The stress may be computed from the slope of the curve. Figure 9-89 shows the rate dependence of the tension and also the stress of ceric oxide. The stress is reduced considerably when ambient air (with its water vapor) is re-admitted to the coating tank.

Although Eqs. (9-20a) and (9-20b) are not developed further, it should be recognized that there are inherent errors in the model: (1) It is assumed that the layer itself has no stiffness; (2) the fact that the beam is allowed to bend relieves a portion of the stress and thereby introduces an inaccuracy; and (3) Poisson's ratio is neglected.

9.7.2.3 Stress from the deflection of a circular disk

Campbell [70] reviews the literature on this measurement method. Figure 9-90 shows that the deflection of a thin disk is measured interferometrically as a layer is deposited on its bottom. A 0.5 mm thick fused silica disk is widely used. The tension $S\,h$ for a given deflection δ is

$$S\,h = \frac{4\ Y\ a^2\ \delta}{3\,(1-\nu)\,D^2} \tag{9-21}$$

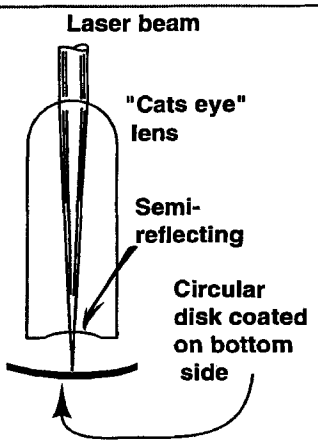


Fig. 9-90 A circular disk is deformed by the mechanical stress of a film deposited upon its lower surface. The laser beam records interferometrically its mechanical displacement.

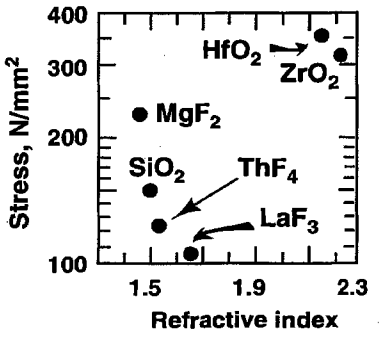


Fig. 9-91 Mechanical stress and refractive index of coating materials deposited at substrate temperatures in the range 200 °C to 250 °C. Stress data from Ennos [66].

where *a, D, Y* and ν are the thickness, diameter, Young's modulus and Poisson's ratio of the disk, respectively. The advantage of the "cat's eye" interferometer in Fig. 9-90 is its relative insensitivity to vibration. The fringe intensity is often displayed on a chart recorder. There is, fortunately, an asymmetry in the fringe shape that indicates whether the disk is deflected upward or downward — and hence whether the net stress is tensive or compressive. Jacobson [94] shows the installation of this type of interferometer in a vacuum tank.

9.7.3 Control of the mechanical stress

Equation (9-21) is written with the tension — rather than stress — on its left because it is this parameter that increases with coating thickness. The tension in turn produces a shear stress between the bottom layer and the substrate. The coating fails when this stress exceeds the adhesion.

The literature often uses *stress* and *tension* interchangeably. It is the tension — not stress — that increases with film thickness. The net tension can be balanced by several methods: (1) modification of the process; (2) tension compensation in layers; and, (3) mixtures of materials.

9.7.3.1 Modification of the process

Stress may be either increased or decreased by varying parameters such as, film composition, rate of deposition, residual gases in the chamber and substrate temperature. In addition, ion bombardment during deposition and post-deposition baking of the coatings can alter the stress dramatically.

9.7.3.2 Tension compensation in discrete layers

Figures 3-45 and 9-91 depict the stress vs refractive index for films that were vacuum deposited from thermal or E-beam sources. All stresses are tensive with the exception of ZnS, ZnSe, SiO_2, La_2O_3 and Y_2O_3. Kurdock [78] reported a compressive stress in MgO, although its affinity for water vapor does not recommend using thick layers of it as a coating material. Virtually all sputtered materials are compressively stressed.

It may be possible to compensate the tension. An example is Fig. 9-92 in which there are five alternate layers of compressively and tensively stressed materials, as identified by the shading. The tension [see the left of Eq. (9-21)] is the product of the stress and the metric thickness. Zero tension for these five layers requires that

$$|S_a|\, h_1 - |S_b|\, h_2 + |S_a|\, h_3 - |S_b|\, h_4 + |S_a|\, h_5 = 0 \qquad (9\text{-}22)$$

where S_a and S_b are the respective stresses in the two materials. Such tension compensation can occur in a SiO_2-TiO_2 multilayer by alternating tensively stressed high-index titania layers with compressively stressed low-index silica layers. A MgF_2-CeO_2 multilayer, on the other hand, offers no tension compensation because both materials are tensively stressed.

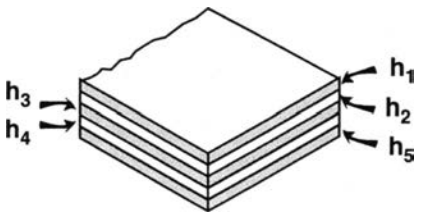


Fig. 9-92 Cross section of a multilayer containing layers that are alternately compressively and tensively stressed.

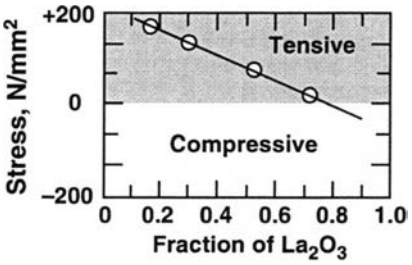


Fig. 9-93 Stress in a co-evaporated PrF_3-La_2O_3 thick film as a function of the percentage composition of La_2O_3. After Kurdock [74].

Another contribution to the stress is the differential expansion coefficient of the coating materials and substrate due to temperature. An example is a fused silica substrate coated at 300 °C. If the mean thermal expansion of the coatings exceeds that of the silica, a net tensive stress occurs when the coated substrate cools to room temperature.

9.7.3.3 Mixtures of materials

Stress compensation is achieved by co-evaporating or co-sputtering compressively and tensively stressed materials. This is illustrated in Fig. 9-93, which shows the stress of a PrF_3-La_2O_3 mixture vs composition [Kurdock 74]. Kurdock [78] describes other stress-free mixtures.

9.8 Appendices

9.8.1 Planetary motion — equations

As depicted in Fig. 9-94, on the planet there is the point P representing an area of infinitesimal size that is being coated. The xy coordinate system (*long-short* dashed in Fig. 9-94) is attached to the "chamber." Its origin is at the center of the sun gear and it does not move. The spindle gear rotation center is at a radial distance r_{sp} from the center of the sun gear. Figure 9-94 shows that spindle gear center has revolved by an angle ξ. At a given instance of time, the coordinates of the spindle gear center in the "chamber" coordinate system are x_0 and y_0.

The Cartesian coordinate system x' and y' (short dash in Fig. 9-94) is locked into the spindle gear center and rotates, as the spindle gear center rotates. The x' axis of this rotating system is always co-incident with the line r_{sp} mentioned above.

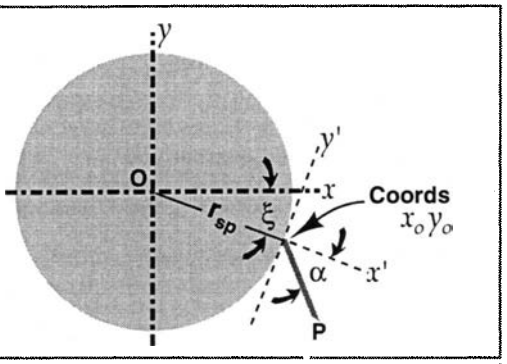


Fig. 9-94 In planetary rotation, r_{sp} is the radial distance from the center to the spindle gear and the shaded area is the locus of its center as it rotates about the sun gear. Please consult the text for further details.

The point P on the planet is at an angle α with respect to the x' axis. This angle is produced by the rotation of the spindle gear about its own axis. P is at a radial distance from the spindle gear center as shown by the heavy shaded line in Fig. 9-94. The Cartesian coordinates (in the rotating system) of P are x' and y'. A coordinate rotation by the angle "negative ξ" transforms the Cartesian coordinates (in the rotating frame of the spindle gear) x' and y' to the "chamber" coordinate system x, y:

$$x = x_0 + x' \cos(-\xi) - y' \sin(-\xi) \tag{9-25a}$$

$$y = y_0 + x' \sin(-\xi) + y' \cos(-\xi) \tag{9-25b}$$

Given the rotation angle ξ of the spindle center about the sun gear, the rotation angle α of the spindle gear is

$$\frac{\alpha}{\xi} = \frac{\text{number of teeth in the sun gear}}{\text{number of teeth in the spindle gear}} \tag{9-25c}$$

9.8.2 Design of an edge filter

The design is (for the *dashed* curve):

air 128.97**L** 65.96**H** 46.18**L** 48.70**H** 77.59**L** 48.70**H** 77.59**L** 48.70**H** 77.59**L** 48.70**H** 77.59**L** 48.70**H** 77.59**L** 48.70**H** 77.59**L** 48.70**H** 77.59**L** 37.73**H** 82.82**L** 35.56**H** 26.49**L** 26.18**H** 35.17**L** 24.32**H** *glass*

where the refractive indices of *glass*, **L** and **H** are 1.52, 1.45 and 2.31, respectively. For the *solid* curve, 2.22 is used for the refractive index of the **H** layer.

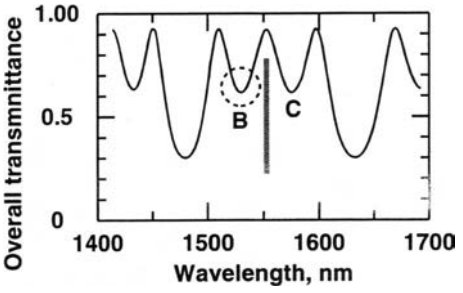


Fig. 9-95 Overall transmittance of

air **H** 20**L** 20**H** 20**L** **H** *sub*

where optical thickness of **H** and **L** is $\lambda_0/4$ at λ_0 of 1552.5 nm. The refractive indices of *sub*, **H** and **L** are 1.50, 2.065 and 1.47, respectively. The vertical shaded line is at 1552.5 nm.

9.8.3 Thickness uniformity of a WDM bandpass — introduction

The hardware configuration of the coating chamber is the following: A coating is deposited upon a plano substrate that is rotated about its center. Two materials are alternately deposited from evaporative or sputter sources. The objective is to use masking (see Fig. 9-43) to obtain the largest area on the substrate over which the optical thickness is essentially identical for the two materials.

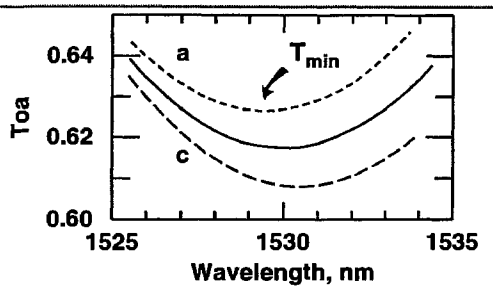


Fig. 9-96 Overall transmittance of the (*dashed*) circled spectral region in Fig. 9-95 (*short dash, upper curve a*)
air 1.001**H** 19.98**L** 20.02**H** 19.98**L** 1.001**H** *sub,* (*solid, middle curve*)
air **H** 20L 20**H** 20L **H** *sub*
and (*long dash, lower curve c*)
air 0.999**H** 20.02**L** 19.98**H** 20.02**L** 0.999**H** *sub.*
The design symbols are defined in the caption to Fig. 9-95.

The first step is to deposit the coating with no masks. The question is: how can the optical thickness be determined at various radial positions across the substrate?

The answer to the foregoing question is: Deposit a modification of the Grèzes-Besset stack as described in §9.6.4.4.1. This is identified as the *test coating* in the discussion below. From spectrophotometer transmittance data, the optical thickness ratio may be determined.

9.8.3.1 Test coating

The following test coating is recommended to determine the relative coating thickness:

air **H** 20**L** 20**H** 20**L** **H** *glass*

This Fabry-Perot (with a composite spacer layer) consists of two reflectors, each consisting of a single quarterwave optical thickness high index layer **H**. Each of the three layers that constitute the composite spacer has an optical thickness of five full waves, which is equivalent to twenty quarterwaves.

Let us suppose that the optical monitoring is centered in the spectral region near a wavelength of 1552.5 nm. The test coating is deposited at this central wavelength and its overall transmittance appears in Fig. 9-95. As defined in §1.2.2.2, the *overall transmittance* accounts for the back surface transmission loss of the substrate.

Our focus is on the transmittance minima designated as *B* and *C* in Fig. 9-95. Figure 9-96 plots on an expanded scale the transmittances of the minimum designated as *B* and circled in Fig. 9-95. The following adjustments to the relative thicknesses of the layers is made:

(a) The **L** layers are thinner by a factor of 0.999 and the **H** layers are thicker by 1.001.
(b) All layers have their original thickness.
(c) The **L** layers are thicker by a factor of 1.001 and the **H** layers are thinner 0.999.

Similar to the performance in Fig. 9-72, Fig. 9-96 shows that the transmittance minima (arrowed as T_{min}) progress downwards for these three conditions, from curve *a* to curve *c*. Although it is not graphed or tabulated, the transmittance minima at C in Fig. 9-95 progress upwards (as shown in Fig. 9-72) for the foregoing three conditions.

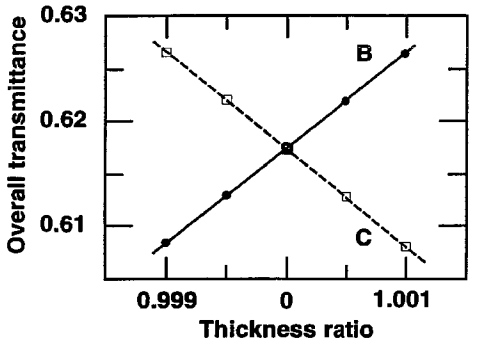


Fig. 9-97 Versus thickness ratio [see Eq. (9-18b)], overall transmittance of the minima in the spectral regions B and C in Fig. 9-95 for the coating appearing in the caption to Fig. 9-95.

Figure 9-97 graphs the numerical values of the minima of the transmission curves vs the thickness ratio, as defined in Eq. (9-18b). This plot is similar to Fig. 9-73. This shows that if the transmittance can be measured to a precision of 0.5 percent, the optical thickness ratio of the high index component and the low index component may be determined to a precision of one part in 500.

In determining the transmittance values of these minima, the digital data from the recording instrument should be analyzed numerically. For example, the 15 or so points adjacent to a minimum should be fit with a least squares parabola and the minimum of the curve determined from the minimum of the parabola.

In some instances, the aforementioned design of

air **H** 20**L** 20**H** 20**L H** *glass*

may be too sensitive and the process will be expedited if thinner layers are used in the spacer, such as the design

air **H** 10**L** 10**H** 10**L H** *glass.*

The problem is that the wavelength of the B and C minima in Fig. 9-95 will move farther away from the central wavelength when the spacer layers are reduced in thickness.

9.8.4 Composite optical monitoring—termination at a maximum or minimum

Assume that a bandpass consists of layers of quarterwave optical thickness (or multiples thereof). The problem is to terminate the deposition of a layer when the transmittance is a maximum or a minimum. That is, when the derivative of transmittance with respect to time is zero. Assume that the transmittance T is measured at equal time internals Δt. Suppose seven data points have been collected, numbered from zero to six. These transmittance values are T_0, T_1, T_2, T_3, T_4, T_5 and T_6 . The finite difference approximation for the time derivative of the transmittance T' at point six (i.e., the most recently measured value) is

$$T' = \frac{10\,T_0 - 72\,T_1 + 225\,T_2 - 400\,T_3 + 450\,T_4 - 360\,T_5 + 147\,T_6}{60\,\Delta t} . \qquad (9\text{-}26)$$

The foregoing is from [Milne 49]

The aforementioned numerical approximation for the derivative is computed and stored in an array. A least squares parabola is fit to the ten or fifteen most recent derivatives. The intersection of this least squares curve with the time axis (i.e., x axis) specifies the estimate of the time at which the derivative goes to zero.

Chapter 10
Miscellaneous topics

10.0 Overview of chapter

The first two sections of Chap. 10 present concepts useful in the design of optical coatings. §10.1 focuses upon the graphical presentation of the admittance or reflection coefficient of a coating and §10.2 on the use of optimization in coating design. Subsequent sections are concerned with the performance of a coating after it has been fabricated. The concern of §10.3 is the transmittance and reflectance of an array of coatings deposited on separate substrates. §10.4 and §10.5 discuss the characterization of coatings by spectrophotometry and other means. §10.6 recapitulates the phase relationships of a multilayer. §10.7 examines how those phase relationships can predict how a wavefront is altered when it reflects from or is transmitted by a coating.

10.1 Graphical aids to multilayer design

In the two decades from 1940 to 1960, digital computers were scarce and quite expensive. The thin film designer had no choice but to depend heavily on graphical design methods. In the twenty-first century and beyond, the digital computer calculates the admittance or reflection coefficients and can present them graphically. Such data furnish insight as to how a multilayer functions.

10.1.1 Use of graphics

§2.8.1.1 and §2.8.2 illustrate the two uses of graphics: (1) actual design; and (2) insight. The text of Macleod [86] is a sterling example of the use of graphical design. It plots more than seventeen admittance diagrams of coatings containing as many as four layers. Graphical methods may hint which layers are most sensitive to thickness or index errors. A caveat: Only the *reflectance* of a stack (whether absorbing or nonabsorbing) can be found directly using these graphical techniques. Table 10-1 summarizes the characteristics of several exact graphical calculators.

Table 10-1 Characteristics of graphical calculators

Type	x	y	Circle 1	Circle 2	Reference
Admittance diagram	Y_r	Y_j	$\lvert\hat{\rho}\rvert$	δ	Ragan [48], Macleod [86]
Smith	ρ_r	ρ_j	y_r	y_j	Smith [69]
Carter	ρ_r	ρ_j	$\sqrt{y_r^2 + y_j^2}$	$\arctan(y_r/y_j)$	Carter [39]
Kard	ρ_r	ρ_j	$\ln(y_r^2 + y_j^2)$	$\arctan(y_r/y_j)$	Kard [56]
Circle	ρ_r	ρ_j	$\lvert\hat{\rho}\rvert$	Phase tkns of layer	Apfel [72]

The x and y columns list the Cartesian coordinates of graphical calculators. *Circle* 1 and *Circle* 2 are plotted on the chart that represents the calculator. Y_r, Y_j, y_r, y_j, ρ_r, ρ_j and δ represent the real and imaginary components of the normalized admittance, real and imaginary parts of the reduced admittance, real and imaginary parts of the amplitude reflection coefficient $\hat{\rho}$ and phase shift upon reflection, respectively.

The use of graphical methods for actual design is not emphasized in this book. The admittance diagram is discussed by Kraus [49] and Ragan [48]. The Smith chart is the subject of a book [Smith 69]. Berning [63] elaborates upon the Carter chart, which is also known to microwave circuit designers as the Z-θ chart. A useful variant of the Carter chart is the Kard calculator [Kard 56]. The apotheosis of the circle diagram in the article of Apfel [72] is decidedly mathematical. All of the aforementioned graphical methods are exact. The vector addition of amplitudes (see §2.8.3) is a useful approximate method.

10.1.2 Graphically presented parameters

Three parameters that may be presented graphically are: (1) the *internal* amplitude reflection coefficient; (2) the normalized admittance; and (3) the *external* amplitude reflection coefficient. In the first example (in §10.1.2), a three-layer all-dielectric stack (mentioned in §2.4.5.1 and §2.4.5.2) is intended to reduce the reflectance of a metallic substrate at a single wavelength. The second example in §10.1.3 is a stack with absorbing layers.

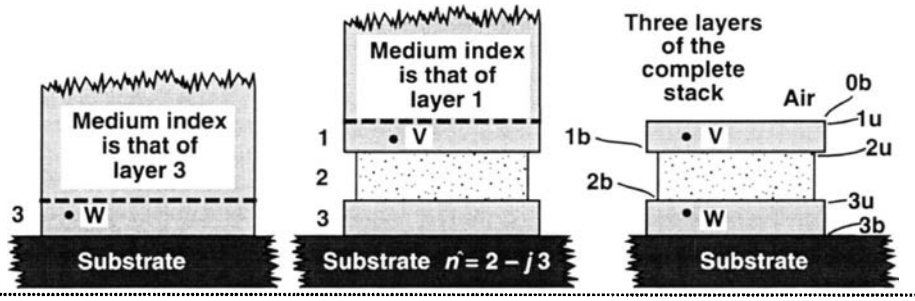


Fig. 10-1 (*Right panel*) The internal amplitude reflection coefficient is measured by moving a generator of light from the bottom of the stack to the top. (*Left panel*) The generator (point *W*) is positioned three-quarters of the layer thickness from the substrate side of the bottom layer. (*Center panel*) The generator (point *V*) is midway in the outermost layer of the stack. The coating design appears in Table 2-9 and the points "0b", "1b", and so on, appear in Table 2-9, and in Figs. 10-2 and 10-3.

10.1.2.1 Internal amplitude reflection coefficient

In this section, $|\hat{\rho}|$ represents the magnitude of the *internal* amplitude reflection coefficient at a prescribed position inside a multilayer and δ represents its phase [see Eqs. (2-16e) and (2-18b)]. Figure 10-1 depicts a "thought experiment" in which the *internal* amplitude reflection coefficient is measured by "burying" the generator of radiation inside the stack. This is termed the *point of observation*. This is then moved towards the incident medium. The process is halted when the top surface of the outermost layer is reached.

A few examples clarify the foregoing process. The dot marked *W* in Fig. 10-1 (*left panel*) shows the point of observation when it is three-quarters of the layer thickness from the substrate in the bottom-most layer. For the purposes of this measurement, the medium (of index 2.05) of the bottom layer *is* the incident medium. The center and outer layers, as well as the incident medium of air, do not exist, as yet. The center panel shows

a point of observation (V dot) that is centered in the outermost layer of the stack. The outer interface of that layer is dashed because it has not yet been added.

