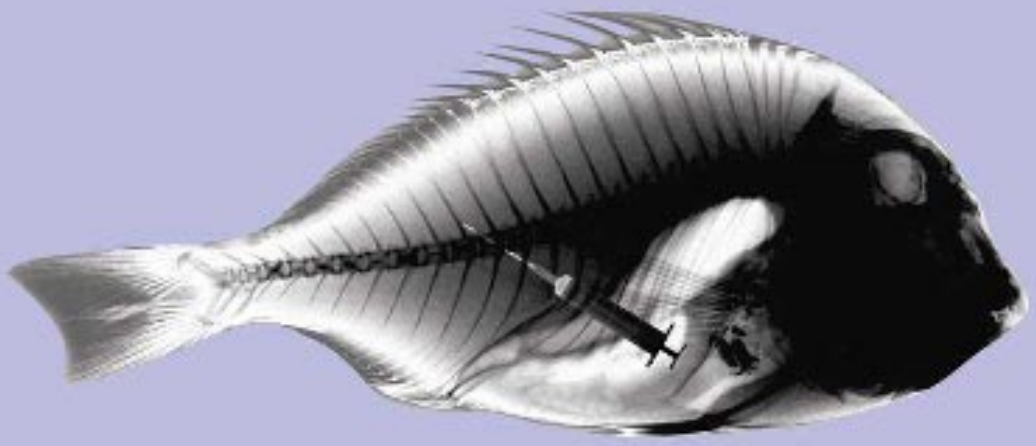
Nondestructive Evaluation

Theory, Techniques, and Applications



edited by

Peter J. Shull

Nondestructive Evaluation

Theory, Techniques, and Applications

edited by

Peter J. Shull

The Pennsylvania State University
Altoona, Pennsylvania



MARCEL DEKKER, INC. NEW YORK • BASEL

Copyright © 2001 by Marcel Dekker, Inc. All Rights Reserved.

Library of Congress Cataloging-in-Publication Data
Nondestructive evaluation: theory, techniques, and applications/edited by Peter J. Shull.
p. cm. – (Mechanical engineering series of reference books; 142)
ISBN: 0-8247-8872-9 (alk. paper)
1. Non-destructive testing. I. Shull, Peter J. II. Series.

TA417.2.N6523 2002
620.1'127–dc21

2002019334

This book is printed on acid-free paper.

Headquarters
Marcel Dekker, Inc.
270 Madison Avenue, New York, NY 10016
tel: 212-696-9000; fax: 212-685-4540

Eastern Hemisphere Distribution
Marcel Dekker AG
Hutgasse 4, Postfach 812, CH-4001 Basel, Switzerland
tel: 41-61-261-8482; fax: 41-61-261-8896

World Wide Web
http://www.dekker.com

The publisher offers discounts on this book when ordered in bulk quantities. For more information, write to Special Sales/Professional Marketing at the headquarters address above.

Copyright © 2002 by Marcel Dekker, Inc. All Rights Reserved.

Neither this book nor any part may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, microfilming, and recording, or by any information storage and retrieval system, without permission in writing from the publisher.

Current printing (last digit):
10 9 8 7 6 5 4 3 2 1

PRINTED IN THE UNITED STATES OF AMERICA

MECHANICAL ENGINEERING

A Series of Textbooks and Reference Books

Founding Editor

L. L. Faulkner

Columbus Division, Battelle Memorial Institute
and Department of Mechanical Engineering
The Ohio State University
Columbus, Ohio