Figure 10-2 shows a parametric plot (vs thickness) of the *internal* amplitude reflection coefficient of the stack mentioned above. The symbols "3b", "3u", "2b", and so on, refer to Table 2-9. There is a discontinuity in the reflection coefficient $\hat{\rho}$ at each interface that represents physically the increase in reflection produced by the Fresnel reflection at that interface. This is illustrated by the fact that point "3u" is the *internal* amplitude reflection coefficient at the *top* surface of the layer next to the substrate and "2b" is the same coefficient at the bottom side of the center layer. At point "1u" the $\hat{\rho}$ is measured at the uppermost part of, but within, layer 1. As the interface is crossed at point "0u", the $\hat{\rho}$ (in air) is close to zero, as it should be for a single-wavelength antireflection coating.

The phrase *transversal of a layer* is subsequently used. It means that the generator of radiation is moved from the emergent side of that layer to its incident side.

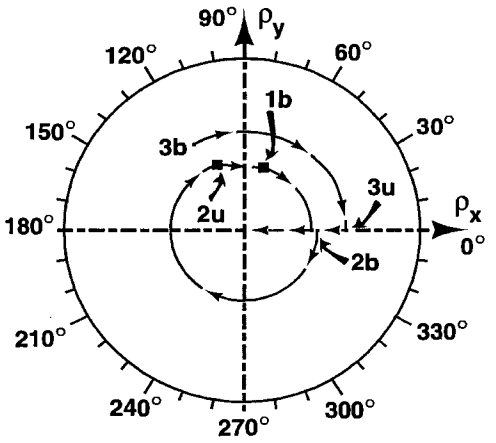


Fig. 10-2 At a wavelength λ_0, internal amplitude reflection coefficient (plotted parametrically vs optical thickness) of the stack captioned in Fig. 10-1. The symbols "3b", "3u", "2b", and so on, refer to the "Nr." column in Table 2-9.

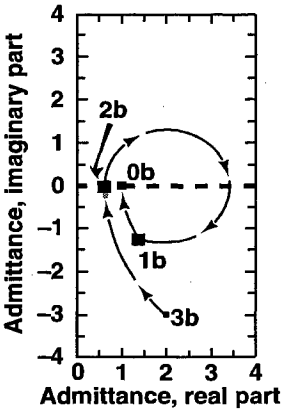


Fig. 10-3 The caption to Fig. 10-2 obtains, with the exception that normalized admittance is plotted.

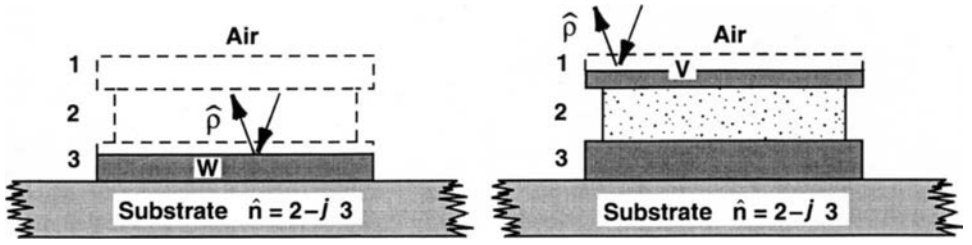


Fig. 10-4 The partial construction of the three-layer coating (captioned in Fig. 10-1) for the purpose of measuring the external reflection coefficient. The letters V and W correspond to the positions in the stack in Fig. 10-1.

10.1.2.2 Normalized admittance

The "thought experiment" mentioned in §10.1.2.1 and illustrated in Fig. 10-1 is still used, with the exception that the normalized admittance $\hat{Y}$ (rather than the *internal* amplitude reflection coefficient) is displayed. Figure 10-3 shows the normalized admittance as the point of observation is moved from the substrate (where the normalized admittance is $\hat{Y} = 2 - j3$) to the top of the stack, where it attains a value of $1.00 + j0$. Because this admittance matches that of air, Eq. (2-16e) shows that this stack produces close to zero reflection at this wavelength.

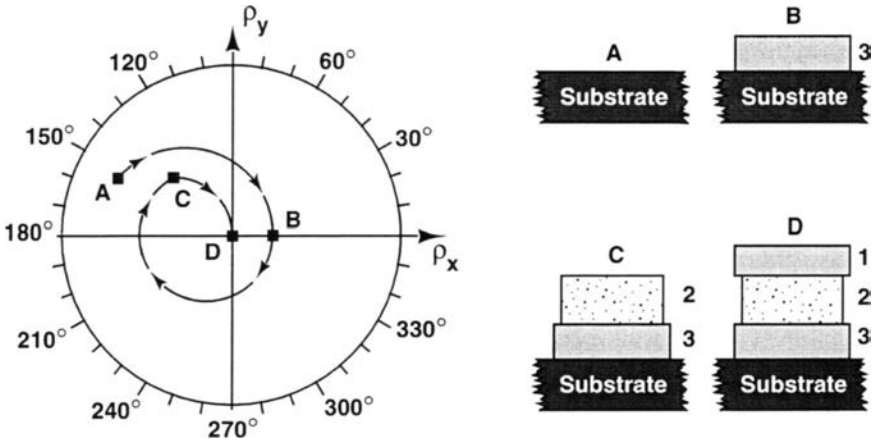


Fig. 10-5 At a wavelength λ_0, external amplitude reflection coefficient (plotted parametrically vs optical thickness) of the stack in Table 2-9. The letters *D*, *C*, *B*, and *A* represent, respectively, the amplitude reflection coefficient (as measured in air) of the uncoated substrate and as the layers in Table 2-9 are successively added to the substrate.

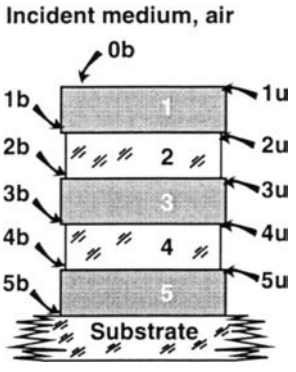


Fig. 10-6 A five-layer stack, the design of which appears in Table 10-2. Shaded layers are absorbing. The symbols such as "1b", "1u", and so on, correspond to the positions in the stack in Table 10-2 and in Figs. 10-7, 10-8 and 10-9.

10.1.2.3 External amplitude reflection coefficient

The admittance $\hat{Y}$ mentioned in §10.1.2.2 is converted into an *external* amplitude reflection coefficient by Eq. (2-16h). This is termed the *external* coefficient because it is measured *externally* to the stack, in the incident medium. Figure 10-4 helps to visualize how the *external* amplitude reflection coefficient is measured for each layer of the coating in Table 2-9.

At point *A*, the amplitude reflection coefficient is that of naked substrate, as measured in air. That substrate is the overcoated with the bottom layer. Point *B* represents the $\hat{\rho}$ of this monolayer, as measured in air. The middle layer is then added to the stack and point *C* represents its $\hat{\rho}$, as measured in the incident medium of air. Point *D* is the $\hat{\rho}$ (which is close to zero) of the completed stack. Apfel [72]

terms the loci of reflection coefficient values traced by this method a *circle diagram.*

10.1.3 Example — an absorbing stack

The next example consists of a stack of five layers **H L H L H** in which the **H** films are slightly absorbing. The optical constant $\hat{n}$ of each high index layer is $\hat{n} = n - jk = 2.30 - j\,0.15$. Figure 10-6 shows the stack. The letters (see "Nr" column in Table 10-2) indicate the various positions in the stack where the reflection coefficient (or admittance) is sampled. The optical thickness of each high index layer is $0.95\ \lambda_0/4$ and the optical thickness of each low index layers is $\lambda_0/4$. Table 10-2 summarizes the reflection coefficient and normalized admittance after each step. The loci of amplitude reflection coefficient or admittance lie on a circle as the *point of observation* is moved through a nonabsorbing layer. When the layer is absorbing, however, those loci lie on an exponential spiral. Given the phase thickness β (in radians) of the layer and the magnitude of the amplitude reflection coefficient $|\hat{\rho}|$ at the bottom interface of the film, the magnitude of the internal amplitude $|\hat{\rho}''|$ at its top interface is

Table 10-2
Reflection coefficients and admittance of a five-layer stack

Nr	Optical thickness	Optical constant	Mod $\hat{\rho}$, internal	Phase of $\hat{\rho}$, internal	Mod $\hat{\rho}$, external	Phase of $\hat{\rho}$, external	Admittance
0b	massive§	$1.00 + j\,0$	0.824	182.9°	0.824	182.9°	$9.67 + j\,2.54$
1u			0.643	188.8°	same	same	same
1	$0.95\,\lambda_0/4$	$2.3 - j\,0.15$					
1b			0.781	360.0°	0.559	1.7°	$0.283 - j\,0.014$
2u			0.660	1.2°	same	same	same
2	$\lambda_0/4$	1.38					
2b			0.660	181.2°	0.741	180.8°	$6.71 + j\,0.326$
3u			0.492	185.0°	same	same	same
3	$0.95\,\lambda_0/4$	$2.3 - j\,0.15$					
3b			0.598	356.0°	0.264	353.6°	$0.584 + j\,0.037$
4u			0.406	356.3°	same	same	same
4	$\lambda_0/4$	1.38					
4b			0.406	176.3°	0.531	177.6°	$3.24 - j\,0.205$
5u			0.171	180.4°	same	same	same
5	$0.95\,\lambda_0/4$	$2.3 - j\,0.15$					
5b			0.208	351.4°	0.206	180.0°	$1.52 + j\,0$
6	massive*	$1.52 + j\,0$					

*Substrate §Incident medium

At a wavelength λ_0, internal reflection coefficient $\hat{\rho}$ (polar form — modulus *Mod* and *phase*) and normalized admittance of a five-layer stack, the design of which appears above. The boldface numerals in *column* 1 refer to the layer, *u* to its upper interface and *b* to its lower interface. The external reflection coefficient $\hat{\rho}$ is measured in air.

$$|\hat{\rho}''| = |\hat{\rho}| \exp(-2\beta\,|k|/n) \ . \qquad (10\text{-}1)$$

The attenuation in the foregoing equation is evident in Fig. 10-7, which depicts the *internal* amplitude reflection coefficient at several points in the stack. Consider the

amplitude reflection coefficient at point "2b", which is at the bottom of the second low index layer — layer 2 in Fig. 10-6. The transversal of this layer and the crossing of the interface into the outermost high index layer produce a reflection coefficient at point "1b" in Fig. 10-7. The transversal of the outermost high index layer *decreases* the amplitude reflection coefficient. This is due to the absorption of this layer and is quantified in Eq. (10-1). In Fig. 10-7, point "1u" lies just beneath the outermost interface of the stack. It is clear in this Fig. that point "1u" lies closer to the origin than point "2b". In other words, the *increase* in the internal amplitude reflection coefficient provided by the crossing of the interface between the high and low index layers is essentially canceled by the absorption in the high index layer.

Figures 10-8 and 10-9 plot, respectively, the *external* amplitude reflection coefficient and normalized admittance of this five-layer stack.

10.1.4 Choice of method

As to the three methods of presenting the data mentioned in §10.1.2, the one that a designer may prefer is purely a matter of personal choice. A feature of the admittance chart is that the admittance may swing from small values to large, as illustrated in Fig. 10-9. This disadvantage is not shared by the charts in Figs. 10-7 and 10-8 that map the amplitude reflection coefficient, although the circles of constant $|\hat{\rho}|$ are compacted together as $|\hat{\rho}|$ approaches 1.00.

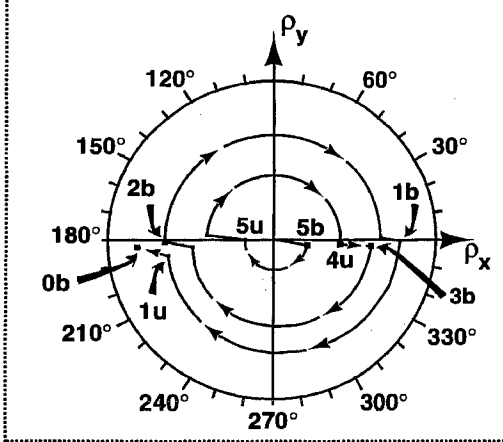


Fig. 10-7 At a wavelength λ_0, internal amplitude reflection coefficient (plotted parametrically vs optical thickness in each layer) of the stack whose design is listed in Table 10-2. The symbols "1b", "2b" and so on correspond to positions in the stack in Table 10-2 and Fig. 10-6.

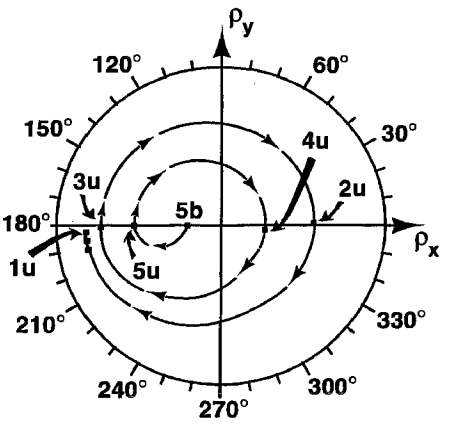


Fig. 10-8 At a wavelength λ_0, external amplitude reflection coefficient (plotted parametrically vs layer thickness) of the stack whose design is listed in Table 10-2. The symbols "1u", "3u" and so on, correspond to positions in the stack in Table 10-2 and Fig. 10-6. Position "5b" represents the uncoated substrate.

An advantage of plotting the *internal* amplitude reflection coefficient is that the physics is not obscured. This is illustrated by the absorbing stack in §10.1.3 — and is shown in Fig. 10-7. The Fresnel reflection increase, i.e., "kick," in reflectance is evident at each interface, as is the attenuation of the internal amplitude reflection coefficient due to absorption. This effect is less evident for the plot of the *external* amplitude reflection

coefficient in Fig. 10-8 and the least evident in the plot of the normalized admittance in Fig. 10-9. One disadvantage of plotting the *internal* amplitude reflection coefficient, however, is that it may exceed 1.00 in an absorbing medium.

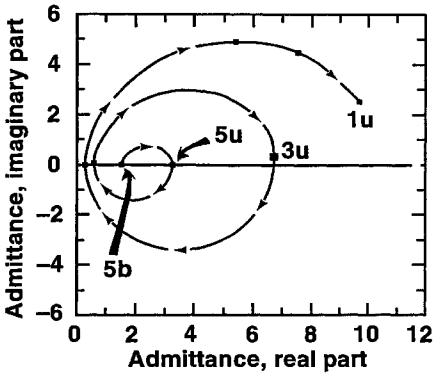


Fig. 10-9 At a wavelength λ_0, normalized admittance (plotted parametrically vs layer 10-2 and Fig. 10-6. Position "5b" represents the uncoated substrate.

10.2 Optimization

10.2.1 Introduction

Optimization, which is synonymous with *refining* or *automatic correction*, is the generic term for an algorithm (usually programmed on a digital computer) that improves the spectral *performance* of a multilayer. The *performance* is usually the transmittance, reflectance, absorption or phase shift upon reflection.

There are at least two possible modes of operation: (1) correction; and (2) minimization. In either mode, the computer code improves the performance by adjusting the *construction parameters* of the multilayer. Examples of construction parameters are: (1) metric thicknesses of the layers; (2) refractive indices of the layers; and (3) constants that establish the dispersion of the layers, such as the *A* and *B* constants in Eq. (2-2c).

10.2.2 Mode of operation — correction

In the correction mode, the designer designates wavelengths in the starting design at which the reflectance or transmittance is to be altered by a prescribed amount. The computer code then adjusts the construction parameters so that this is accomplished.

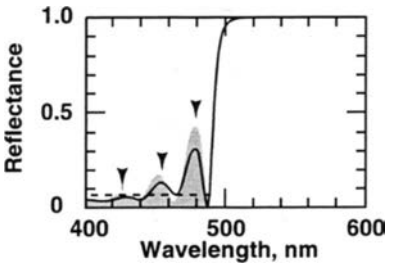


Fig. 10-10 $R(\lambda)$ of shortwave pass filters before correction (*shaded*) and after (*solid*). The arrows designate the maxima that the computer code reduced by correction. Design data are in §10.2.8. The target reflectance is dashed.

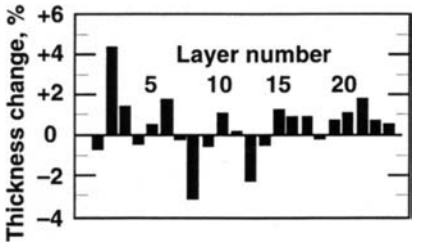


Fig. 10-11 Percentage change in optical thicknesses of the *after correction* coating whose $R(\lambda)$ is shown on the left. Design data appear in §10.2.8.

As an example of the correction mode, suppose that a filter is required to reflect the green and transmit the blue. The reflectance is to lie within the specifications; below the dashed line in Fig. 10-10. The starting design violates these specifications — namely, at the arrowed reflectance peaks at 425 nm, 451 nm and 477 nm. The

reflectance values at those wavelengths are approximately 7%, 17% and 42%, respectively. The execution of a cycle of correction on the computer adjusts those peaks downwards to 0.7 of their original values. The solid curve in Fig. 10-10 shows that the correction cycle pushed the reflectance downwards to 5%, 13% and 30% at those wavelengths. Figure 10-11 shows that the layer thicknesses were altered by only a few percent to achieve this result.

It is advantageous to implement the correction process on computers with slow execution speeds because many fewer arithmetic operations are performed, as compared to minimization (see §10.2.3). An additional advantage is that the layer thicknesses are altered by the least amount.

The process is not without its disadvantages. A latent extremum — say, at wavelength A — may *not* be corrected because it is within the specification. There is also a maximum violation of the specification curve occurring at wavelength B. The computer adjusts the layer thicknesses and causes the extremum at B to come within the specifications. But, in the process, the extremum at A is now in violation of the specifications. An oscillation can easily occur, in which A and B alternately flip back and forth in and out of specifications. This is one of the reasons that the *minimization* technique is more widely used.

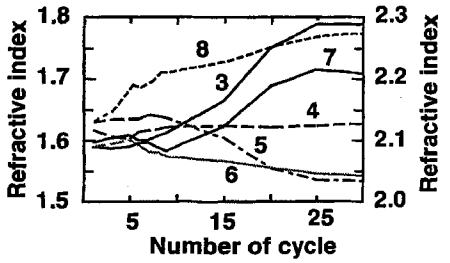


Fig. 10-12 Refractive index of layers during each optimization cycle. *Dashed* lines refer to the left ordinate and the other lines to the right ordinate. The design is discussed in §4.8.5.1. The curves are identified by layer number.

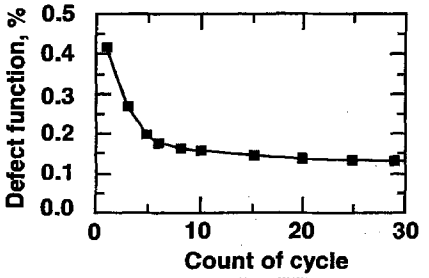


Fig. 10-13 Defect function [see Eq. (10-2a)] vs cycle of optimization of the multilayer captioned in Fig. 10-12 and discussed in §4.8.5.1.

10.2.3 Mode of operation — minimization

The optimization process *minimizes* the *defect* function $\mathcal{D}$, which is called the *merit* function by other authors [Dobrowolski 90]. There are several ways of defining it. For reflectance, the root-mean-square defect is

$$\mathcal{D} \equiv \left[\frac{\sum_{i=1}^{m} W_i \, (R_i - R_{T,i})^2}{\sum_{i=1}^{m} W_i} \right]^{\frac{1}{2}} \qquad (10\text{-}2a)$$

where m is the number of points at which the reflectance is computed and W_i is the relative weight. $R_{T,i}$ is the target reflectance (i.e., the specification) at the ith wavelength

and R_i is the reflectance at that same wavelength. §10.2.5.4 discusses the use of an exponent other than 2 in Eq. (10-2a). The reciprocal of $\mathcal{D}$ is the merit function M

$$M \equiv 1/\mathcal{D} \ . \qquad (10\text{-}2b)$$

The merit function M is to be *maximized* and the defect function $\mathcal{D}$ is *minimized.*

Although the reflectance R is used as the *target* in Eq. (10-2a), many other physical quantities associated with optical coating may be used, including the transmittance T, absorption A, phase shift upon reflection and phase retardance (alias phase shift upon transmission). At nonnormal incidence, the list is extended to the differential phase shift upon reflection Δ, as well as R_p, R_s, T_p, T_s, A_p and A_s. As an example, in an absorbing multilayer it is possible to optimize both the reflectance and transmittance simultaneously. Bloom [81] discusses the software packages that perform the minimization.

10.2.3.1 Example of optimization

The goal is to achieve close to zero reflectance over a 2:1 bandwidth on a BK-7 substrate. Figure 4-44 shows the reflectance of the starting design, in which the antireflection coating of Ikeda [73] is modified by splitting each of the three halfwave layers into two quarterwave layers. Each layer has the same optical thickness. The computer code then adjusted the indices, with the restriction that no index could exceed 2.30 nor be less than 1.38. Figure 10-12 depicts the changes in refractive index and Fig. 10-13 the decrease in the defect function as the optimization progresses. The solid curve in Fig. 4-45 shows the result. R is less than 0.3% over a 2:1 bandwidth.

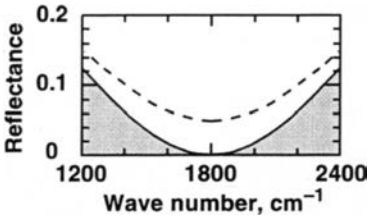


Fig. 10-14 $R(\sigma)$ of a single layer of index 2.00 (*lower curve*) and 2.50 (*upper curve*) deposited upon a substrate of 4.00. The optical thickness of each layer is $\lambda_0/4$ at $\lambda_0 = 5.56$ μm.

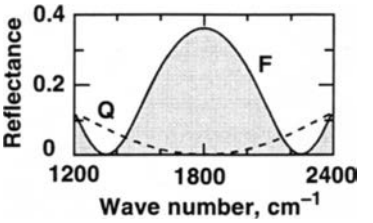


Fig. 10-15 $R(\sigma)$ of a single layer of index 2.00 deposited upon a substrate of 4.00. Optical thickness of the layer is $\lambda_0/4$ for the Q curve (*dashed*) and λ_0 for the F curve (*solid*). λ_0 is 5.56 μm

10.2.4 Design procedure

The following is a generalized design procedure.