1. *Spring Designer's Handbook*, Harold Carlson
2. *Computer-Aided Graphics and Design*, Daniel L. Ryan
3. *Lubrication Fundamentals*, J. George Wills
4. *Solar Engineering for Domestic Buildings*, William A. Himmelman
5. *Applied Engineering Mechanics: Statics and Dynamics*, G. Boothroyd and C. Poli
6. *Centrifugal Pump Clinic*, Igor J. Karassik
7. *Computer-Aided Kinetics for Machine Design*, Daniel L. Ryan
8. *Plastics Products Design Handbook, Part A: Materials and Components; Part B: Processes and Design for Processes*, edited by Edward Miller
9. *Turbomachinery: Basic Theory and Applications,* Earl Logan, Jr.
10. *Vibrations of Shells and Plates*, Werner Soedel
11. *Flat and Corrugated Diaphragm Design Handbook,* Mario Di Giovanni
12. *Practical Stress Analysis in Engineering Design,* Alexander Blake
13. *An Introduction to the Design and Behavior of Bolted Joints,* John H. Bickford
14. *Optimal Engineering Design: Principles and Applications,* James N. Siddall
15. *Spring Manufacturing Handbook,* Harold Carlson
16. *Industrial Noise Control: Fundamentals and Applications,* edited by Lewis H. Bell
17. *Gears and Their Vibration: A Basic Approach to Understanding Gear Noise,* J. Derek Smith
18. *Chains for Power Transmission and Material Handling: Design and Applications Handbook,* American Chain Association
19. *Corrosion and Corrosion Protection Handbook,* edited by Philip A. Schweitzer
20. *Gear Drive Systems: Design and Application,* Peter Lynwander
21. *Controlling In-Plant Airborne Contaminants: Systems Design and Calculations,* John D. Constance
22. *CAD/CAM Systems Planning and Implementation,* Charles S. Knox
23. *Probabilistic Engineering Design: Principles and Applications,* James N. Siddall
24. *Traction Drives: Selection and Application,* Frederick W. Heilich III and Eugene E. Shube
25. *Finite Element Methods: An Introduction,* Ronald L. Huston and Chris E. Passerello
26. *Mechanical Fastening of Plastics: An Engineering Handbook,* Brayton Lincoln, Kenneth J. Gomes, and James F. Braden
27. *Lubrication in Practice: Second Edition,* edited by W. S. Robertson
28. *Principles of Automated Drafting,* Daniel L. Ryan
29. *Practical Seal Design,* edited by Leonard J. Martini
30. *Engineering Documentation for CAD/CAM Applications,* Charles S. Knox
31. *Design Dimensioning with Computer Graphics Applications,* Jerome C. Lange
32. *Mechanism Analysis: Simplified Graphical and Analytical Techniques,* Lyndon O. Barton
33. *CAD/CAM Systems: Justification, Implementation, Productivity Measurement,* Edward J. Preston, George W. Crawford, and Mark E. Coticchia
34. *Steam Plant Calculations Manual,* V. Ganapathy
35. *Design Assurance for Engineers and Managers,* John A. Burgess
36. *Heat Transfer Fluids and Systems for Process and Energy Applications,* Jasbir Singh