(1) Establish the *desideratum.* At various wavelengths the *target* reflectance, transmittance or other parameters are defined and entered into the maw computer. Unless there are compelling reasons to do otherwise, the weights W_i [see Eq. (10-2a)] should be set equal.

(2) Choose a starting design. This is discussed in §10.2.4.1.

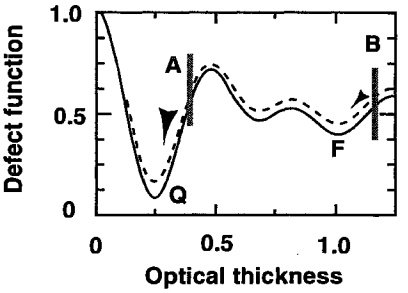


Fig. 10-16 Normalized defect function [see Eq. (10-2c)] vs optical thickness (in waves) of a single layer deposited upon a substrate of 4.00. The film indices are 2.00 (*lower curve*) and 2.50 (*upper curve*). The shaded vertical bars at *A* and *B* are the starting points for optimization (see Fig. 10-17).

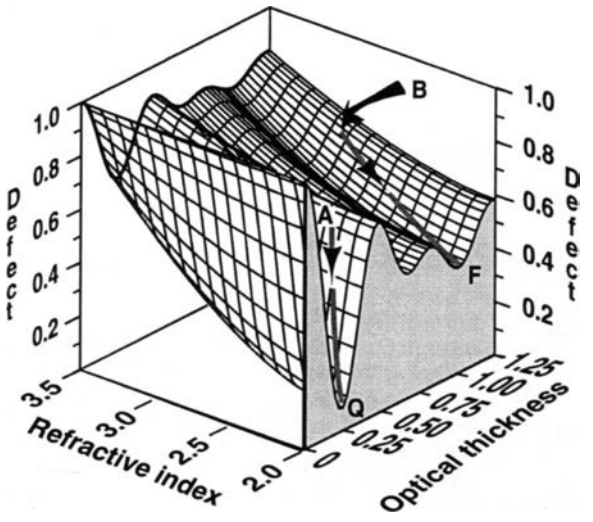


Fig. 10-17 The caption to Fig. 10-16 obtains, with the exception that the variables are optical thickness and refractive index of the layer. Point *F* is a local minimum and point *Q* is a global minimum.

(3) Unleash the awesome might of the magnificent optimization program. The computer gobbles the input design, ruminates and voilà, a final design is spewed forth.

(4) If that final design is acceptable, the laurels of victory rest upon the brow of the designer, who can retire from the field of battle, hors de combat.

(5) Adjustments are needed if the foregoing step fails to occur. This is discussed below.

10.2.4.1 Starting design — general

Multilayer design is an art, rather than a science. Only in rare instances can the cunningness and perspicacity of a computer code compensate for poor judgment in the selection of the starting design. Guidelines are: (1) most optimization procedures seek a *local* minimum of the defect function; and (2) the starting design must contain a

sufficient number of layers "to do the job." Examples, as in §4.1.3, illustrate these points.

10.2.4.2 Achievement of a local minimum

An example of the first guideline in §10.2.4.1 is an antireflection consisting of a single layer of index n_1 and optical thickness $n_1 h_1$ deposited on a germanium substrate ($n_s = 4.00$). Figure 10-14 illustrates that its spectral region of low reflectance extends from a wave number of 1200 cm^{-1} to 2400 cm^{-1}. λ_0 is the reciprocal of the bandwidth center of 1800 $cm^{-1} = 5.56$ μm. The defect function is proportional to the integrated area under that reflectance curve.

Figure 10-14 plots $R(\sigma)$ for a film of optical thickness $\lambda_0/4$ and index $n_1 = 2.00$, which is the square root of the substrate [see Eq. (2-23j)].
and is less than 1.00, as long as the layer refractive index is bounded, $n_0 < n_1 < n_s$ .The upper curve in Fig. 10-14 shows the increase in reflectance (and, hence, an increase in the defect) that occurs when n_1 is non-optimal — i.e., its index is not the square root. An even greater increase in integrated reflectance occurs when the thickness is increased from a quarterwave to a full wave, as shown in Fig. 10-15. The normalized defect function $\mathcal{D}'$ is

$$\mathcal{D}' \equiv \frac{\text{Defect of coated substrate}}{\text{Defect of uncoated substrate}} \qquad \text{(10-2c)}$$

Finally, Fig. 10-16 depicts the normalized defect function vs the optical thickness of the layer. It has a sharp minimum at Q, which corresponds to the quarterwave thickness curve in Fig. 10-15. There are other minima in the defect curve, as, for example, at F, where the optical thickness is a full wave.

The Q point in Fig. 10-16 is termed a *global* minimum. This corresponds to the solid curve in Fig. 10-14. The minimum at F is a *local* minimum. That is, if the user begins an optimization process with an optical thickness of 0.375 of a wave (corresponding to point A on that graph), most optimization codes will seek out a local minimum and will *decrease* the thickness until point Q is reached, as shown by the arrow in that Fig. If, however, the starting design had a layer thickness of 1.15 waves (at point B), then the computer code would decrease the thickness until the *local* minimum at F is attained.

Suppose that there are two construction parameters — both the refractive index and optical thickness are varied. Figure 10-17 graphs isometrically the normalized defect function vs those parameters. The entire defect surface gains in altitude as the refractive index strays from its optimal value of 2.00. Suppose that the starting design has an index of 3.0 and an optical thickness of 1.15 waves. The optimization procedure would use the "steepest descent" algorithm to drive the thickness and index to point F in Fig. 10-17. This is a local minimum. At point Q there is a global minimum.

As the final example, Fig. 6-32 depicts the transmittance of a linear polarizer that achieves its broader bandwidth by burying a layer of halfwave optical thickness in the interior of the stack. The remainder of the layers are matched quarterwaves. Suppose a starting design consisting entirely of quarterwave layers is used and the optimization computer code is to broaden the bandwidth. The thin film optimization software produced prior to 1997 was not "smart" enough to add the aforementioned halfwave.

§10.2.6.2 addresses the problem of finding a global minimum, as contrasted with a local minimum.

10.2.4.3 Starting design — layer count

It is axiomatic that certain designs must contain at least a minimum number of layers. A simple example is that it is questionable whether the broadband antireflection coating whose reflectance is depicted in Fig. 4-2 could be produced with three layers. A mirror whose reflectance exceeds 95% in the visible portion of the spectrum usually has at least 15 layers. A starting design containing fewer than the needed number of layers is destined for failure. Experience is the teacher of how many layers to use.

10.2.4.4 Starting design — antireflection coating

A "trick" in choosing a starting design for an antireflection coating is to assume that the coating is initially contiguous to a fictitious incident medium of index n_0 higher than the actual material. Suppose that an antireflection coating is designed that matches BK-7 glass to air. The optimization is carried out for a design that matches BK-7 to a fictitious incident medium of index $n_0 = 1.3$. The index n_0 is then decreased somewhat, say, to 1.2, and the optimization repeated. This process is executed again and again until an index for n_0 of 1.0003 is reached.

10.2.4.5 Starting design — Darwinian survival

Coursen (Quoted in Aguilera [88]) suggested as a starting design for an antireflection coating

air (**L H**)^{10} **L** *substrate*

where the optical thickness of **L** and **H** is less than $\lambda/8$ in the region where the coating is to function. Computer optimization unleashes the awesome forces of Darwinian survival. Some layers survive and others perish. Layers are discarded and coalesced by the optimization process so that in many instances the final design contains one-half to one-quarter of the number of thin layers in the starting design. The author has used this approach frequently to design antireflection coatings, beamdividers and phase retarders (see §5.7.4). One reason that it works so well is that the change of R or T with thickness is more linear when the layers are relatively thin. It follows that because the optimization process relies on linear algebra, it functions more efficiently. §10.2.6 discusses alternative starting designs.

10.2.5 Miscellaneous topics

This section discusses practical problems encountered in the implementation of optimization computer codes.

10.2.5.1 Weights

The term "*automatic* correction" to describe an optimization process is often a misnomer. In most cases there is little *automatic* in the optimization process because human judgment is required to keep it on track. This is illustrated in Fig. 4-45, which shows that there is a sharp increase in reflectance near 800 nm. A greater weight on the target at that wavelength could dampen that rise. Many optimization codes input the *sensitivity*, which is the reciprocal of the weight W_i.

10.2.5.2 Layers of negative thickness

It is not unusual for an optimization code to produce layers that are either: (1) exceedingly thin; or (2) of negative thickness. Several approaches are available to cope with these events: (1) Permit the optimization to proceed a bit with a layer of negative thickness. Eventually it may cease its downward trend and become positive again. ("Hope springs eternal in the human breast.") (2) Constrain the optimization so that layers below a prespecified minimum thickness are not permitted. (3) Discard the layer the moment its thickness goes to zero. (4) Permit the thickness of the layer to become negative and replace it eventually with a full wave optical thickness layer plus its negative optical thickness. This last possible action is recommended only if the spectral bandwidth is relatively narrow. In each case, a different design will evolve.

As an example, I was entranced by the screen of the computer as the optimization program improved the design of an antireflection coating similar to that described in §4.1.3.3. One of the layers had a negative thickness for five consecutive optimization cycles. I was about to halt the process and discard the layer when a telephone interrupted, as the computer continued relentlessly to optimize. Upon my return, the thickness had again become positive.

10.2.5.3 Constraints

The goal is to achieve a *practical* solution. In other words, solutions that require layers with negative thicknesses or refractive indices as large as 15 are less than useful. There is a simple way to impose a constraint upon a construction parameter [Bloom 86]. This is illustrated by an optimization code that is not permitted to produce a metric thickness h in excess of 345 nm. The designer of the computer code simply adds to the defect function a term

$$(h/345)^{100} \ . \tag{10-2d}$$

The contribution of that term is small as long as the number within the brackets is fractional. It becomes large the instant h exceeds 345.

10.2.5.4 Alternative definition of the defect function

The objective is to bias the optimization so that the values of reflectance, transmittance or phase shift furthest from the target are weighted the most heavily. An alternative defect function is

$$\mathcal{D} \equiv \left[\frac{\sum_{i=1}^{m} W_i \ (R_i - R_{T,i})^q}{\sum_{i=1}^{m} W_i} \right]^{\frac{1}{q}} \ . \tag{10-2e}$$

As an illustration, suppose the target is 50% reflectance and at certain wavelengths the starting design manifests a reflectance of 90%. If a larger value of q is used in Eq. (10-2e), say 10, the net effect is that the optimization procedure concentrates its attention on the largest defects and neglects the smaller ones.

10.2.6 Alternative starting designs

§10.2.4.5 suggests that a starting design for optimization is a stack of many relatively thin layers. There are other approaches, such as: (1) The "flip-flop" method; (2) global optimization; (3) needle layer optimization; and (4) linear programming.

10.2.6.1 "Flip-flop" method

In the "flip-flop" procedure [Southwell 85], the starting design is a stack of relatively thin layers of alternating high **H** and low **L** refractive indices. At a place in the interior of the stack, the physical locations of two adjacent layers are altered by interchanging them. The defect function before and after the interchange is computed. A decrease in the defect means that the interchange is permanent — otherwise, the interchange is rejected. This procedure is repeated with different pairs until no decrease of the defect function is produced by any of the interchanges.

10.2.6.2 Global optimization

A global optimization procedure strives to avoid a design that produces a *local* minimum (see §10.2.4.2). The global minimum is (by definition) the lowest of all possible minima. The code developed by Dobrowolski [65], for example, starts with a design in which the thickness of each layer is chosen from a list of prescribed values. The refractive index of each layer is selected from a list. Each of the foregoing indices and thicknesses is then tried for each layer.

Suppose, for example, that five thicknesses and four refractive indices are possible for each layer. If there are six layers in the stack, the number of permutations is $(5 \times 4)^6 = 5.3 \times 10^5$. The defect function is found for each design. The winner is the one with the lowest defect function. It is possible that the extension of the layer count to twelve would tax the computational power of even a supercomputer.

10.2.6.3 Needle layer optimization

This optimization method was developed by Tikhonravov [Furman 92]. It provides a means of augmenting the number of layers in a stack by the insertion of one or more thin layers — the needle layers. Where are these layers inserted? The program scans each layer and searches for a sensitive one and in particular the point at which the sensitivity *within that layer* is negative and an extremum. The sensitivity is the change of the defect produced by the insertion of the needle layer, which is sometimes less than 1 nm thick.

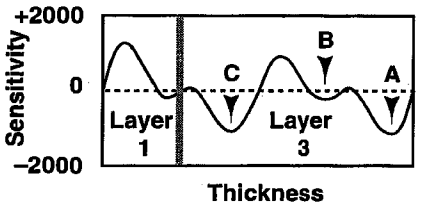


Fig. 10-18 A thin titania layer (vertical shaded line) is sandwiched between two MgF_2 layers. The caption is continued below.

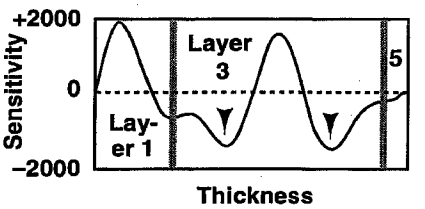


Fig. 10-19 Sensitivity after insertion of a needle layer of titania at point *A* in Fig. 10-18. The caption is continued below.

For both Figs., sensitivity of the defect function vs layer thickness is graphed for a defect function consisting of a low R over 2:1 bandwidth. The incident medium is on the left and the BK-7 substrate is on the right.

This sensitivity appears in Fig. 10-18 for a three-layer stack. The arrowed positions indicate the extrema of the sensitivity. A needle layer is then inserted at point A. In other words, the computer code splits asunder layer three — the layer count is incremented by two. The resulting five-layer coating is delineated in Fig. 10-19. The optimization then proceeds by adjusting (i.e., optimizing) the layer thicknesses of this new (augmented) design. Variations of this method allow the simultaneous insertion of needle layers at multiple points.

Often the starting design for needle layer optimization is just one thick layer. The starting design in Fig. 10-18, for example, is a magnesium fluoride layer 1.5 waves in optical thickness at 533 nm.

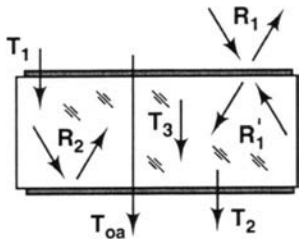


Fig. 10-20 A slab with parallel plane surfaces is coated on both sides. The coating on its upper side has a transmittance T_1 and reflectance R_1' from its emergent (i.e., substrate) side. The coating on its lower side has a transmittance T_2 and reflectance R_2 as measured inside the slab. T_{oa} is its overall transmittance.

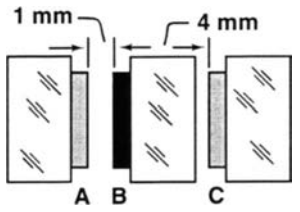


Fig. 10-21 A bandpass filter (*B*) is sandwiched between blocking rejection filters (*A* and *C*) that extend the spectral region over which the overall transmittance is relatively low. The physical separation (in mm) of the filters is shown.

$\lambda_0 = 577$ nm.

10.2.6.4 Linear programming

As mentioned in §4.9.4, this is a method of adjusting the refractive indices of a stack containing layers of equal optical thickness [Walsh 77]. Its advantage is that linear programming constrains the refractive indices between a lower and upper limit.

10.2.7 Further exploration of the subject

The discussion above focuses on the *use* of an optimization code, rather than on *how* such a code is written. Thelen [89] lists an optimization code in Basic computer language. Dobrowolski [90] compares the computational efficiency of optimization numerical procedures.

10.2.8 Appendix — design of edge filters

The designs of the short-wave pass edge filters (*shaded* in Fig. 10-10) are

air (0.5**L** **H** 0.5**L**)^{11} 0.5**L** *glass*

and, after correction, the adjusted layer thicknesses (*solid* in Fig. 10-10) are

air 0.490**L** 1.031**H** 1.002**L** 0.984**H** 0.993**L** 1.005**H** 0.986**L** 0.957**H** 0.983**L** 0.998**H** 0.99**L** 0.965**H** 0.983**L** 1.0**H** 0.997**L** 0.997**H** 0.986**L** 0.995**H** 0.998**L** 1.005**H** 0.995**L** 0.993**H** 0.988**L** *glass.*

The refractive indices of **L**, **H** and *glass* are 1.45, 2.30 and 1.517, respectively. **L** and **H** represent layers of optical thickness $\lambda_0/4$ at

10.3 Overall transmittance of an array of coated objects

10.3.1 Introduction

Many examples could be cited that illustrate the meaning of a "coated object" above, such as a substrate coated on both sides or an array of slabs with coatings on their surfaces. The question is, what is the overall transmittance or overall reflectance? A partial answer is that it depends upon the apparatus used to measure it. Different results occur when the transmittance is measured with a conventional spectrophotometer (with its relatively broad spectral bandwidth) or a dye laser with quite narrow bandwidth.

The transmittance of a multilayer computed using Equations such Eq. (2-27a) or Eq. (2-34g) is *from* the incident medium *into* the substrate. The computation is complete *if* the substrate is the detector of radiant flux, which is rare. In many cases there is a Fresnel reflection loss on its emergent side of the substrate or, as shown in Fig. 10-20, an additional coating. What is the overall transmittance T_{oa} of such a configuration?

A still more complex situation occurs when a bandpass filter is sandwiched between highly reflecting blocking filters, as shown in Fig. 10-21. The function of these *blockers* is discussed in §1.6.5. Given a computer code that calculates the R and T of each filter, how is the T_{oa} of the system found? How does this computed T_{oa} compare with the R and T measurement of a spectrophotometer?

The answers to the questions posed in the last sentences of the foregoing paragraphs are: (1) the matrix theory of §2.4.4 must be used; or (2) simpler equations in §10.3 are utilized. The choice between those options is made once the bandwidth of the monochromator and the physical separation of the elements are known.

Another topic in this section is the transmittance of a multilayer in a beam of uncollimated flux, as the cone depicted in Fig. 1-17. A typical spectrophotometer illuminates a multilayer with a convergent beam that ranges from $f/4$ to $f/8$. §10.3.7 develops equations for the transmittance of a multilayer in convergent flux.

10.3.2 Effects of illumination with non-monochromatic light

10.3.2.1 Spectral bandwidth

§2.10.1 states the assumptions about the flux (i.e., radiant energy) that illuminates the multilayer stacks which are considered throughout Chap. 2. In the language of Born & Wolf [75], it is assumed to be quasi-monochromatic, that is, nearly monochromatic. The radiation emerging from a monochromator, however, is far from monochromatic. In §10.3, the effects of the finite spectral bandwidth of that radiation is examined.

10.3.2.2 Bandwidth and coherence

The spectral bandwidth of the radiation is related to its *temporal coherence*. Two extreme cases are cited: (a) the flux emanating from a frequency-stabilized and mode-stabilized CW laser has a quite narrow spectral bandwidth and hence manifests a high degree of temporal coherence; or (b) the flux from an incandescent filament over a broad range of wavelengths has a degree of temporal coherence that is essentially zero and thus is termed *incoherent*. Between these two extremes, the flux is *partially* coherent — using the terminology of Chap. 10 of Born & Wolf [75]. If the flux is coherent, then matrix theory (see §2.4.4) must be used to find the transmittance of a slab, as, for example, that in Fig. 10-20. The formalism of §10.3.5 is used if the flux is incoherent.

At the risk of oversimplification, the following procedure is used to establish the coherence. The term *system* refers to an array of reflecting and transmitting surfaces, such as the single slab in Fig. 10-20 or the more complicated array of elements in Fig. 10-21. It is assumed that the *system* is irradiated with flux characterized spectrally by a bandwidth $\Delta\lambda$ and center wavelength λ_0. The flux from a monochromator and the emission line from a low pressure mercury discharge lamp are examples. The coherence length d_c of the radiation is

$$d_c = \lambda_0^2 / \Delta\lambda \qquad (10\text{-}3a)$$

which shows that the length decreases with increasing bandwidth.

Given the smallest of the physical separation of the elements in a *system*, the flux may be regarded as *incoherent* if its coherence length is much less than that separation.

10.3.2.3 Coherence length — examples

In the first example, the *system* mentioned in the foregoing paragraph consists of a thin sheet of plastic 40 μm in thickness coated on one side with a multilayer. It is illuminated with flux from a monochromator (λ_0 = 1250 nm) that has a bandwidth of 5 nm. The coherence length d_c of this flux is

$$d_c = (1.25)^2 / 0.005 \approx 300\ \mu m \approx 0.3\ mm \qquad (10\text{-}3b)$$

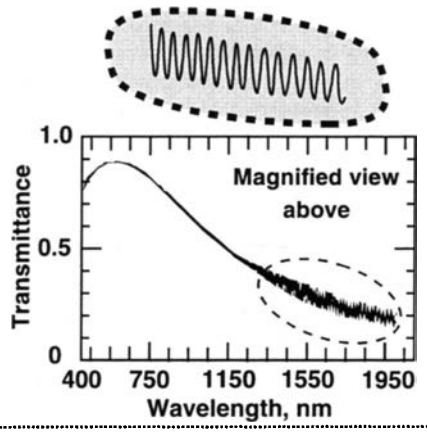


Fig. 10-22 Transmittance of a thin polyester sheet coated with a multilayer containing a noble metal. From the data of Southwall [90].

which is considerably greater than the thickness of the sheet — the element. The flux is, therefore, *coherent*. A consequence is that there is optical interference between the surfaces of the sheet of plastic. This is resolvable in a spectrophotometer, as depicted in Fig. 10-22. The undulations in transmittance are manifestations of the interference. The spacing between them would change if the thickness of the plastic were altered.

In the next example, that aforementioned light source illuminates the same coating, with the exception that the multilayer is deposited on a glass slab 5 mm in thickness. In this case, the coherence length of 0.3 mm is much less than the slab thickness. The flux is incoherent, which means it is valid to use Eq. (10-7e) to find its overall transmittance.

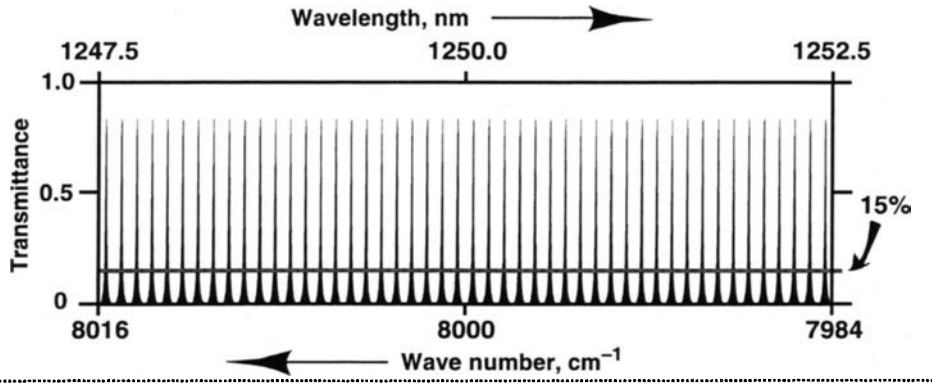


Fig. 10-23 $T(\sigma)$ of a nonabsorbing single layer of refractive index 1.507 and metric thickness 5 mm. Other details appear in §10.3.2.5. The upper abscissa is wavelength.

As the final example, suppose that there are three reflecting elements in an array, which are designated as *A*, *B* and *C* in Fig. 10-21. There is a separation of 1 mm between elements *A* and *B* and 4 mm between elements *B* and *C*. The smaller of the two separations (i.e., 1 mm) is used as a criterion for determining whether the coherence length is smaller than the physical separation of the elements.

10.3.2.4 Transmittance in non-monochromatic flux

This is an illustration (not a proof) of the validity of the procedure in the previous paragraph. It is assumed that a glass slab (index $n_c = 1.507$ and thickness $h_c = 5$ mm) is coated on both sides. The slab is nonabsorbing, which means that its internal transmittance T_3 (see §10.3.4.2) is 100%. On one side the multilayer has a reflectance R_1 of 80%. The opposite side has a coating with a reflectance R_2 of 60%. Both coatings have negligible absorption, so that $T_1 = 1 - R_1$ and $T_2 = 1 - R_2$. The aforementioned R values are assumed to be relatively constant over the 5 nm bandwidth of the light source.

The transmittance vs wave number is shown in Fig. 10-23. The curve consists of a series of narrow spikes, similar to those in Fig. 2-97. The wave number separation $\Delta\sigma_f$ between the centers of adjacent spikes is found from Eq. (2-83i)

$$\Delta\sigma_f = (2\, n_c\, h_c)^{-1} = (2 \times 1.507 \times 5000)^{-1} = 66.4\ 10^{-6}\ (\mu\text{m})^{-1} = 0.664\ \text{cm}^{-1}\ . \qquad (10\text{-}3c)$$

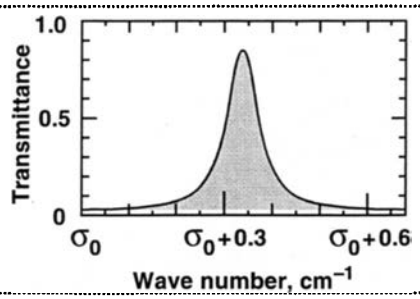


Fig. 10-24 Transmittance (over a more restricted range of wave number) of the coating captioned in Fig. 10-23. σ_0 is 0.8 reciprocal micrometer (i.e., 8000 cm^{-1}).

Next, assume that this slab is illuminated with a light source that has a center wavelength of 1250 nm [equivalent to 0.8 $(\mu m)^{-1}$ in wave number] and bandwidth of 5 nm. This 5 nm converts into a bandwidth (in wave number) of

$$\Delta\sigma_b = 3.2\times10^{-3}\ (\mu m)^{-1}\ . \qquad (10\text{-}3d)$$

The ratio of $\Delta\sigma_b$ to $\Delta\sigma_f$ is the number of transmission spikes within the bandwidth, which is approximately 48 in this example. The abscissa length in Fig. 10-23 is precisely the bandwidth $\Delta\sigma_b$.

Assuming that the detector has a uniform responsivity over the spectral region of the monochromator bandwidth extending from σ_1 to σ_2, the overall transmittance T_{oa} is found by averaging over that passband

$$T_{oa} = (\sigma_2 - \sigma_1)^{-1} \int_{\sigma_1}^{\sigma_2} T(\sigma)\, d\sigma \qquad (10\text{-}3e)$$

where $T(\sigma)$ is obtained from Eq. (2-44f). The evaluation of the integral is greatly simplified by noting that the integrated transmittance is approximately 48 times that of a single spike, as shown in Fig. 10-24. The average transmittance is thus obtained by integrating over a single band. Substitution of Eq. (2-44f) into Eq.(10-3e) provides

$$T_{oa} = \pi^{-1} \int_0^{\pi} \frac{T_1\, T_2\, (1 + R_m^2)^{-1}}{1 + b\cos\beta}\, d\beta \qquad (10\text{-}3f)$$

where the integrand has been converted to phase thickness β and the constant b is

$$b \equiv -2\, R_m\, (1 + R_m^2)^{-1}\ . \qquad (10\text{-}3g)$$

Exclusive of the multiplier π^{-1} and numerator, Eq. (10-3f) evaluates to $\pi\ (1 - b^2)^{-(1/2)}$. Substitution of R_m from Eq. (2-45a) provides for Eqs. (10-3f) and (10-3g)

$$T_{oa} = \frac{T_1\, T_2}{1 - R_1'\, R_2} \qquad (10\text{-}4a)$$

which is identical to Eq. (10-7e) — with $T_3 = 1$. The mean transmittance, as computed above, is termed the overall transmittance T_{oa} throughout this book. It is found by summing the irradiance reflected from, or transmitted by, each *element* in an array. As explained in §10.3.4, the slab in this example is considered to be an array of three elements — a multilayer on each surface and the slab.

This example should not be abandoned without a numerical result. As shown in Fig. 10-24, although the peak transmittance of each spike is 85%, the average transmittance (*shaded* line in Fig. 10-23) is only 15%, which is obtained by numerical evaluation of Eq. (10-4a)

$$T_{oa} = \frac{0.2 \times 0.4}{1 - 0.8 \times 0.6} = 0.15\ . \qquad (10\text{-}4b)$$

10.3.2.5 Simulation of the glass slab

The transmittance curve in Fig. 10-23 is simulated by a single layer of refractive index 1.507. A fictitious material of refractive index 27.0 is the incident medium. At its upper interface, this provides a Fresnel radiant reflectance R_1 of

$$R_1 = (27.0 - 1.507)\ (27.0 + 1.507)^{-2} = 0.8\ . \qquad (10\text{-}4c)$$

A similar calculation shows that the substrate is a material of fictitious index 11.86, which provides the radiant reflectance of 60% at the emergent interface.

10.3.3 Equations for reflectance and transmittance — incoherent illumination

10.3.3.1 Introduction

§2.4.6 uses both recursion and matrix methods to calculate the R and T of a multilayer. The same applies to arrays in which the beams inter-reflect and are collected incoherently [Dioffo 68].

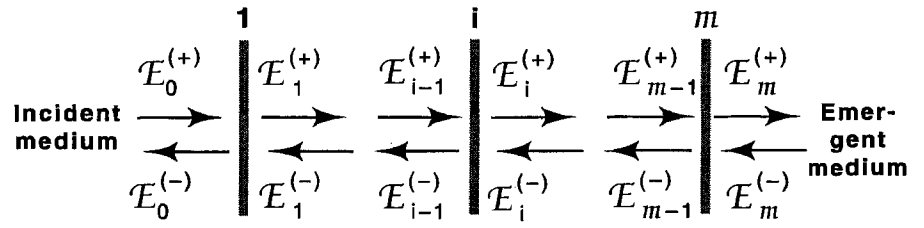


Fig. 10-25 The components of the irradiance $\mathcal{E}_i^{(+)}$ and $\mathcal{E}_i^{(-)}$ flow in the $+z$ and $-z$ directions, respectively, at the ith element.

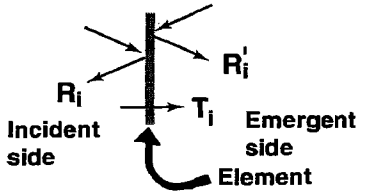


Fig. 10-26 The ith element of an array has a reflectance R_i from its front side, R_i' from its reverse side and transmittance T_i.

10.3.3.2 Basic equations

§2.10.2.4 introduces the SI symbol $\mathcal{E}$ for irradiance, which is the radiant flux per unit area. Figure 10-25 shows the incident and reflected components of the irradiance as they propagate through an array of *elements*. In the most general case, each of the m elements is both reflecting and transmitting. The source of flux is at the extreme left of the array in Fig. 10-25 and the detector is on the right. The following definitions are made (see Fig. 10-26):

$\mathcal{E}_i^{(+)}$ is the component of the irradiance propagating in the $+z$ direction on the incident side of the ith element;

$\mathcal{E}_i^{(-)}$ is the component of the irradiance propagating in the negative z direction on the incident side of the ith element;

R_i is the radiant reflectance at the incident side of the ith element;

R_i' is the radiant reflectance at the emergent side of the ith element; and

T_i is the radiant transmittance of the ith element.

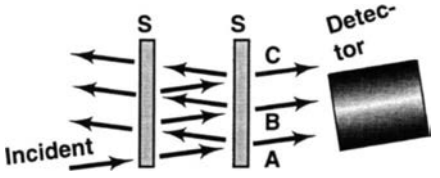


Fig. 10-27 The incidence of the flux at nonnormal incidence produces a *walk off* of the inter-reflected beams between the filters *S*. Thus beams *A* and *B* are collected by the detector, but *C* is not.

An inherent assumption is that the detector collects all beams that bounce back and forth between the elements. At normal incidence, this requirement is easily met, but at nonnormal incidence, there may be a *walk-off* of the beams, as shown in Fig. 10-27. The analysis in the remainder of this section would be inadequate to compute the net transmittance of the slabs shown in that figure.

The basic equations in §2.4.2.1 are based upon the balance of flow. They apply with equal validity to coherent or incoherent systems. Substituting *radiant* reflectances and transmittances into Eq. (2-21c) provides

$$\mathcal{E}_{i-1}^{(-)} = T_i^{-1} \left[R_i \mathcal{E}_i^{(+)} + (T_i^2 - R_i R_i') \mathcal{E}_i^{(-)} \right] \tag{10-5a}$$

$$\mathcal{E}_{i-1}^{(+)} = T_i^{-1} \left[\mathcal{E}_i^{(+)} - R_i' \mathcal{E}_i^{(-)} \right] . \tag{10-5b}$$

On the emergent side of the mth interface, there is (by definition) no reflected component of the flux. Mathematically, this means that

$$\mathcal{E}_m^{(-)} = 0 . \tag{10-5c}$$

Equations (10-5a) and (10-5b) are applied recursively, similar to the procedure in §2.4.5.2. That is, $\mathcal{E}_m^{(+)}$ is set equal to 1.00 and then $\mathcal{E}_{m-1}^{(+)}$, $\mathcal{E}_{m-1}^{(-)}$, $\mathcal{E}_{m-2}^{(+)}$, $\mathcal{E}_{m-2}^{(-)}$ and so on are found until the incident medium is encountered. The overall reflectance R_{oa} and overall transmittance T_{oa} are

$$R_{oa} \equiv \mathcal{E}_0^{(-)} / \mathcal{E}_0^{(+)} \tag{10-5d}$$

and

$$T_{oa} \equiv 1 / \mathcal{E}_0^{(+)} . \tag{10-5e}$$

10.3.3.3 Matrix methods

There are several ways in which Eqs. (10-5a) and (10-5b) are written in matrix form. One is similar to Eq. (2-21c)

$$\begin{bmatrix} \mathcal{E}_{i-1}^{(+)} \\ \mathcal{E}_{i-1}^{(-)} \end{bmatrix} = T_i^{-1} \begin{bmatrix} M_{11}^{(i)} & M_{12}^{(i)} \\ M_{21}^{(i)} & M_{22}^{(i)} \end{bmatrix} \begin{bmatrix} \mathcal{E}_i^{(+)} \\ \mathcal{E}_i^{(-)} \end{bmatrix} \equiv T_i^{-1} \, \mathrm{D}_i \begin{bmatrix} \mathcal{E}_i^{(+)} \\ \mathcal{E}_i^{(-)} \end{bmatrix} \tag{10-5f}$$

where

$$\mathrm{D}_i \equiv \begin{bmatrix} 1 & -R_i' \\ R_i & T_i^2 - R_i R_i' \end{bmatrix} . \tag{10-5g}$$

The advantage of factoring the scalar T_i^{-1} out of the matrix is that the effects of inter-reflections are seen more clearly. The matrix product [see §2.6.2.1(b)] is formed

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \equiv \mathrm{D} \equiv \prod_{i=1}^{m} \mathrm{D}_i \tag{10-5h}$$

where D_i appears in Eq. (10-5g) and m is the number of elements. R_{oa} and T_{oa} follow

$$R_{oa} = M_{21}/M_{11} \tag{10-6a}$$

and

$$T_{oa} = T_\pi \; M_{11}^{-1} \tag{10-6b}$$

where the *product transmittance* T_π is

$$T_\pi \equiv T_1 \; T_2 \; T_3 \; \ldots \; T_{m-1} \; T_m \; . \tag{10-6c}$$

The foregoing formulation [Dioffo 68] suggests an approach that minimizes the overall transmittance. For the sake of argument, suppose that the transmittance of the elements in a system can be maintained at a constant level, but the absorption increased so that the reflectance is reduced. Because T_π in Eq. (10-6c) is fixed, the reflectances of the elements should be reduced so that M_{11} in Eq. (10-6b) is as close to unity as possible. Then, and only then, is T_{oa} given by Eq. (10-6c). The *absorbances add* when this occurs (see §10.3.6.2).

An alternate form of the matrices [Harrison 72] is

$$\begin{bmatrix} \mathcal{E}_{i-1}^{(+)} + \mathcal{E}_{i-1}^{(-)} \\ \mathcal{E}_{i-1}^{(+)} - \mathcal{E}_{i-1}^{(-)} \end{bmatrix} = (2\,T_i)^{-1} \begin{bmatrix} \mathcal{B}_{11}^{(i)} & \mathcal{B}_{12}^{(i)} \\ \mathcal{B}_{21}^{(i)} & \mathcal{B}_{22}^{(i)} \end{bmatrix} \begin{bmatrix} \mathcal{E}_i^{(+)} + \mathcal{E}_i^{(-)} \\ \mathcal{E}_i^{(+)} - \mathcal{E}_i^{(-)} \end{bmatrix} \tag{10-6d}$$

where the matrix elements are

$$\mathcal{B}_{11}^{(i)} \equiv (1 + R_i)\,(1 - R_i')\,) + T_i^2 \quad , \tag{10-6e}$$

$$\mathcal{B}_{12}^{(i)} \equiv (1 + R_i)\,(1 + R_i')\,) - T_i^2 \quad , \tag{10-6f}$$

$$\mathcal{B}_{21}^{(i)} \equiv (1 - R_i)\,(1 - R_i')\,) - T_i^2 \tag{10-6g}$$

and

$$\mathcal{B}_{22}^{(i)} \equiv (1 - R_i)\,(1 + R_i')\,) + T_i^2 \quad . \tag{10-6h}$$

Although there is little advantage computationally in expressing the matrix in the foregoing form, it is not without insight. When all elements are nonabsorbing, the difference

$$[\,\mathcal{E}_0^{(+)} - \mathcal{E}_0^{(-)}\,]\,[\mathcal{E}_m^{(+)}\,]^{-1} = 1 \tag{10-6i}$$

is unity and the sum is

$$[\,\mathcal{E}_0^{(+)} + \mathcal{E}_0^{(-)}\,]\,[\mathcal{E}_m^{(+)}\,]^{-1} = 1 + 2\,F_{oa} = 1 + 2\,R_{oa}/T_{oa} \; . \tag{10-6j}$$

The matrices simplify [T. Smith 26] and the result is

$$F_{oa} = \sum_{i=0}^{m} F_i \tag{10-6k}$$

or, in other words, the R/T of the array is the sum of the R/T of each element, provided that each is nonabsorbing. §10.3.6.1 illustrates the use of the above equation.

10.3.4 Examples of an "element"

The *element* through which the flux propagates can be: (1) A surface (§10.3.4.1); (2) a bulk absorber (§10.3.4.2); or (3) a multilayer. In the first two instances the matrix simplifies.

10.3.4.1 Fresnel radiant reflection at a surface

In this case, $R \equiv R'$ and $R + T = 1$. The matrix in Eq. (10-5g) simplifies to

$$\mathrm{D}_i = \begin{bmatrix} 1 & -R_i \\ R_i & 1-2R_i \end{bmatrix} . \qquad (10\text{-}7a)$$

10.3.4.2 Absorbing medium

This medium is characterized by its *internal transmittance* [see Eq. (2-14a)]. This is the transmittance due to the attenuation of the absorbing medium itself, neglecting any surface reflection losses. One convention is to establish the internal transmittance T_0 for a *standard thickness* h_0. If an absorber of thickness h_1 is used, its internal transmittance T_1 is

$$T_1 = [T_0]^{h_1/h_0} . \qquad (10\text{-}7b)$$

To cite a numerical example, suppose that the internal transmittance T_0 is 82% for a 3 mm slab. From Eq. (10-7b), what is its internal transmittance T_1 when its thickness is 2 mm?

$$T_1 = [0.82]^{2/3} = 88\% . \qquad (10\text{-}7c)$$

The matrix [used in Eq. (10-5h)] for a bulk absorber of internal transmittance T_i is

$$\mathrm{D}_i = \begin{bmatrix} 1 & 0 \\ 0 & T_i^2 \end{bmatrix} . \qquad (10\text{-}7d)$$

The zero antidiagonal elements are manifestations of the fact that (by definition) it is nonreflecting.

10.3.5 Application to the transmittance of a single slab

10.3.5.1 Basic equation

An equation is written for the transmittance and reflectance of a slab. Some definitions are:

R_1' is the reflectance of the emergent side of the top surface of the slab, as viewed from the interior of the slab;

R_1 is the reflectance of the incident side of the top surface of the slab, as measured from the incident medium;

T_1 is the transmittance of the top surface of the slab;

R_2 is the reflectance of the incident side of the bottom surface of the slab, as measured from the interior of the slab;

T_2 is the transmittance of the bottom surface of the slab; and

T_3 is the internal transmittance of the slab itself.

The nomenclature appears in Fig. 10-20. The overall transmittance T_{oa} of the slab is

$$T_{oa} = \frac{T_1 T_2 T_3}{1 - R_1' R_2 T_3^2} \qquad (10\text{-}7e)$$

where T_1 and T_2 are the transmittances of the surfaces. It is possible that the coatings on the surfaces of the slab are, themselves, absorbing. If the internal transmittance T_3 of the

slab is less than 100%, this is due to its absorption alone. In other words, T_3 is measured when both sides of the slab are coated with "perfect" antireflection coatings.

The foregoing equation provides insight. The denominator is due to inter-reflections between the two surfaces. That is, if either R_1' or R_2 is zero, then no multiple bounces can occur and the overall transmittance is

$$T_{oa} = T_1 T_2 T_3 \qquad (10\text{-}8a)$$

or, in other words, the absorbances add. The T_3 appears in the denominator as "T squared," because the flux must pass twice through the medium in the interior of the slab to emerge as a multiple reflection.

A special case of the foregoing equation occurs when the slab is a nonabsorbing medium of refractive index n surrounded on both sides with air. Substitution of Eq. (2-12e) into Eq. (10-7e) provides

$$T_{oa} = \frac{2n}{n^2 + 1} \quad . \qquad (10\text{-}8b)$$

At nonnormal incidence, *effective indices* can be substituted into the foregoing equation.

The overall reflectance R_{oa} is

$$R_{oa} = R_1 + \frac{T_1^2 R_2 T_3^2}{1 - R_1' R_2 T_3^2} \quad . \qquad (10\text{-}9a)$$

The second term in the foregoing equation is due to reflections within the slab itself. If the slab is highly absorbing, for example, T_3 is quite small and the second term vanishes.

10.3.6 Examples

10.3.6.1 Slab polarizer (a pile of plates)

The objective is to cascade several tilted slabs (with parallel faces) so that they transmit as much p polarized flux as possible. The s polarization is to be rejected. Two configurations of this IR linear polarizer are depicted in Fig. 10-28. On the left the slabs are parallel and on the right (*antiparallel*) the slabs are canted sufficiently so that there are no inter-reflections between the second slab and the first. By canting the slabs at Brewster's angle (see §2.2.3.3.2), the p transmittance is close to 100%. The slabs are fabricated from cadmium telluride, $n = 2.67$.

Parallel The parallel surfaces permit the detector to collect the inter-reflected beams between all four surfaces. The first step is to compute the reflectance R_s and transmittance T_s at the surface of each slab. Because at Brewster's angle $U = W$ in Eq. (2-48f) and from the definitions of U and W in Eqs. (2-48d) and (2-48e), it follows that

$$R_s = [(1 - n^2) / (1 + n^2)]^2 = 0.568 \qquad (10\text{-}9b)$$

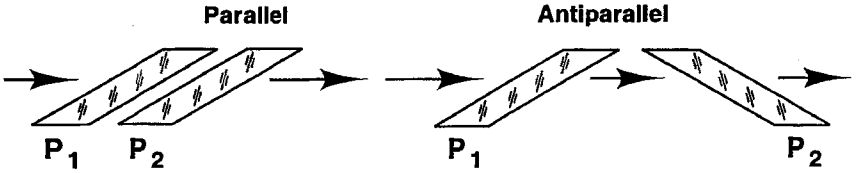


Fig. 10-28 A *piles of plates* linear polarizer consists of two nonabsorbing slabs inclined at Brewster's angle, where they are parallel (*left*) and antiparallel (*right*).

for $n = 2.67$. Equation (10-6k) is used to find the overall transmittance by use of the variable R/T. The F [see Eq. (2-24b)] of each surface is the same and is 1.317. It follows that

$$F_{oa} = 4F = 4 \times 1.317 = 5.27 \qquad (10\text{-}9c)$$

and from Eq. (2-24d) the overall transmittance is

$$T_{oa} = (1 + F_{oa})^{-1} = (1 + 5.27)^{-1} = 0.16 \ . \qquad (10\text{-}9d)$$

That is, one-sixth of the (unwanted) s polarized flux leaks through the slabs, due to inter-reflections.

Antiparallel The slabs are antiparallel. The overall transmittance T_2 of a single slab is similar to Eq. (10-9c), with the exception that two surfaces (rather than four) are used. This is

$$T_2 = (1 + 2 \times 1.317)^{-1} = 0.275 \ . \qquad (10\text{-}9e)$$

Because there are no inter-reflections between the slabs, the *product transmittance* in Eq. (10-6c) for s polarized flux is

$$T_{oa} = T_2^2 = (0.275)^2 = 0.076 \ . \qquad (10\text{-}9f)$$

This is an improvement, although the polarizer has been made bulkier.

The degree of polarization P_t [see Eq. (1-5c)] quantifies the efficiency of this device as a polarizer. Its s transmittance is given above and the p transmittance is 100%:

$$P_t = (1 - 0.076) \times (1 + 0.076)^{-1} = 0.86 \ . \qquad (10\text{-}9g)$$

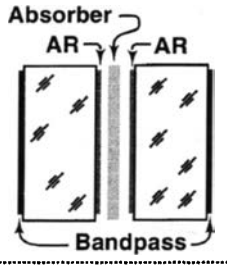


Fig. 10-29 Two dielectric multilayer bandpasses (whose optical properties are in Table 10-3) are coated on slab substrates with negligible absorption. The rear surfaces of the slabs are antireflection coated. A weakly absorbing material is inserted between the slabs.

10.3.6.2 Dielectric multilayer stacks

This section addresses the problem of improving the blocking (see §1.6.5) provided by dielectric coatings.

In an example, two dielectric multilayer bandpasses are positioned in tandem. As illustrated in Fig. 10-29, each is coated on a substrate with an antireflection coating on its opposite side. Table 10-3 lists the reflectance and transmittance of each multilayer in both the passband and stop band spectral region. The goal is to decrease the overall transmittance in the stop band spectral region.

Case A: In this configuration, the two multilayers are placed in tandem so that the incident flux traverses both of them. The internal transmittance of each substrate is essentially 100%. Eq. (10-7e) is used to find the overall transmittance T_{oa}. Table 10-3 shows that the transmittance of each multilayer is reasonably high in the passband region and that the reflectance is only a few percent.

$$T_{oa} = \frac{0.88 \times 1.0 \times 0.88}{1 - (1.0)^2 \times 0.025 \times 0.03} = \frac{0.7744}{0.9993} = 0.775 \ . \qquad (10\text{-}10a)$$

The "1.0" above is the internal transmittance of the substrate. The denominator of the foregoing equation is close to 1. This is not the case in the stop band region, where the T_{oa} is

$$T_{oa} = \frac{0.002 \times 1.0 \times 0.002}{1 - (1.00)^2 \times 0.995 \times 0.994} = \frac{(0.002)^2}{0.011} = 3.6\times10^{-4} \ . \qquad (10\text{-}10b)$$

Table 10-3 Overall transmittance of tandem bandpass filters

Spectral region	R	R'	T	T_{oa}	T''_{oa}
Stop band	0.994	0.995	0.002	3.6×10^{-4}	5.6×10^{-5}
Passband	0.025	0.030	0.88	0.775	0.752

Obverse reflectance R, reverse reflectance R' and transmittance T of identical bandpass filters in two spectral regions. T_{oa} is the overall transmittance [see Eq. (10-7e)] when the two filters are in tandem. T''_{oa} is the overall transmittance when the filters are sandwiched between a nonreflecting medium that absorbs 3%.

The effect of multiple reflections is evident. The relatively high reflectance in the stop band spectral region causes the denominator of Eq. (10-10b) to be small (compared to unity), thus increasing T_{oa} considerably.

Case B: A weak absorber is added. An absorber with internal transmittance of 97% is inserted between the multilayers. The passband overall transmittance is

$$T_{oa} = \frac{0.88 \times 0.97 \times 0.88}{1 - (0.97)^2 \times 0.025 \times 0.03} = 0.752 \qquad (10\text{-}10c)$$

and the stop band transmittance is

$$T_{oa} = \frac{0.002 \times 0.97 \times 0.002}{1 - (0.97)^2 \times 0.995 \times 0.994} = \frac{0.97 \times (0.002)^2}{0.0694} = 5.6\times10^{-5} \ . \qquad (10\text{-}10d)$$

A comparison of Eqs. (10-10b) and (10-10d) shows that the stop band transmittance is reduced by a factor of six. The introduction of the weak absorber reduces the passband transmittance by only a mere 2% — from 77.5% to 75.2%. This is an advantageous trade and is often used to reduce the offband transmittance of bandpasses.

The conclusion is that it is often advantageous to insert weakly absorbing media, as shown in Fig. 10-29, in order to enhance the offband rejection of tandem filter arrays.

10.3.7 Transmittance in convergent flux

10.3.7.1 Introduction

The definition of radiant transmittance in Eq. (2-11a) is still valid if the flux is uncollimated, provided the total flux in the incident and transmitted beams is collected by the detector. It is assumed that the radiation is incoherent over the cone of the incident flux, as shown in Fig. 1-17. The resulting transmittance T_b is termed *biconical* because a cone of flux is collected in both the incident and emergent beams, as defined in Eq. (1-6c).

10.3.7.2 Definition of biconical transmittance

The *intensity* $I(\xi,\phi)$ of the source of flux impinging upon the multilayer must be known. *Intensity*, as used in radiometry, is the flux per solid angle. The angles ξ and ϕ are defined in the spherical coordinate system shown in Fig. 10-30 as the polar and azimuthal angles, respectively. The source is at the origin.

\The biconical transmittance T_b is

$$T_b = \frac{\int_{\phi=0}^{2\pi} \int_{\xi=0}^{\xi_{max}} I(\xi,\phi)\ T(\xi,\phi)\ \sin\xi\, d\xi\, d\phi}{\int_{\phi=0}^{2\pi} \int_{\xi=0}^{\xi_{max}} I(\xi,\phi)\ \sin\xi\, d\xi\, d\phi} \qquad (10\text{-}11a)$$

where the factor "$\sin\xi\ d\xi\ d\phi$" represents the differential solid angle and ξ_{max} is the maximum semi-angle of the convergent cone. Of special interest is a Lambertian source in which $I(\xi,\phi) \equiv I_0 \cos\xi$. Equation (10-11a) simplifies to

$$T_b = (\pi \sin^2\xi_{max})^{-1} \int_{\phi=0}^{2\pi} \int_{\xi=0}^{\xi_{max}} T(\xi,\phi) \sin\xi \cos\xi\, d\xi\, d\phi \qquad (10\text{-}11b)$$

If the multilayer is perpendicular to the principal ray in the cone, the foregoing reduces to

$$T_b = 2\,(\sin\theta_{max})^{-2} \int_{\theta=0}^{\theta_{max}} T(\theta) \sin\theta \cos\theta\, d\theta \qquad (10\text{-}11c)$$

where $\xi = \theta$ is the angle of incidence upon the multilayer. θ_{max} is the maximum semi-angle of the convergent cone.

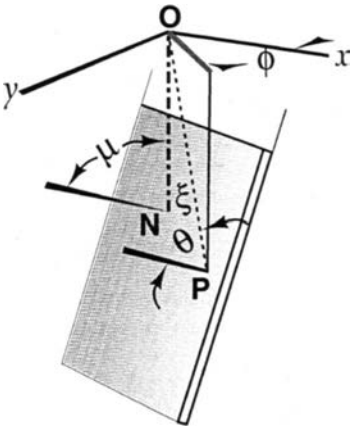


Fig. 10-30 A slab is tilted about the x axis at an angle μ, as measured to the z axis (*long-short dash*). A skew ray $\overline{OP}$ is represented as a line (*short dash*) from the origin O to P.

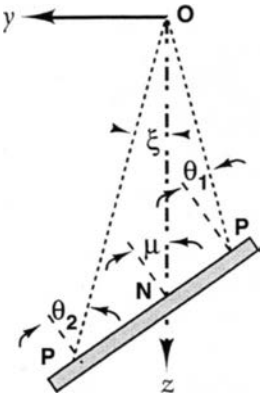


Fig. 10-31 The caption to Fig. 10-30 obtains, with the exception that the meridional rays (*short-dash*) in the y-z plane are shown. N is a normal to the slab and ξ is the angle of the ray to the z axis

10.3.7.3 Evaluation of the biconical transmittance

The numerical evaluation of Eq. (10-11a) becomes more complicated when the slab is tilted, as shown in Fig. 10-30. μ is the tilt angle as measured to the normal. Figure 10-31

depicts a cross section of the cone. A principal ray and two meridional rays impinge upon the slab. The minimum and maximum values of the angle of incidence θ (upon the slab) are

$$\theta_{min} = \mu - \xi_{max} \tag{10-11d}$$

and

$$\theta_{max} = \mu + \xi_{max} \ . \tag{10-11e}$$

There are several methods of evaluating Eq. (10-11a). Carniglia [83] uses a coordinate system that transforms from a double integral to a single integral.

There is a procedure that is computationally efficient, especially if the software package has an efficient sorting routine. The first step is to convert from $T(\xi,\phi)$ into $T(\theta)$, where θ is the "local" angle of incidence for a given ray. The spherical coordinate system is shown in Fig. 10-30. The apex of the cone is at the origin. The slab intersects the z axis at N. The distance $\overline{O\,N}$ is $\mathcal{L}$. For any ray intersecting the slab at a distance z from the origin, the angle of incidence θ is

$$\theta = \arccos(\cos\mu \cos\xi \, \mathcal{L} \, z^{-1}) \tag{10-12a}$$

which is converted into

$$\theta = \arccos(\cos\mu \cos\xi - \sin\mu \sin\xi \sin\phi) \ . \tag{10-12b}$$

The four steps in the procedure to evaluate Eq. (10-11a) are: (1) select a method of numerical integration, such as the trapezoidal rule or Simpson's rule; (2) determine the number of intervals used in the numerical integration over ξ and ϕ, which are m and q, respectively; (3) using Eq. (10-12b), compute the angle of incidence θ_{ij} at each midpoint of the intervals in step (2); (4) approximate Eq. (10-11a) by the sums

$$T_b \approx \frac{\sum_{i=0}^{m} \sum_{j=0}^{q} W_{ij} \, T(\theta_{ij})}{\sum_{i=0}^{m} \sum_{j=0}^{q} W_{ij}} \tag{10-12c}$$

where the weighting factor is

$$W_{ij} \equiv I(\xi_i , \phi_j) \sin\xi_i \, P_{ij} \tag{10-12d}$$

and the numerical weight P_{ij} is determined by the method of integration. As an illustration, Simpson's rule is used and the sum is over i. For $i = 0$ or $i = m$, the numerical weight (i.e., P_0 or P_m) is 0.33. The numerical weight P_i is 1.333 for all other even-numbered intervals and 0.667 for the odd intervals.

The result of the procedure in the foregoing paragraph is a doubly scripted array of angle of incidence values consisting of the "theta values" θ_{ij} and the weight W_{ij} associated with each theta. The theta values are now sorted from high to low. The weight associated with each theta value should tag along with it in the sort. The next step is to use a multilayer code to compute the transmittance at theta values close to those in the sorted array. Interpolation is used to find the transmittance at intermediate theta values. The interpolation is facilitated because $T(\theta)$ ascends as a monotone in θ. The sums in the numerator and denominator of Eq. (10-12c) are evaluated to find an approximate T_b .

10.3.7.4 Lambertian source and an aplanatic lens

It is shown in this section that the illumination of an optical component by an aplanatic lens is equivalent to illuminating it with a Lambertian source. An *aplanatic* lens is corrected for aberrations such as third-order coma and spherical. In Fig. 10-32, it is shown that an incident ray parallel to the optical axis at height y converges towards the focus at an angle θ

$$y = f \sin\theta \qquad (10\text{-}12e)$$

where f is the lens focal length. In other words, the principal planes of the lens are spheres. The intensity at the focus is equivalent to a distant Lambertian source.

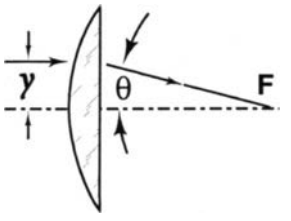


Fig. 10-32 Meridional ray at height y converges upon focus F of lens at angle θ.

Given the transmittance $T(\theta)$ of the coating as a function of angle θ, the infinitesimal flux $d\Phi$ collected in an annulus of area $2\pi y\, dy$ is

$$d\Phi = 2\pi \mathcal{E}_0\, y\, dy\,. \qquad (10\text{-}12f)$$

Because y and θ are related by Eq. (10-12e),

$$\Phi = 2\pi \mathcal{E}_0 f^2 \int_0^{\theta_m} T(\theta) \sin\theta \cos\theta\, d\theta \qquad (10\text{-}12g)$$

where θ_m is the maximum semi-angle of a ray (see Fig. 1-98). The foregoing equation is equivalent to Eq. (10-11c). This weighting factor "$\sin\theta \cos\theta$" is significant. It implies that for $\theta < 45°$, flux at higher angles contributes relatively more than flux at smaller θ.

There are several commercial software packages that compute T_b. In one study, Eq. (1-25) was not used because the lens was non-aplanatic [Baker 67].

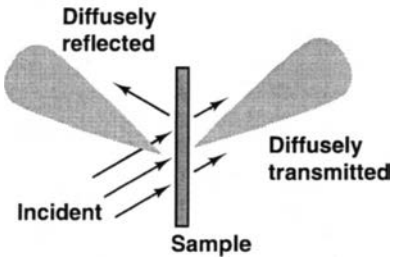


Fig. 10-33 Collimated incident flux (represented by parallel arrows) is diffusely reflected and diffusely transmitted, as represented by the shaded plumes.

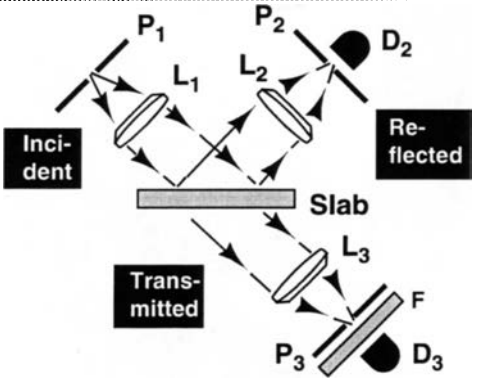


Fig. 10-34 Light from pinhole P_1 is collimated by lens L_1, thus illuminating the slab. Similar configurations transfer the light to the detectors D_2 and D_3 in the other beams, thus measuring the *specular* reflectance and *direct* transmittance.

10.4 Performance of coatings — their optical characteristics

10.4.1 Introduction

The *optical characteristics* above refer to the amount of light a coating reflects, transmits or absorbs. In the UV and IR spectral regions, "light" in the previous sentence refers to

radiant energy or, to be more precise, radiant *flux*. §10.5 focuses on non-optical properties of coatings such as adhesion and hardness.

10.4.1.1 Overview of section

§1.2.2 discusses the reflectance R and transmittance T of a coated part, as well as its overall transmittance. The tacit assumption in §1.2.2 is that the specular reflectance is measured. In §10.4.2, this is expanded to include diffuse R and diffuse T. §10.4.3 surveys the spectrophotometer that measures R and T. This section includes practical problems encountered in measurements.

10.4.2 Specular and diffuse measurements

The arrows in Fig. 10-33 show a collimated beam of light impinging upon a coated sample. The arrows in the reflected (i.e., specular) and transmitted beams represent the light that has retained its collimation. In addition, however, there is a diffuse (i.e., scattered) component, as represented by a shaded plume. The distinction between specular and diffuse is discussed below.

10.4.2.1 Specular R and direct T

The specular reflectance R is measured with the apparatus shown in Fig. 10-34. A pinhole P_1 is illuminated and the emerging light is collimated by a diffraction-limited lens. Φ_i is the radiant flux (i.e., energy per unit time) in the beam impinging upon the coated part.

A similar optical system exists for the reflected beam. The pinhole P_2 at the focus of the lens functions as a *spatial filter* that rejects any scattered light. The flux in the reflected beam is Φ_r. The *specular reflectance* R is similar to Eq. (2-11f)

$$R = \Phi_r / \Phi_i \ . \tag{10-13a}$$

The *direct transmittance* is the analog of specular reflectance. Figure 10-34 shows that a lens collects the transmitted flux Φ_t and concentrates it upon the pinhole P_3. The direct transmittance is

$$T = \Phi_t / \Phi_i \ . \tag{10-13b}$$

As mentioned in §1.2.2, the T measured above is the overall transmittance. But assume, for purposes of argument, that the rear side of the part in Fig. 10-34 is coated with a highly transmissive antireflection coating. The transmittance and overall transmittance are virtually the same. The next section describes the apparatus in which the scattered flux is included in the measurement.

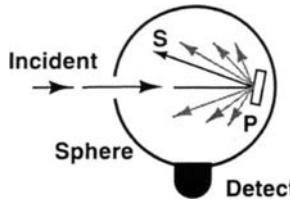


Fig. 10-35 *Directional hemispherical total reflectance* of the part P is measured by tilting it inside the integrating sphere so that both the specular and diffuse components reflect from the diffusely reflecting interior of the sphere.

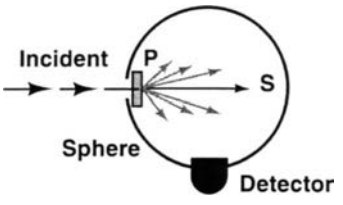


Fig. 10-36 *Total directional transmittance* of the part P is measured. Both the diffuse and direct components impinge upon the diffusely reflecting interior of the sphere.

10.4.2.2 Reflectance — both diffuse and hemispherical

In all coatings and substrates, both surface roughness and inclusions in the bulk materials scatter a portion of the incident flux. This is measured with an integrating sphere, as depicted in Fig. 10-35. It measures both the specularly and diffusely reflected (i.e., scattered) components. The interior of the sphere is coated with a highly diffuse reflecting material, such as barium sulfate. A small hole in the sphere admits the incident beam, which reflects from the part under test P. Both the specular component of the beam S and the diffusely scattered reflected flux hit the diffusely reflecting walls of the sphere. This flux bounces around the interior of the sphere many times and eventually reaches the detector D. The reflectance measured in this manner is termed the *diffuse reflectance* or *directional hemispherical total reflectance.* This reflectance R_t is the sum of the specular reflectance R and scattered component R_u

$$R_t = R + R_u \ . \tag{10-13c}$$

10.4.2.3 Transmittance — diffuse and total directional

The roughness mentioned above scatters a fraction of the transmitted light. This is measured using the same integrating sphere described in §10.4.2.2. The sample is positioned over the input port, as depicted in Fig. 10-36. The directly transmitted beam S and the transmitted scattered flux impinge upon the walls of the sphere. This flux bounces around the interior of the sphere. A small part of it reaches the detector. The transmittance measured in this manner is termed the *diffuse transmittance* or *total directional transmittance.* This transmittance T_t is the sum of the direct transmittance T [see Eq. (10-13b)] and scattered component T_u

$$T_t = T + T_u \ . \tag{10-13d}$$

10.4.2.4 What is measured?

The apparatus described above can be costly to install and time consuming to use. The spectrophotometer is the instrument of choice because it is less expensive and performs a spectral scan more rapidly.

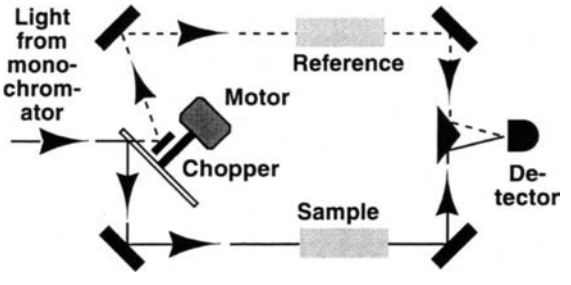


Fig. 10-37 In a dual-beam spectrophotometer, a mirror and open slot on a rotating chopper alternately shunt the monochromatic beam between the sample and reference.

10.4.3 Spectrophotometry — a brief overview

Spectrophotometer describes generically an automated instrument that measures reflectance or transmittance vs wavelength. The span of wavelength ranges from the UV (120 nm) to the far IR (100 μm). The following description of this instrument is greatly oversimplified. There are, however, at least three basic steps in its operation:

(1) Produce a beam of monochromatic light (i.e., radiant flux);

(2) a detector measures the flux with sample out of the beam; and

(3) a detector measures the transmitted or reflected flux with sample inserted into the beam.

In the discussion below, *sample* refers to the coated part — the R and T of which are to be measured by the instrument. The use of a dual-beam instrument obviates the necessity to move the sample into and out of the beam.

10.4.3.1 Production of a "monochromatic" beam

"Monochromatic" refers to "a reasonably narrow bandwidth." The following pieces of hardware have been used to produce the aforementioned beam:

(a) A Lyot-Öhman filter;
(b) multilayer bandpass (see §1.6);
(c) a laser, such as a tunable diode laser; or
(d) a diffraction grating;
(e) a prism.

In addition, the Fourier transform spectrometer is widely used in the IR. It does not produce monochromatic light. Instead, it convolves the transmittance of a sample with the transmittance of an interferometer. Because the latter is known mathematically, computer software may be used to deconvolve these two transmittances. Some instruments use a single detector, while others contain an array of detectors.

10.4.3.2 Flux ratio

The flux ratio is determined from steps (2) and (3) in §10.4.3. The measurements are made point-by-point and wavelength-by-wavelength if the hardware in method (1) of §10.4.3.1 is used. Although this usually produces accurate results, it is slow, time consuming and expensive. For this reason, the "automated" feature of the dual-beam spectrophotometer is preferred.

10.4.3.3 Dual-beam spectrophotometer

Figure 10-37 depicts a portion of a dual-beam instrument in which the light from a monochromator is alternately shunted between the sample and reference paths by a rotating mirror. That light is collected by the detector. A few specialized instruments with integrating spheres measure the total directional transmittance (see §10.4.2.3). A typical spectrophotometer measures the direct transmittance *and* the scattered transmitted flux (i.e., diffuse transmittance). In other words, it collects rays that lie at angles within a few degrees of an axial ray.

(1) These instruments are designed with a primary mission of measuring the absorbance of a chemical in a cuvette of prescribed length. They are modified to measure the transmittance of a multilayer.
(2) A typical beam size of an instrument that functions in the visible part of the spectrum is 5 mm × 15 mm. Smaller samples may be measured, but with impaired accuracy.
(3) Reflectance measurements are made using special attachments. The R at "normal incidence" is approximated with a relatively small incidence angle, such as 12°. The R of a curved surface may be measured, usually with reduced accuracy. The determination of a surface reflectance below 0.1% is prone to systematic errors, the principal one being the reflectance of the reverse side of the substrate.

(4) The insertion of a reflectance attachment or a thick substrate causes defocus of the beams converging on the detector, with an attendant loss of precision.
(5) Due to scattered light within the instrument, a typical lower limit of the transmittance is 5.0 absorbance units in the visible part of the spectrum and 3.0 in the IR.

10.4.3.4 Use of a witness

Often a coated part is too large to fit into the sample compartment of a spectrophotometer. In other instances, the coated part may be highly curved or so small that the spectrophotometer beam must be severely masked to accommodate it. For these reasons, the measurements are performed on a witness piece. Its attributes and use are discussed in §1.3.6.2.

10.4.3.5 *R* and *T* at nonnormal incidence

Suppose that a beamdivider such as the one shown in Figs. 1-180 or 1-196 is inserted into a spectrophotometer. The objective is to measure the R_a or T_a for randomly polarized light, as defined in Eqs. (1-5a) and (1-5b), respectively. The problem is that the light produced by most spectrophotometers is strongly polarized. There are two solutions:

(1) Insert linear polarizers into both sample and reference beams so that separate measurements are made for p and s linear polarizations. A computer generated numerical average produces R_a or T_a .
(2) The sample is measured in two orientations, as described below.

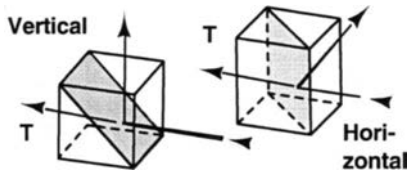


Fig. 10-38 Polarization effects in the light from a spectrophotometer are averaged by measuring the transmittance in two orientations of the sample.

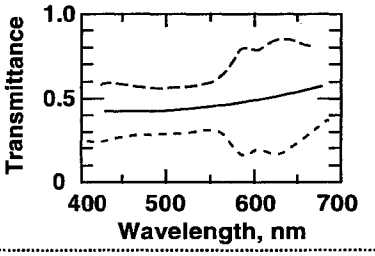


Fig. 10-39 In a spectrophotometer (Cary 5E), measured transmittance of a slab at 45° incidence in the two orientations (*dashed*) shown in Fig. 10-38. The arithmetic mean is *solid*.

In the second option, the beamdivider is rotated about an azimuthal axis collinear with the beam in the spectrophotometer, as depicted in Fig. 10-38. The transmittance is measured in one orientation, rotated 90° and measured again. This is illustrated in Fig. 10-39, which shows the spectral transmittance of an inclined slab as measured on a Varian-Cary model 5E spectrophotometer. The upper and lower curves are for the orientations in Fig. 10-38. The arithmetic average T is the middle curve in Fig. 10-39. The averaging of polarization effects produces a curve that approximates more closely T_a in Eq. (1-5b).

10.4.3.6 Effects of beam convergence

The $f$number of a typical spectrophotometer is $f/6$. In other words, the flux is not collimated as it is in Fig. 10-34. This $f$number is often too small to measure accurately the $T(\lambda)$ of a coating such as a narrowband bandpass or an immersed coating (see Fig. 1-196). To use the language of §1.6.6, the *angle shift* of the bandpass makes its measured bandwidth larger than it actually is. The cone angle can also make the measured slope (see §1.5.2.2) of an edge filter larger than it actually is.

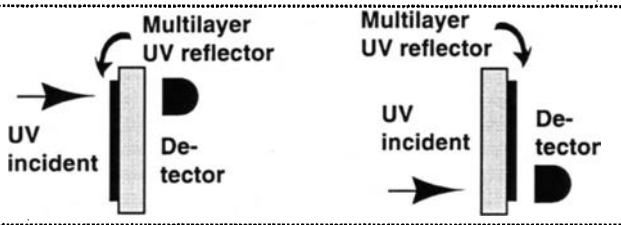


Fig. 10-40 A fluorescent slab is irradiated by UV. *(Left panel)* The UV reflecting coating permits minimal UV from entering the slab. The fluorescence is almost nonexistent. *(Right panel)* The UV penetrates the slab; both the fluorescent light is collected by a detector.

10.4.3.7 Fluorescence

This is a subtle effect and is best illustrated by an example. Consider a slab of "ionically colored glass," such as that produced by Schott or Hoya. Figure 10-40 shows that this slab is coated on one side with a mirror that reflects strongly from 300 nm to 400 nm. The wavelength of the flux incident upon the sample is 350 nm, which is measured in two orientations — with the coated side *towards* the detector and *away* from the detector. The spectrophotometer outputs a different transmittance for these orientations. However, §1.2.2.1 indicates that $T(\lambda)$ is *independent* of the direction of propagation. How can this happen?

The answer lies in the fact that the 350 nm UV flux causes the ionically colored glass to fluoresce strongly, as, for example, at a wavelength of 510 nm. This fluorescent light is subsequently collected by the detector in the spectrophotometer. In one orientation, the 350 nm radiation is reflected by the coating and is not permitted to reach glass. The fluorescence is minimal. When the sides are reversed, the 350 nm radiation easily penetrates the glass and produces fluorescence, which the coating readily transmits and the detector collects.

10.4.3.8 Temperature and humidity effects

§1.5.5 mentions that the wavelength of an edge filter is shifted by humidity. The wavelength shift of a bandpass filter with temperature is enumerated in §1.6.4.2. The optical properties of a coating should be measured in conditions similar to the environment in which the coating will operate. If it is used in a normal ambient environment, it should be allowed to stabilize a few hours after removal from the coating chamber before it is inserted into a spectrophotometer.

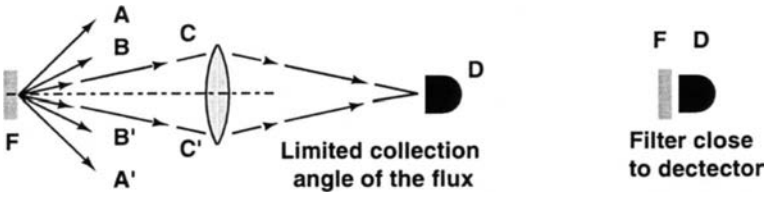


Fig. 10-41 Bandpass filter *F* is irradiated. (*Left*) Due to lens, the flux scattered over a limited cone angle is collected by the detector *D*. (*Right*) Considerably more scattered flux is collected by the detector because of closeness of the filter and detector.

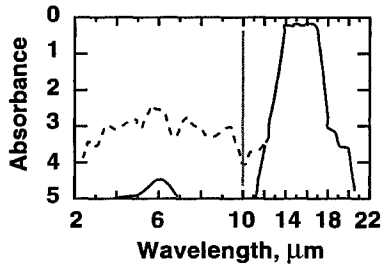


Fig. 10-42 Transmittance (on an absorbance scale) of a multilayer filter as measured on a conventional spectrometer (*solid*) and in the "filter close to detector" configuration (*dashed*) shown in Fig. 10-41. There is a change in scale of the wavelength at 10 μm. After [Stierwalt 91].

10.4.3.9 Proximity effect

Figure 10-41 shows a configuration in which this effect is observed [Stierwalt 91]. There is an IR multilayer bandpass that is irradiated by a radiation source (on its left). This source is not shown in Fig. 10-41. The flux transmitted by the bandpass is relayed to a detector *D* by a lens.

The arrows in Fig. 10-41 indicate the relative amount of scatter emerging from the bandpass at various angles. There is not much scattered flux at angles up to 15°, which is the cone over which the lens gathers the light and relays it to the detector. A consequence is that the more intense scattered radiation at higher angles (represented by arrows *A'*, *B'*, *A* and *B* in Fig. 10-41) does not reach the detector. The rejection in the offband spectral region is adequate, as shown by the *solid* $T(\lambda)$ curve in Fig. 10-42.

The right panel of Fig. 10-41 shows the configuration in which the detector is moved in proximity to the bandpass. This allows the detector to collect the scattered flux. The $T(\lambda)$ curve in Fig. 10-42 shows that the scattered flux has degraded the offband rejection by a factor of 100 or greater.

10.4.4 Absorption

When the radiant absorption *A* is relatively large, such as 10%, *R* and *T* are measured on a spectrophotometer and Eq. (1-3a) is solved for *A*.

There are many ways of measuring lower levels of absorption. One technique is to heat a laser mirror by reflecting a laser beam from it. A thermocouple next to the mirror measures its temperature increase. For a given laser power *W*, the *initial* rate of rise of temperature $\mathcal{T}$ (above ambient) is approximately

$$\frac{d\mathcal{T}}{dt} \approx \frac{WA}{C} \qquad (10\text{-}13e)$$

where *C* is the heat capacity (units of J/K) of the mirror. The rate of rise and *W* are measured, which allows the foregoing to be solved for *A*. This technique measures routinely the absorption of argon ion laser mirrors to levels of 0.04%.

10.4.5 Loss

The *loss* (see §1.2.2.3) includes any degradation due to flux that is absorbed, scattered or transmitted. The apparatus in Fig. 10-43 measures such loss [Deacon 92]. A short pulse of a few ns duration from a dye laser irradiates a cavity consisting of two confocal spherical mirrors. The mirrors are coated with dielectric reflectors.

The incident pulse fills the cavity. That radiation then leaks from the cavity in a series of pulses. The temporal spacing between them is related to the mirror separation. The pulse amplitude is an exponential decay, which is related to the loss. An ultrafast silicon photodiode measures the time constant of this decay.

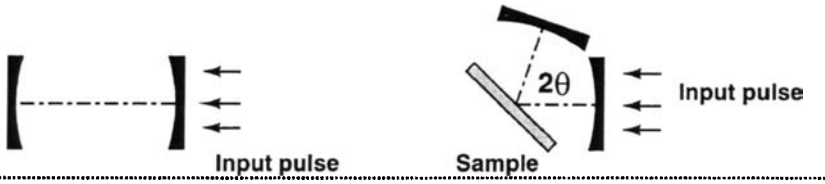


Fig. 10-43 (*Left*) An incident laser pulse resonates in a cavity consisting of two confocal mirrors. (*Right*) The cavity is folded to include a plano sample. The mirror diameters and angle 2θ are exaggerated for purposes of illustration.

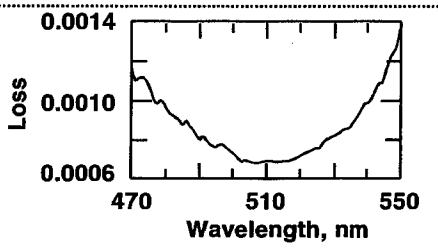


Fig. 10-44 Measured loss of an argon ion laser mirror using the apparatus shown in Fig. 10-43.

The cavity is then folded (*right* panel of Fig. 10-43) and a plano mirror with additional loss is inserted. The incidence angle θ is typically about 10° to 15°. The decay is then remeasured and the loss of the plano mirror is determined by subtraction. The wavelength of the dye laser can be varied. Figure 10-44 depicts the loss of an output mirror for an argon ion laser. The minimum of this curve corresponds to the wavelength of maximum reflectance of a quarterwave stack (see §5.2.4.3.1).

10.4.6 Other optical tests

The differential phase shift upon reflection may be determined with an ellipsometer. Bennett [85a, 85b, 85c] describes instruments that measure both the scatter of a coating and its surface microtopography. The optical figure of the coated part is often measured in an interferometer, as discussed in §10.7.

10.5 Performance of coatings — their non-optical characteristics

The tests in this section quantify the ability of a coated part to operate in a hostile environment, such as high humidity. The adhesion of the coating and its ability to withstand mechanical abrasion are also important. Most of these tests are conducted according to Military Standard Specifications. These are listed in the references under MIL. The Society of Automotive Engineers has embodied many of the MIL specifications.

10.5.1 Adhesion

The "Scotch tape test" is an unsophisticated (although widely used) way of testing the adhesion of a coating. Cellulose tape coated with an adhesive is placed on the surface, rubbed and removed slowly. The older type of tape should be used — the tape that was in vogue in the 1940s when this specification was devised. The newer (more transparent) type of tape does not adhere as tenaciously to the surface because the physical thickness of the adhesive is less. Details of this test appear in §4.6.8.1 of MIL-F-48616.

10.5.2 Moderate abrasion

This is the "cheesecloth rub" test. The surface is given 50 strokes with a pad of cheesecloth, the area of which is 9.5 mm by 6.4 mm. The force on the pad is 0.45 kg. After it has been rubbed, the surface is cleaned and light is reflected from it. It passes if the rub produced no visually apparent defects.

10.5.3 Severe abrasion

This "eraser rub" test is similar to the test of §10.5.2. It was devised in the 1940s and shows every indication that it will endure. Twenty strokes (back and forth) are applied to the coating with 0.91 kg of force (see section 4.4.7 of MIL-C-14806A). The eraser is special because it contains only pumice and rubber [Summers 85].

An extraneous comment is that the mechanical friction between the eraser and the coating is important. Any lubrication helps. Suppose that an antireflection coating fails the eraser rub. The next step is to anoint the surface of the antireflection coating with lanolin and then wipe it clean. It could easily pass abrasion the second time because of the ultrathin layer of grease on its surface. Such procedures cannot be used in the IR, of course, because of the absorption of organic materials in that spectral region.

10.5.4 Humidity

There are several ways to conduct the humidity test.

10.5.4.1 Constant humidity and temperature

According to MIL-C-14806A, the part is immersed in at least 95% humidity at 49 °C for 48 hours.

10.5.4.2 Humidity — cycled temperature

It is curious that often a coating will not degrade if is placed in a humid atmosphere at *constant* temperature, whereas the cycling of the temperature will hasten its entry to Valhalla. MIL-C-14806A specifies a temperature ranging from 28 °C to 71 °C and a humidity level of 90%.

10.5.5 Salt fog

The coating is immersed in a concentrated aqueous solution of common table salt for 24 hours.

10.5.6 Temperature shock

Temperature cycling tests the adhesion of the coating and also whether the coated substrate will withstand thermal shock. An example appears in section 3.11.3 of MIL-C-14806. The temperature is cycled between 71 °C and –54 °C in a time period of five minutes. There are IR coatings that function at even lower temperatures.

10.5.7 Specialized tests — fungus and sand erosion

The foregoing tests can be performed relatively easily on one's own premises. The tests in this paragraph are best performed by a special laboratory. In one pass through the gauntlet, they attempt to grow a fungus on the coating. Having failed to convert it into Roquefort, they pummel it in a miniature Mohave sandstorm, which tests its abrasion resistance like nothing else can.

10.5.8 Sequence of the tests

Each of the MIL specifications defines the *sequence* of operations. An example is that the transmittance is measured *after* the part is coated. It may be required that the part is again subjected to humidity and subsequently remeasured.

10.5.9 Other analytical methods

There are many methods of determining the structure of a coating as well as the chemical constituents of each layer. Microfractographs, as shown in Figs. 3-20 and 3-21, are produced by replication, shadowing with a heavy metal and transmission electron microscopy. Auger electron analysis, secondary ion mass spectroscopy or Rutherford backscatter can determine the chemical composition of each layer in a stack. X-ray or electron diffractometry furnishes information on the crystalline structure of the films. Bartella [85] describes how at least ten analytical tools found the exact composition and thickness of a multilayer containing both metallic and dielectric films.

10.6 Phase relations in multilayers

10.6.1 Introduction

The starting point of the discussion is the multilayer depicted in Fig. 10-45. The phase shift upon reflection δ [see Eq. (2-18b] associated with its amplitude reflection coefficient $\hat{\rho}$ ([as obtained, for example, from Eq. (2-33e)] is measured from its obverse side. It is also possible to define $\hat{\rho}'$ (and the corresponding phase shift δ') as the reflection coefficient from the reverse side (i.e., substrate side) of the stack. The amplitude transmission coefficient τ appears in §2.4.2.4.1.

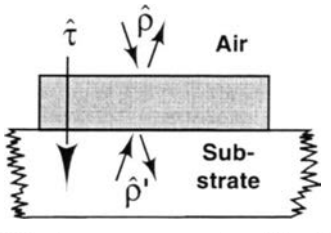


Fig. 10-45 A multilayer manifests an amplitude reflection $\hat{\rho}$ from its obverse side, $\hat{\rho}'$ from its substrate side and amplitude transmission coefficient τ.

10.6.1.1 Relationships between the phases

§2.2.5.4 establishes that the symbols *r* and *t* are reserved for the reflection and transmission coefficients at an interface, whereas $\hat{\rho}$ and $\hat{\tau}$ refer to a multilayer. Figure 10-45 illustrates that $\hat{\rho}$ and $\hat{\rho}'$ are measured at *different* physical locations. $\hat{\rho}$ is determined at the surface of the top layer of the stack, whereas $\hat{\rho}'$ is measured inside the substrate at the interface between it and the last layer. Given that Eq. (2-6g) obtains at an interface, it is *not* true, in general, that

$$\hat{\rho} = -\hat{\rho}' \ . \qquad (10\text{-}14a)$$

The transmission coefficient from right-to-left could also be defined as $\hat{\tau}'$. An equation analogous to Eq. (2-6h) is

$$\hat{\tau}\,\hat{\tau}' + \hat{\rho}^2 = 1 \ . \qquad (10\text{-}14b)$$

It is also not necessarily valid.

The amplitude transmission coefficient τ appears in Eq. (2-22f). On page 52 of Tikhonravov [92], it is proven that the phase retardance ν (i.e., arg of τ for "left to right" propagation) is equal to the phase retardance ν' (i.e., arg of $\hat{\tau}'$ for "right to left" propagation), provided the incident medium and substrate are nonabsorbing. The foregoing is true for an absorbing or nonabsorbing multilayer.

A relationship [Ou 89] for a nonabsorbing multilayer that relates the transmissive phase retardance ν (which is the arg of τ)to the reflective phase shifts is

$$\delta + \delta' + 2\,\nu = \pm\,\pi \qquad \text{(10-14c)}$$

where the arithmetic is modulo 2π. Several examples in subsequent sections illustrate the importance of this relationship.

10.6.2 Transmissive phase retardance

The phase retardance is, to a first approximation, related to the "single-pass path difference" as a ray propagates through a coating containing ℓ layers

$$PD]_{single\text{-}pass} \approx \sum_{i=1}^{\ell} h_i\, n_i \cos\theta_i \quad . \qquad \text{(10-14d)}$$

When 2π times the foregoing is divided by the wavelength (or multiplied by wave number), the phase retardance (in radians) is approximately

$$\nu \approx 2\,\pi\,(PD) / \lambda \approx 2\,\pi\,(PD)\,\sigma \quad . \qquad \text{(10-14e)}$$

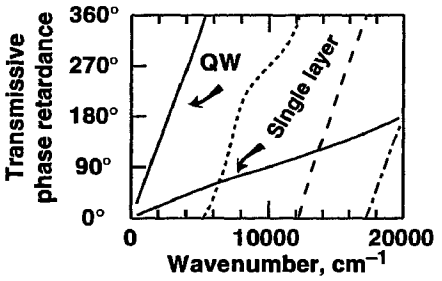


Fig. 10-46 Transmissive phase retardance vs wave number of

air **H** *glass*

and a quarterwave stack (curves with greater slope)

air **H** (**L H**)^6 *glass*

where the refractive indices of *air*, *glass*, **L** and **H** are 1.00, 1.52, 1.45 and 2.32, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$ at λ_0 of 1.00 μm. (*Short dash*): add 360° to phase. (*Long dash*): add 720° to phase. (*Long-short dash*): add 1080° to phase.

The foregoing is valid, to the extent that multiple reflections between the layers may be neglected. The important fact is that: the thicker the coating, the greater is the change of ν with respect to frequency (i.e., wave number). This is illustrated in Fig. 10-46, which shows the transmissive phase retardance of a single layer and a quarterwave stack that is seven times as thick. A final comment is that part of the curve in Fig. 10-46 is purely fictitious, if it were to be reproduced experimentally. The reason is that the oxide materials of such a coating are absorbing at wave numbers less than 3000 cm^{-1}.

10.6.2.1 Physical interpretation of retardance

An illustration is the phase retardance for a two-layer coating. The matrix in Eq. (2-22b) is written for each of the two layers. A matrix similar to Eq. (2-22c) is written for the interface (at the substrate) represented by the Fresnel amplitude coefficient r_2. After multiplication of the matrices [see Eq. (2-26e)], the amplitude transmission coefficient τ for this coating is

$$\tau = \frac{t_0\, t_1\, t_2 \exp[-j(\beta_1+\beta_2)\,]}{1 + r_0\, r_1 \exp(-j2\beta_1) + r_1\, r_2 \exp(-j2\beta_2) + r_0\, r_2 \exp[-j(2\beta_1+2\beta_2)]} \quad . \qquad (10\text{-}14f)$$

The "$(\beta_1+\beta_2)$" in the numerator represents the sum of the "single-pass" retardances through each of the two layers and is identical with Eq. (10-14e).

The denominator of Eq. (10-14f) is attributed to the multiple bounces of the waves within the layers. These are sometimes termed the *zigzag* reflections. As shown in Fig. 10-47, the term "$r_0\, r_1 \exp(-j2\beta_1)$" represents the retardance produced by multiple bounces of the waves in the upper layer and "$r_1\, r_2 \exp(-j2\beta_2)$" the zigzag reflections within the second layer. The term "$r_0\, r_2 \exp[-j(2\beta_1+2\beta_2)]$" represents the retardance produced by the multiple bounces of the waves between the upper interface (i.e., the incident medium) and the lower interface (at the substrate). Each of these multiple bounces adds retardance to the denominator of Eq. (10-14f).

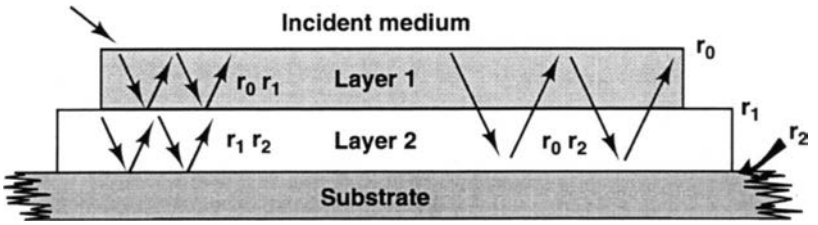


Fig. 10-47 Multiple reflections of beams within the layers of a two-layer coating. r_0, r_1 and r_2 are the amplitude reflection coefficients at the upper, inner and lower interfaces, respectively.

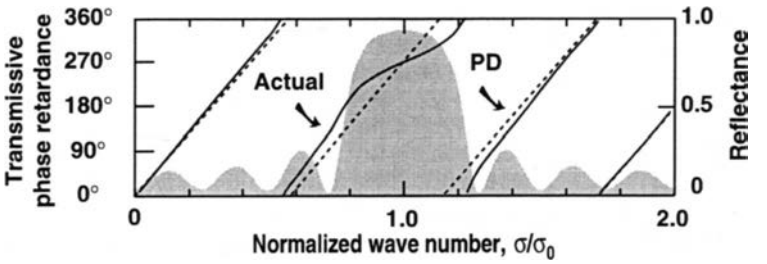


Fig. 10-48 Transmissive phase retardance (*solid*), approximate retardance *PD* [see Eq. (10-14e)], (*dashed*) and reflectance (*shaded*, left ordinate) of the seven layer stack captioned in Fig. 10-46.

In other words, for a stack with ℓ layers, the "single-pass" retardance is a valid approximation, as long as the additional retardance contributed by the multiple reflections of the waves between the layers is not significant.

Figure 10-48 plots the exact transmissive phase retardance (*solid*), whereas the *dashed curve* is the approximate retardance from Eq. (10-14e). At selected wave numbers, there is at least a 30° spread, whereas at a few others, the curves coincide.

10.6.3 Reflective phase shift

Curves of reflective phase shift of specific designs appear in Figs. 2-29, 2-30, 2-31, 2-32, 5-97, 5-58 and 5-99. This section takes a broader view of the phase and relates it to the radiant reflectance through the Kramers-Kronig relationship.

10.6.3.1 Dispersion theory

Dispersion theory, also known as Kramers-Kronig analysis, establishes a relationship between the real and imaginary parts of a complex variable, both of which are functions of frequency (i.e., wave number σ). Schatz [63], for example, started with reflectance measurements and used Kramers-Kronig to find the optical constants in the spectral region near a resonance absorption. The treatment here follows Lichten [85] and Stern[63].

Consider a multilayer whose amplitude reflection coefficient $\hat{\rho}$ as a function of wave number is known. The Naperian logarithm of $\hat{\rho}$ [see Eq. (2-18b)] is

$$\hat{w} \equiv A(\sigma) + j\,B(\sigma) \equiv \ln|\hat{\rho}| + j\,\delta \tag{10-15a}$$

where $\hat{w}$ is a complex variable. The result of dispersion theory is that the phase shift at a wave number σ_0 is

$$\delta(\sigma_0) = KKPV + B \tag{10-15b}$$

where *KKPV* represents the "Kramers-Kronig principal value"

$$KKPV = \frac{2\sigma_0}{\pi} \int_0^{\infty} \frac{A(\sigma)}{\sigma^2 - \sigma_0^2}\, d\sigma \tag{10-15c}$$

and B is the Blaschke factor

$$B = 2 \sum_{i=-\infty}^{+\infty} \arctan[\,(2\pi\sigma_0 - \sigma_i)/\mu_i\,] \;. \tag{10-15d}$$

In Eq. (10-15d), σ_i is proportional to the wave number at which a zero exists in the complex frequency plane. μ_i is the imaginary part of the frequency at which the zero occurs. Tikhonravov [97a] observed that the aforementioned σ_i values occur in pairs, symmetrical about the vertical axis in the complex frequency plane. Using this fact, Eq. (10-15d) may be modified to

$$B = 2 \sum_{i=0}^{\mathcal{N}} \arctan\left[\frac{2\pi\mu_i\sigma_i}{(2\pi\sigma_0)^2 - |\sigma_i|^2}\right] \;. \tag{10-15e}$$

The number of zeros $\mathcal{N}$ is either zero or infinite, provided the layers are optically homogeneous. Figure 10 in Tikhonravov [97a] depicts inhomogeneous layers that manifest a finite number of zeros in the lower complex frequency plane.

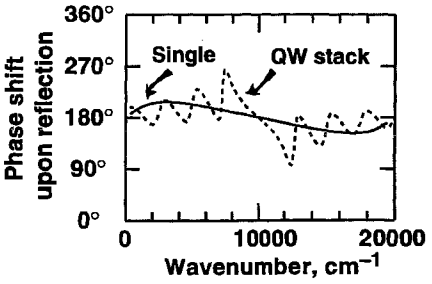


Fig. 10-49 Phase shift upon reflection (from the air side) vs wave number of the two coatings captioned in Fig. 10-46.

The aforementioned parameters (i.e., μ_i and σ_i) depend upon the design of the multilayer, as well as the refractive index of the incident and emergent media. For some designs, all zeros are in the upper half complex frequency plane and, consequently, the Blaschke factors in Eq. (10-15d) do not exist.

Tikhonravov [97a] cites earlier publications that explain how to compute μ_i and σ_i above. A commercial software package [Optilayer 97] enables the user to display the zeros in the complex frequency plane, which, in turn, quantifies μ_i and σ_i .

Because the wave numbers σ_i are approximately equally spaced for optically homogeneous layers (see Tikhonravov [97a]), the effect of the Blaschke factor is to add 360° to the phase shift on a quasi-periodic basis.

10.6.3.1.1 Minimal phase systems — examples

Minimal phase is borrowed from the theory of feedback, i.e., control theory (see Solodovnikov [60]). It refers to a system in which the phase shift is bounded, which implies that the Blaschke factors [in Eq. (10-15d)] do not exist.

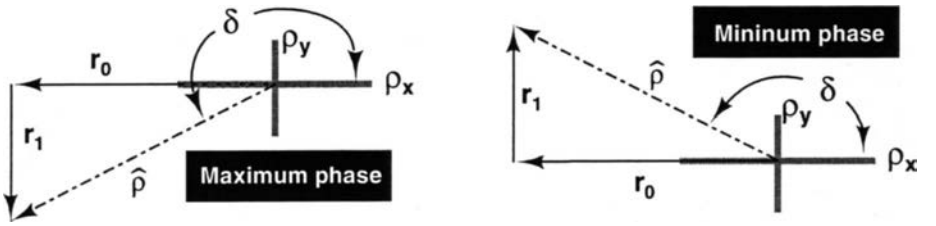


Fig. 10-50 Vector addition of amplitudes computation of the amplitude reflection coefficient $\hat{\rho}$ (*long-short dash*) and phase shift upon reflection δ from the air side of a single nonabsorbing layer of index 2.32 deposited upon a substrate of index 1.52.

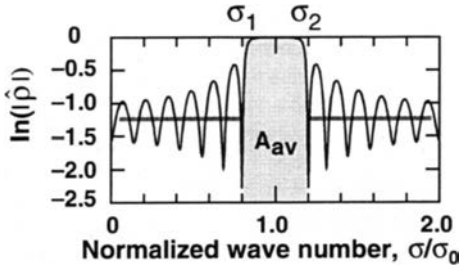


Fig. 10-51 Naperian log of the magnitude of the amplitude reflection coefficient $\hat{\rho}$ vs σ / σ_0 of the quarterwave stack

air **H** (**L H**)^6 *glass*

where the refractive indices of *air*, *glass*, **L** and **H** are 1.00, 1.52, 1.45 and 2.32, respectively. The optical thickness of **L** and **H** is $\lambda_0/4$. A_{av} is the approximate mean of the $\ln(|\rho|)$ in the spectral region 0 to σ_1.

Figure 10-49 depicts the reflected phase shift (from the air side) of a single high index layer on a glass substrate. A vector diagram aids in visualizing its behavior as a minimal phase system. Figure 10-50 shows that the length of the vector r_0 at the outer interface exceeds that of r_1 — the vector representing the Fresnel amplitude reflection coefficient at the film-glass interface. Ab intuito, the reflective phase shift δ is bounded between the approximate values shown in 10-49.

Also graphed in Fig. 10-49 is the phase of a seven layer quarterwave stack

air (**H L**)^3 **H** *glass* .

It is evident that this coating represents a minimal phase system.

10.6.3.1.2 Quarterwave stack — slope of the phase

The foregoing shows that if the magnitude of the amplitude reflection coefficient is known at all frequencies, then the phase shift upon reflection may be computed. An additional proviso is that the Blaschke factors are absent.

As an example, Fig. 5-97 shows the phase shift of a quarterwave stack. The goal is to find the *slope* of the phase shift at the stop band center. Fig. 10-51 plots the natural log of the magnitude of the amplitude reflection coefficient. The stop band occurs at a normalized frequency g of 1.0, which corresponds to a wave number of σ_0. The wave numbers σ_1 and σ_2 delineate the edges of the stop band. The ordinate is essentially zero within the stop band spectral region, which is a consequence of the fact that the reflectance is close to 100%. At frequencies outside of the stop band, the ordinate attains an average value A_{av}, as shown in Fig. 10-51. Following Lichten, the integral in Eq. (10-15c) is split into two parts — a summation of frequencies below and above the stop band edges

$$\delta(\sigma_0) \approx \frac{2\sigma_0}{\pi} \int_0^{\sigma_1} \frac{A_{av}}{\sigma^2 - \sigma_0^2} d\sigma + \frac{2\sigma_0}{\pi} \int_{\sigma_2}^{\infty} \frac{A_{av}}{\sigma^2 - \sigma_0^2} d\sigma \qquad (10\text{-}16a)$$

which evaluates to

$$\delta(\sigma_0) \approx \frac{A_{av}}{\pi} \left[\ln\left(\frac{\sigma_0 - \sigma_1}{\sigma_0 + \sigma_1} \right) + \ln\left(\frac{\sigma_2 + \sigma_0}{\sigma_2 - \sigma_0} \right) \right] \qquad (10\text{-}16b)$$

and equals zero when

$$\sigma_0 = \sqrt{\sigma_1 \, \sigma_2} \quad . \qquad (10\text{-}16c)$$

Differentiating Eq. (10-16b) with respect to σ and evaluating it at $\sigma = \sigma_0$, where σ_0 is given in Eq. (10-16c), leads to

$$\left[\frac{d\delta}{d\sigma} \right]_{\sigma = \sigma_0} \approx \frac{4 A_{av}}{\pi (\sigma_2 - \sigma_1)} \quad . \qquad (10\text{-}16d)$$

Although Lichten [85] finds a numerical value for the foregoing, no attempt is made to do so here. Its principal value is to show the functional dependence — namely, that the slope is inversely proportional to the stop band width. The slope in Eq. (10-16d) is negative because A_{av} is negative.

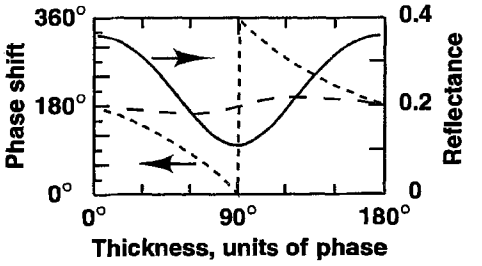


Fig. 10-52 Layers of index n_a and n_b are deposited upon a nonabsorbing substrate of index 4.00. The incident medium is air. (*Solid* curve — *right* ordinate): radiant reflectance vs phase thickness β for both layers. Phase shift upon reflection (*left* ordinate): n_a = 1.414 (*short dash*) and n_b = 2.828 (*long dash*).

It is serendipitous that the Blaschke contribution is absent for this seven layer stack that starts with an H and ends with an H. There are many ways of spawning a plethora of non-zero Blaschke factor — a few are: The addition of a L layer to the top of the stack, measurement of the reflective phase shift from the substrate side of the stack or replacing the top quarterwave thickness H layer with an eighthwave layer.

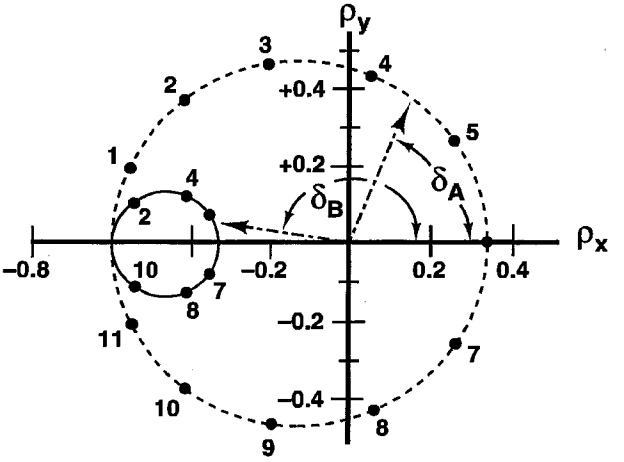


Fig. 10-53 Caption to Fig. 10-52 obtains, with the exception that the amplitude reflection coefficient in air is plotted parametrically at 30° increments of β. $n_a = 1.414$ — outer circle (*dashed*); $n_b = 2.828$ — smaller circle (*solid*).

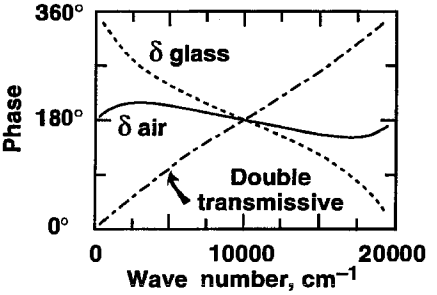


Fig. 10-54 Versus wave number, double the transmissive phase retardance v, phase shift upon reflection from the air side δ and phase shift upon reflection from the substrate side δ' of the single-layer coating captioned in Fig. 10-46.

10.6.3.1.3 Uniqueness of the $R(\sigma)$

The properties of the simple one-layer coatings below show that, although their spectral reflectances are identical, the phase shifts upon reflection are different.

Single layers of indices $n_a = 1.414$ and $n_b = 2.828$ are deposited on a substrate of index 4.00. It can be shown that layer indices that satisfy

$$n_a\, n_b = n_0\, n_s \qquad \text{(10-16e)}$$

manifest identical reflectance curves, as shown in Fig. 10-52. The phase shift curves are decidedly different, however. The plot of the δ(β) in Fig. 10-52 shows that the reflective phase shift of the higher index layer never wanders far from $\delta = 180°$. This is a minimal

phase system. The $\delta(\beta)$ of the lower index layer steadily decreases from 180° to zero as $\beta \rightarrow 90°$ and changes by a total of 360° as β increases from zero to 180° . The Blaschke factors are evidently present.

The reason for this behavior is evident from the parametric plot of the external amplitude reflection coefficient in Fig. 10-53. The circle for the lower index film encloses the origin, whereas that of the higher index layer does not. Of necessity, the $\delta(\beta)$ of the lower index film must change by 360° each time the layer thickness increases by $\lambda/2$.

Tikhonravov [97b] illustrates the non-uniqueness of the reflectance. Four nonabsorbing coatings are shown that have different refractive index profiles.

10.6.3.1.4 Non-minimal phase system

Figure 10-54 graphs the reflective phase from the substrate (i.e., glass) side of the single-layer coating mentioned in §10.6.3.1.1. The reason that the reflective phase is unbounded is evident from the vector diagram in Fig. 10-55. The length of the vector r_0 at the outer interface is less than the Fresnel amplitude reflection at the film-glass interface. The consequence is that the locus of the net amplitude reflection coefficient r encompasses the origin, similar to the behavior of the lower index film in §10.6.3.1.3.

10.6.4 Constraints upon the phase shifts

For a nonabsorbing stack, it is true that

> The phase shift upon reflection from the obverse side may be minimal phase, or the phase shift from the reverse side may be minimal phase, but both simultaneously may not be minimal phase.

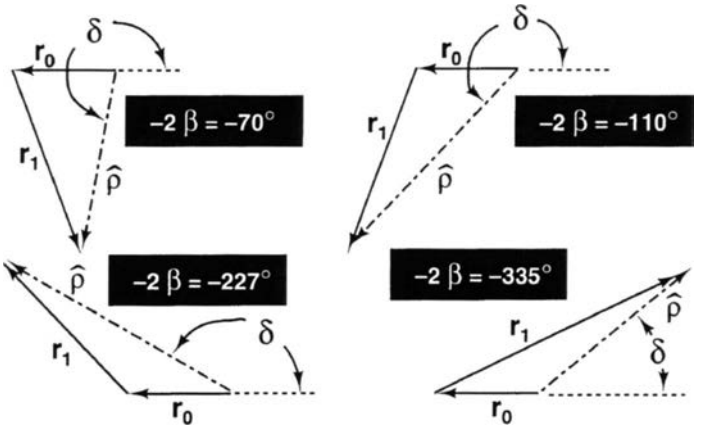


Fig. 10-55 The caption to Fig. 10-50 obtains, with the exception that *glass* is the incident medium and *air* is the substrate.

Minimal phase above is used in the same sense as in §10.6.3.1.1. Proof follows from Eq. (10-14c). Because the transmissive phase retardance is not bounded but, rather, increases monotonically with wave number, either δ or δ' (or possibly both) must be unbounded.

The phases — transmissive and reflective — of a single layer shown in Fig. 10-54 serve as an example. The sums at several wavelengths illustrate Eq. (10-14c). The

reflectance phase shift of the coating is "minimal phase" from its air side, whereas from its substrate the reflective phase shift is unbounded.

A consequence is that if, for example, there is to be a 90° phase difference between the light reflected from the front side and from the reverse side of a multilayer beamdivider, there is no choice but to make the coating absorbing. The design of such a coating is a balancing act of no mean proportions. The trick is not only to produce the correct reflective and transmissive phase shifts, but in addition to produce equal R_p and R_s. The absorption should also be minimized. There have been numerous studies of such coatings. A few are: Zelenchuk [85], Rabinovitch [94] and Monzón [97].

10.6.4.1 Reference of the phase to another surface

In the analysis of the effect of thickness errors upon the phase, it is advantageous to reference the phase to the substrate. At a particular wavelength λ, the phase shift upon reflection δ_0 at the top surface of a multilayer is referenced to the phase δ_e at the surface of the emergent medium — the substrate — by

$$\delta_e = \delta_0 + 4\,\pi\, n_0\, h_t \cos\theta_0 / \lambda \qquad \text{(10-17a)}$$

where h_t is the metric thickness of the entire multilayer. n_0 and θ_0 are, respectively, the refractive index and incidence angle of the incident medium.

The foregoing equation can be proven in several ways. The simplest is to substitute the phase thickness β_t for β_1 in Eqs. (2-21f) and (2-21g). The ratio is

$$\hat{\rho}(0) = \frac{\hat{E}_{0,i}^{(-)}}{\hat{E}_{0,i}^{(+)}} = \frac{\hat{E}^{(-)}(h_t)}{\hat{E}^{(+)}(h_t)}\ \exp(-j\,2\beta_t)\ . \qquad \text{(10-17b)}$$

Substituting $\hat{\rho}(0) \equiv |\hat{\rho}_0|\ \exp(j\,\delta_0)$ and a similar expression for the electric field ratio on the right of the equation leads to Eq. (10-17a).

10.7 The influence of a coating upon a transmitted or reflected wavefront

10.7.1 Introduction

The transmitted and reflected wavefronts are altered by a coating. These effects are analyzed in §10.7.2 and §10.7.3, respectively.

10.7.2 Transmitted wavefront

The starting point of the discussion is Eq. (2-4c), which describes mathematically a plane wave propagating in the +z direction. The normal to the wavefront makes an angle θ_0 with respect to the surface. The first exponential term in Eq. (2-4c) is the retardance *parallel* to the surface and the second term is the retardance normal to the surface. These are discussed below.

10.7.2.1 Transmitted wavefront — retardance normal to the surface

The second term in Eq. (2-4c) is the retardance for that component of the wave propagating in a direction normal to the surface. For a multilayer, this is quantified by the parameter ν in Eq. (2-22f) — the transmissive phase retardance that is, for example, graphed in Fig. 10-46.

There are two approximations for this transmissive phase retardance. The first is the path difference *PD* in Eq. (10-14d). In another approximation, the optical thickness OT [see Eq. (2-37a)] of the multilayer is found from the *PD* in Eq. (10-14d) at normal

incidence. The multilayer of total metric thickness h_t is replaced with a fictitious layer of the same metric thickness and mean refractive n_m, where

$$n_m \equiv \mathrm{OT}/\, h_t \ . \qquad \text{(10-17c)}$$

At non-normal incidence, the ray is assumed to refract in this fictitious layer using Snell's law.

The second approximation has been used by lens designers from time immemorial — the retardance of the coating is neglected.

10.7.2.2 Transmitted wavefront — retardance parallel to the surface

Consider a coating with ℓ layers. At a given angle θ_0 in the incident medium, angle θ_i of the refracted ray in the ith layer and lateral displacement y_i of the ray in that layer

$$\beta]_{\,y\ component} = 2\pi\,\lambda^{-1} \sum_{i=1}^{\ell} y_i\, n_i \sin\theta_i \ = \ 2\pi\,\lambda^{-1}\, n_0 \sin\theta_0 \sum_{i=1}^{\ell} h_i \tan\theta_i \quad . \text{(10-18a)}$$

Snell's law [i.e., Eq. (2-3f)] makes it possible to remove the "$n_i \sin\theta_i$" from the sum. The ray displacement y of the total stack is depicted in Fig. 10-56.

The total phase retardance is the algebraic sum of ν in Eq. (2-22f) and (10-18a).

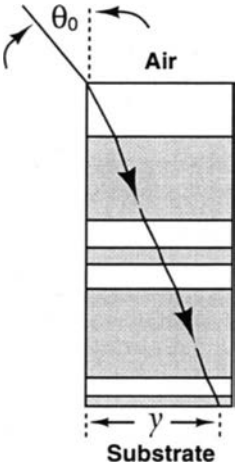


Fig. 10-56 Lateral displacement y of a ray as it propagates from air to the substrate in the eight-layer antireflection coating captioned in Fig. 4-2. The angle of incidence is 40° in air.

10.7.2.3 Ray tracing — exact and approximate retardance

Suppose that a multilayer antireflection coating is deposited upon a lens. If the antireflection coating design is known, its retardance may be computed and included in the ray trace. For many years, however, lens designers have ignored the retardance of the optical coating. What error is introduced by this neglect?

Three cases are considered: (1) An exact computation — the ray is traced through each layer and ν is computed exactly using matrix theory (see §2.4.4.1); (2) the multilayer is approximated by a fictitious medium that manifests the mean refractive index of the multilayer; and (3) the phase retardance of the multilayer is ignored. Each case is discussed below.

An antireflection for the visible spectral region and an IR bandpass serve as examples.

10.7.2.4 Antireflection coating — exact computation

The multilayer antireflection coating captioned in Fig. 4-2 is the first example. For a ray incident at 40° in air, its physical displacement parallel to the substrate surface is found, as well as its phase retardance in the z direction.

The matrix theory in §2.4.4.3 was used to compute the transmissive phase depicted in Fig. 10-57. This Fig. shows that the retardance is 720° , i.e., two waves, at 514 nm. The "single-pass" retardance from Eq. (10-14e) is also shown. Only in the blue spectral region does it differ significantly from the exact.

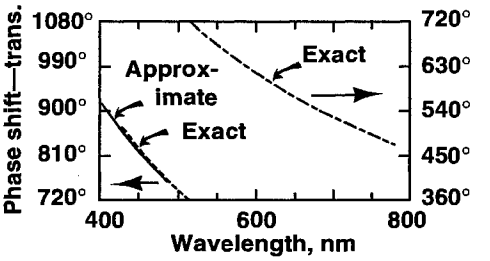


Fig. 10-57 Exact transmissive phase retardance [see Eq. (2-22f)] of the eight-layer antireflection coating captioned in Fig. 4-2. The right ordinate pertains to the *long-short dash* curve and the left ordinate to the *short-dash* curve. The approximate retardance (*solid*) is from Eq. (10-14e).

Consider next the performance when the ray is incident upon the multilayer at 40°. Table 10-4 lists the angle of the refracted ray in each layer. The product of the tangent of this angle and the metric thickness of that layer gives the displacement of the ray, as shown in Fig. 10-56. The conclusion is that the ray is displaced laterally a distance of 0.24 μm. A separate calculation shows that at 514 nm the transmissive phase retardance for p polarization is 667° , which is smaller than the value cited above of 720° at normal incidence.

10.7.2.5 Antireflection coating — approximate retardance

Table 10-4 shows that this antireflection coating is 4.0960 QWOTs thick and that its metric thickness is 0.58631 μm, which from Eq. (2-47i) implies that its mean index is

$$n_m = 4.0960/(4 \times 0.58631) = 1.7465 \ . \qquad (10\text{-}18b)$$

A calculation shows, for example, that at 514 nm the phase retardance produced by a single transversal of this fictitious layer is 717° , which is 3° smaller than the exact value in Eq. (10-14e).

The final task is to compute the lateral displacement y of the ray by assuming that it refracts in a medium of index of 1.7465 [from Eq. (10-18b)]

$$y \approx 0.5863 \tan\{\arcsin[\sin(40°)/1.7465]\} = 0.2321 \ \mu m \qquad (10\text{-}18c)$$

which is about 4% smaller than the exact value of 0.24 μm cited in the final column of Table 10-4.

10.7.2.6 Antireflection coating — neglect of the retardance

A question is: what are the consequences if a ray is traced through air (of refractive index 1.0003) and the retardance of the coating is neglected? An approximation for the error ε

is obtained by using the average index

$$\varepsilon \approx h_t\,(n_m - 1.0003)/\lambda = \quad 0.438\ \mu m/\ \lambda \qquad (10\text{-}18d)$$

Table 10-4 Optical properties of the layers of an antireflection coating

Layer	QWOT	N_i	h_i	θ_i	$\cos\theta_i$	$n_i\ h_i\ \cos\theta_i$	y_i
1	0.5400 μm	1.38	0.09783 μm	27.76°	0.8849	0.1195 μm	0.0515 μm
2	1.1800 μm	1.95	0.15128 μm	19.25°	0.9441	0.27850 μm	0.0528 μm
3	0.2900 μm	1.45	0.05000 μm	26.31°	0.8964	0.06500 μm	0.0247 μm
4	0.2380 μm	1.95	0.03051 μm	19.25°	0.9441	0.05617 μm	0.0107 μm
5	0.2560 μm	1.45	0.04414 μm	26.31°	0.8964	0.05737 μm	0.0218 μm
6	1.2600 μm	1.95	0.16154 μm	19.25°	0.9441	0.2974 μm	0.0564 μm
7	0.1910 μm	1.45	0.03293 μm	26.31°	0.8964	0.0428 μm	0.0163 μm
8	0.1410 μm	1.95	0.01808 μm	19.25°	0.9441	0.03329 μm	0.0063 μm
Totals	4.0960 μm		0.58631 μm			0.95000 μm	0.2405 μm

Refractive index n_i and metric thickness h_i of each of the eight layers of the antireflection coating captioned in Fig. 4-2. θ_i is the angle of a refracted ray and y_i is its displacement in a direction parallel to the surface of the substrate. The angle of incidence is 40° in air.

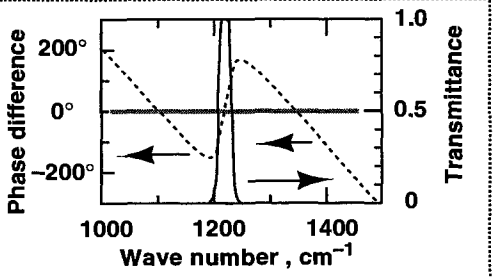


Fig. 10-58 $T(\sigma)$ (*right ordinate*) and phase difference (*left ordinate* — see text for explanation) of

air 0.353**H** 0.251**L** 0.353**H** **L** **H** 2**L** **H** (**L** **H**)^4 2**L** **H** (**L** **H**)^4 2**L** **H** **L** 1.215**H** 0.514**L** 0.215**H** *germanium*

where *air*, **L**, **H** and *germanium* have refractive indices of 1.00, 2.22, 4.10 and 4.00, respectively. The optical thicknesses of **L** and **H** are λ_0 /4 at λ_0 of 8.2 μm.

In an imaging system that functions in the blue part of the spectrum, this neglect of the coating introduces a focal shift of a full wave.

10.7.2.7 Bandpass — normal incidence

An all-dielectric bandpass that transmits at 8.2 μm in the IR was designed using the procedure of §7.7.3. As shown in the caption to Fig. 10-58, this bandpass contains 31 layers of germanium and zinc sulfide. Its retardance, which is 63.34 μm, appears in Table 10-5. From Eq. (10-18d), the error ε introduced by substituting this coating for an "air" path is

$$\varepsilon = 63.34\ \mu m - 1.0003 \times 22.55\ \mu m = 40.8\ \mu m \qquad (10\text{-}18e)$$

and amounts to 5 waves at 8.2 μm. It would to be difficult to ignore such a large focal shift. That is, the *edge thickness* of the component on which the coating is deposited should be shortened by 40.8 μm to compensate for the retardance introduced by the bandpass.

How closely does the optical thickness of this coating approximate the actual retardance? Figure 10-58 graphs its transmittance. Also plotted is the *difference* between the exact phase retardance and the approximate value in Eq. (10-14e). At the wavelengths in the passband where the transmittance is close to 50%, the difference is close to 100° — in excess of one-quarter of a wave. The conclusion is that, in a precise optical system, the matrix method should be used to compute the retardance.

Table 10-5 Optical properties of the layers of a bandpass

Material	h_i	n_i	OT	θ_i	$\cos\theta_i$	$h_i\, n_i \cos\theta_i$	$h_i \tan\theta_i$
ZnS	15.48 μm	2.22	34.37 μm	8.9	0.988	33.96 μm	2.41 μm
Ge	7.06 μm	4.10	28.98 μm	4.8	0.997	28.88 μm	0.59 μm
Totals	22.55 μm		63.34 μm			62.84 μm	3.0 μm

Optical thickness OT of the bandpass captioned in Fig. 10-58. The meaning of the parameters h_i, n_i and θ_i appear in the caption to Table 10-4, with the exception that θ_i is computed for an angle of incidence in air of 20°.

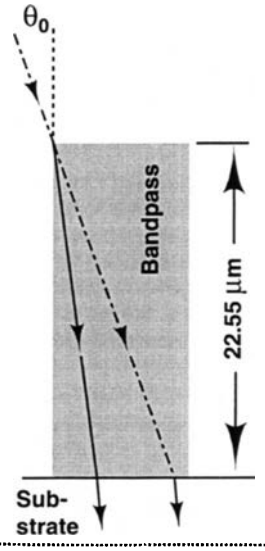


Fig. 10-59 Path of a ray traced in air (*long-dash*) at $\theta_0 = 20°$. The metric thickness of the bandpass is 22.55 μm. The *solid* line is where the ray emerges into the substrate when the bandpass is present.

10.7.2.8 Bandpass — nonnormal incidence

Suppose that an optical systems engineer assumes that the ray propagates through air and neglects the retardance of the multilayer whose metric thickness is 22.55 μm. What is the error introduced when a ray incident at 20° in air propagates through it? Because the index of air is 1.0003, the *assumed* path difference between normal incidence and 20° is

$$PD]_{assumed} = 22.55\ \mu m \times 1.0003 \times (1 - \cos 20°)$$
$$= 1.356\ \mu m\ . \qquad (10\text{-}19a)$$

The actual path difference is approximated by subtracting the optical thickness from the final column in Table 10-5. The error ε is

$$\varepsilon \approx 1.356\ \mu m - (63.34\ \mu m - 62.84\ \mu m)$$
$$= 0.86\ \mu m \qquad (10\text{-}19b)$$

or about a tenth wave at 8.2 μm. The foregoing may, or may not, be significant, depending upon the precision of the imaging system.

Finally, Fig. 10-59 shows that when the ray propagates through air, it is displaced laterally by 22.55 μm × tan20° = 8.2 μm. Table 10-5 shows that, due to the relatively large refractive index of the layers in the bandpass, the ray refracted by a substantial amount. The result is that it is laterally displaced by 3 μm, which is 36% of aforementioned 8.2 μm.

10.7.2.9 Effect of thickness nonuniformity

Suppose that the manufacturing process produces a metric thickness of a coating that is thinner by an amount δh at its periphery, as depicted in Fig. 10-60. For purposes of

illustration, let us assume that the coating is 2% thinner. Table 10-5 shows that the optical thickness of this bandpass is 63.34 μm. 98% of the retardance occurs in this coating. However, the wave also propagates through air for a distance that is 2% of the total metric thickness.

The path difference in retardance *PD* between this thinner portion of the coating and the "correct" portion is

$$PD = 63.34\ \mu m - (0.98 \times 63.34\ \mu m + 0.02 \times 22.55\ \mu m) = 0.82\ \mu m\ . \qquad (10\text{-}19c)$$

This is a tenth wave at 8.2 μm.

10.7.3 Phase of the reflected beam

Suppose that a multilayer reflector is used in a coherent optical system, such as an interferometer, imaging system or laser cavity. It is only coincidental that the top surface of the multilayer is the surface that the wavefront reflects from. The pseudo-surface that it reflects from may be modeled as the position of the node of the standing wave (see §2.3.2). Depending upon the value of the phase shift, Fig. 2-27 shows that this surface may be physically above or below the top surface of the multilayer.

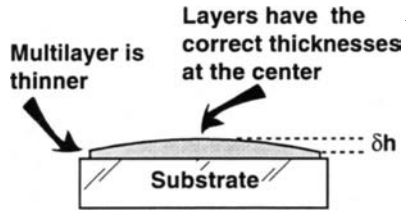


Fig. 10-60 The metric thickness of a multilayer coating is smaller by an amount δh at its periphery. Vertical dimensions are exaggerated for purposes of illustration.

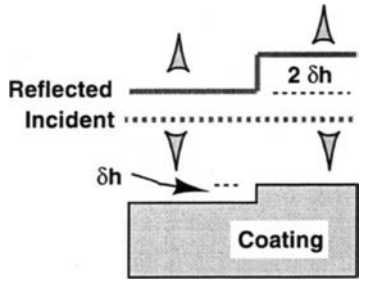


Fig. 10-61 The dotted line represents a plane wavefront of a wave incident upon a reflector. A step of thickness δh in its surface produces a step retardance of $2\,\delta h$ in the reflected wavefront.

Thickness changes in a multilayer may cause the pseudo-surface to move a distance that may exceed, or be less than, the actual physical deformation of the multilayer. Several examples illustrate this phenomenon.

10.7.3.1 Effects of a thickness change — opaque overcoat

A configuration that is simple to model occurs when an overcoat of opaque metal is applied to the multilayer, as shown in Fig. 10-61. This Fig. shows that a plane wavefront *I* is incident upon a step (of vertical distance of δh in its thickness) in the coating. There is a displacement of $2\,\delta h$ in the reflected wavefront *W* because the wave must travel both *into* and *away from* the thinner section of the coating The phase distortion ε produced by this thickness change is

$$\varepsilon = 2\,\delta h/\lambda \text{ waves}\ . \qquad (10\text{-}20a)$$

The foregoing equation is valid because the phase shift upon reflection is determined only by the optical constants of the metal and, as long as the metal is opaque, it is independent of its thickness. The same is not true of a relatively transparent multilayer.

10.7.3.2 Effects of a thickness change — quarterwave stack

A reflector at 1.06 μm consisting of thirteen layer quarterwave layers serves as the first example. It is composed of materials of refractive indices 1.46 and 2.30. Its design is

air $(\mathbf{H}\,\mathbf{L})^6\,\mathbf{H}$ *glass* .

Suppose that the layers are 1% thinner at the periphery of the substrate than at the center, as depicted in Fig. 10-60. Several scenarios are considered: (1) The multilayer is overcoated with an opaque metal and its figure measured with an interferometer; (2) the multilayer itself is measured at 1.06 μm; (3) the multilayer is measured at 633 nm. The ensuing discussion shows that the contour of the surface, as observed in an interferogram, may be quite different, depending upon the wavelength at which it is observed.

10.7.3.2.1 Overcoated multilayer

Given the QWOTs [see Eq. (2-47i)], the first step is to find the metric thickness of the multilayer. This is quite close to 1.90 μm. At 1.06 μm, Eq. (10-20a) predicts that a 1% thickness nonuniformity produces a phase distortion ε of

$$\varepsilon = 2 \times (1.90\ \mu\text{m}) \times 0.01 \times (1.06\ \mu\text{m})^{-1} = 0.04 \text{ waves} = 13^\circ \qquad (10\text{-}20b)$$

for the multilayer overcoated with the opaque metal.

10.7.3.2.2 Methods of computation

There are at least two methods of finding the phase distortion of the multilayer due to a thickness change. The first is to compute the phase shift upon reflection of a stack that is 1% thinner and has an *additional* layer of air on top that has a thickness of 0.019 μm (which is 1% of 1.90 μm). In this way, the reference planes for the "correct" and the "thinner" stacks are the same. The other method is to have a computer code correct the phase to the substrate using Eq. (10-17a).

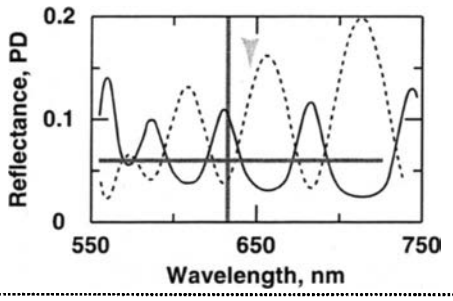


Fig. 10-62 For the reflected wave, $R(\lambda)$ (*dashed*) and phase distortion (in waves — *solid*) due to a simultaneous 1% change in thickness of each layer. The design is

air $(\mathbf{H}\,\mathbf{L})^6\,\mathbf{H}$ *glass*

where the indices of **L** and *glass* are 1.46 and 1.52, respectively. The refractive index of **H** is given by Eq. (2-2c) with $A = 5$ and $B = 0.3$. The optical thickness of each layer is $\lambda_0/4$ at $\lambda_0 = 1.06$ μm.
The vertical line is at 633 nm and the horizontal line is explained in the text.

10.7.3.2.3 Phase distortion of reflector at 1.06 μm

A computer code using the characteristic matrix method was employed to find the phase shift upon reflection δ. The result is that at the stack center of 1.06 μm, an interferometer would observe a phase distortion of 10.2° for this 1% thickness change. How does this compare to the multilayer overcoated with an opaque metal? Equation (10-20b) shows that this phase distortion of 13° is slightly greater than the 10.2° cited above.

The problem is that interferometers that function at 1.06 μm are rare. What is the phase distortion when the interferogram is made at 633 nm? The answer is that it may be worse (by nearly a factor of two) or better than it actually is, depending upon the wavelength centering of the coating.

10.7.3.2.4 Phase distortion at 633 nm — centered coating

A calculation similar to that in Eq. (10-20b) shows that an ε of 21° (i.e., 0.058 wave) is measured at 633 nm when the thickness change is 1% and the sample is overcoated with an opaque metal. This 0.058 wave is represented by the horizontal line in Fig. 10-62. It is larger than the value at 1.06 μm because of the smaller wavelength.

Figure 10-62 plots both the reflectance and phase distortion (due to 1% thickness change) vs wavelength. An interferogram is made at 633 nm of a perfectly centered quarterwave stack. "Perfectly centered" means that each layer has an optical thickness of precisely a quarterwave at 1.06 μm. Figure 10-62 shows that the phase distortion is slightly in excess of a tenth wave, which is nearly double the 0.058 wave cited above.

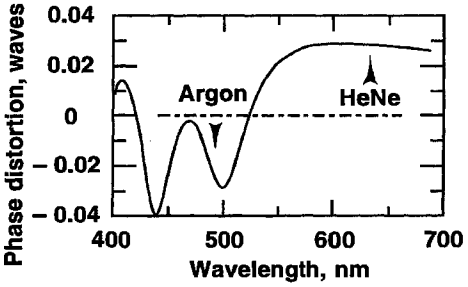


Fig. 10-63 At fixed reference plane, phase change (in waves) due to a 1% change in thickness of each layer. The design is

air 92.85**H** 103.03**L** 68.08**H** 125.61**L** 66.67**H** 103.39**L** 58.187**H** 96.365**L** 53.527**H** 74.46**L** 38.74**H** 85.13**L** 54.18**H** 68.05**L** 24.72**H** *glass*

where the indices of **H**, **L** and *glass* are 2.30, 1.35 and 1.52, respectively. Layer thicknesses are metric in nm

10.7.3.2.5 Phase distortion at 633 nm — noncentered coating

Suppose, however, there is a flaw in the coating process so that the entire multilayer is centered "short" by 2%. That is, each layer (on the substrate, at both its physical center and periphery) has an optical thickness that is 2% thin. The net result is to shift the reflectance curve and "phase distortion curve" in Fig. 10-62 to shorter wavelengths. The conclusion is that the phase distortion at 633 nm is about 0.03 wave, which is one-third the value of 0.1 wave cited in §10.7.3.2.4. The conclusion is:

As measured on an interferometer at *other than* the design wavelength, the figure of an optical surface with multilayer coatings *with nonuniform thicknesses* may be fallacious. (Latin — *fallio*: "I deceive.")

10.7.3.3 Phase distortion — broadband reflector

This fifteen-layer reflector covers the visible spectrum. Its $R(\lambda)$ is similar to curve (d) Fig. 1-127. What phase distortion ε is observed when the coating (metric thickness = 1.11 μm) is 1% thin? Eq. (10-20a) shows that ε is 0.035 wave at 633 nm.

Figure 10-63 depicts ε, which is the sensitivity (in waves) of the phase shift upon reflection to a 1% thickness change. A positive value implies that the phase distortion is the *same* as that produced by an opaque metal overcoat. Figure 10-63 shows that this occurs at the wavelength of a HeNe red laser. As noted by earlier workers [Bouwhuis 62, Ramsay 67, Ciddor 68], the ε may be negative in certain spectral regions (as, for example, λ = 488 nm). This implies that ε is in the *opposite* direction to the thickness change.

Suppose that a plano substrate is coated with the multilayer captioned in Fig. 10-64 and that the thickness of each layer is thinner by 1% at its periphery. This thinner edge would be observed in an interferometer if the multilayer were overcoated with an opaque metal. The *short dashed* curve in Fig. 10-64 is the contour that an interferometer at 633 nm would measure, which is slightly less convex than the surface itself. If the wavelength were moved to 488 nm, the surface would appear concave, as indicated by the *long-short dash* curve in Fig. 10-64. It follows that there are other wavelengths at which the coated substrate (the upper surface of which is, in fact, slightly convex) would appear "plano" because of the insensitivity of ε to thickness changes.

10.7.3.4 Phase distortion of an antireflection coating

There are no examples in this section, but a warning is proffered that if any coating has quite low reflectance in any spectral region, then the variation of the phase shift upon reflection with wavelength is quite large in that region. A consequence is that the phase distortion caused by thickness changes is also quite large.

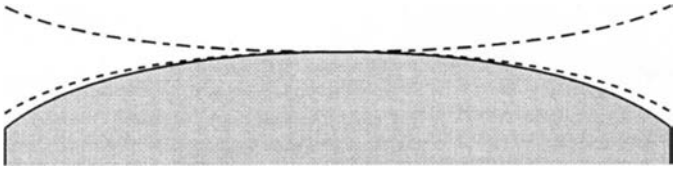


Fig. 10-64 The upper surface (*shaded*) of the coating captioned in Fig. 10-63 is curved due to an identical fractional decrease of thickness in each of the layers. Dashed curves represent the figure as measured with an interferometer at 633 nm (*short dash*) and 488 nm (*long-short dash*).