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

37. *Potential Flows: Computer Graphic Solutions,* Robert H. Kirchhoff
38. *Computer-Aided Graphics and Design: Second Edition,* Daniel L. Ryan
39. *Electronically Controlled Proportional Valves: Selection and Application,* Michael J. Tonyan, edited by Tobi Goldoftas
40. *Pressure Gauge Handbook,* AMETEK, U.S. Gauge Division, edited by Philip W. Harland
41. *Fabric Filtration for Combustion Sources: Fundamentals and Basic Technology,* R. P. Donovan
42. *Design of Mechanical Joints,* Alexander Blake
43. *CAD/CAM Dictionary,* Edward J. Preston, George W. Crawford, and Mark E. Coticchia
44. *Machinery Adhesives for Locking, Retaining, and Sealing,* Girard S. Haviland
45. *Couplings and Joints: Design, Selection, and Application,* Jon R. Mancuso
46. *Shaft Alignment Handbook,* John Piotrowski
47. *BASIC Programs for Steam Plant Engineers: Boilers, Combustion, Fluid Flow, and Heat Transfer,* V. Ganapathy
48. *Solving Mechanical Design Problems with Computer Graphics,* Jerome C. Lange
49. *Plastics Gearing: Selection and Application,* Clifford E. Adams
50. *Clutches and Brakes: Design and Selection,* William C. Orthwein
51. *Transducers in Mechanical and Electronic Design,* Harry L. Trietley
52. *Metallurgical Applications of Shock-Wave and High-Strain-Rate Phenomena,* edited by Lawrence E. Murr, Karl P. Staudhammer, and Marc A. Meyers
53. *Magnesium Products Design,* Robert S. Busk
54. *How to Integrate CAD/CAM Systems: Management and Technology,* William D. Engelke
55. *Cam Design and Manufacture: Second Edition; with cam design software for the IBM PC and compatibles, disk included*, Preben W. Jensen
56. *Solid-State AC Motor Controls: Selection and Application,* Sylvester Campbell
57. *Fundamentals of Robotics,* David D. Ardayfio
58. *Belt Selection and Application for Engineers,* edited by Wallace D. Erickson
59. *Developing Three-Dimensional CAD Software with the IBM PC,* C. Stan Wei
60. *Organizing Data for CIM Applications,* Charles S. Knox, with contributions by Thomas C. Boos, Ross S. Culverhouse, and Paul F. Muchnicki
61. *Computer-Aided Simulation in Railway Dynamics,* by Rao V. Dukkipati and Joseph R. Amyot
62. *Fiber-Reinforced Composites: Materials, Manufacturing, and Design,* P. K. Mallick
63. *Photoelectric Sensors and Controls: Selection and Application,* Scott M. Juds
64. *Finite Element Analysis with Personal Computers,* Edward R. Champion, Jr., and J. Michael Ensminger
65. *Ultrasonics: Fundamentals, Technology, Applications: Second Edition, Revised and Expanded,* Dale Ensminger
66. *Applied Finite Element Modeling: Practical Problem Solving for Engineers,* Jeffrey M. Steele
67. *Measurement and Instrumentation in Engineering: Principles and Basic Laboratory Experiments,* Francis S. Tse and Ivan E. Morse
68. *Centrifugal Pump Clinic: Second Edition, Revised and Expanded,* Igor J. Karassik
69. *Practical Stress Analysis in Engineering Design: Second Edition, Revised and Expanded,* Alexander Blake
70. *An Introduction to the Design and Behavior of Bolted Joints: Second Edition, Revised and Expanded,* John H. Bickford
71. *High Vacuum Technology: A Practical Guide,* Marsbed H. Hablanian
72. *Pressure Sensors: Selection and Application,* Duane Tandeske
73. *Zinc Handbook: Properties, Processing, and Use in Design,* Frank Porter
74. *Thermal Fatigue of Metals,* Andrzej Weronski and Tadeusz Hejwowski
75. *Classical and Modern Mechanisms for Engineers and Inventors,* Preben W. Jensen
76. *Handbook of Electronic Package Design,* edited by Michael Pecht
77. *Shock-Wave and High-Strain-Rate Phenomena in Materials*, edited by Marc A. Meyers, Lawrence E. Murr, and Karl P. Staudhammer
78. *Industrial Refrigeration: Principles, Design and Applications*, P. C. Koelet
79. *Applied Combustion*, Eugene L. Keating
80. *Engine Oils and Automotive Lubrication*, edited by Wilfried J. Bartz

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

81. *Mechanism Analysis: Simplified and Graphical Techniques, Second Edition, Revised and Expanded*, Lyndon O. Barton
82. *Fundamental Fluid Mechanics for the Practicing Engineer*, James W. Murdock
83. *Fiber-Reinforced Composites: Materials, Manufacturing, and Design, Second Edition, Revised and Expanded*, P. K. Mallick
84. *Numerical Methods for Engineering Applications*, Edward R. Champion, Jr.
85. *Turbomachinery: Basic Theory and Applications, Second Edition, Revised and Expanded*, Earl Logan, Jr.
86. *Vibrations of Shells and Plates: Second Edition, Revised and Expanded*, Werner Soedel
87. *Steam Plant Calculations Manual: Second Edition, Revised and Expanded*, V. Ganapathy
88. *Industrial Noise Control: Fundamentals and Applications, Second Edition, Revised and Expanded*, Lewis H. Bell and Douglas H. Bell
89. *Finite Elements: Their Design and Performance*, Richard H. MacNeal
90. *Mechanical Properties of Polymers and Composites: Second Edition, Revised and Expanded*, Lawrence E. Nielsen and Robert F. Landel
91. *Mechanical Wear Prediction and Prevention*, Raymond G. Bayer
92. *Mechanical Power Transmission Components*, edited by David W. South and Jon R. Mancuso
93. *Handbook of Turbomachinery*, edited by Earl Logan, Jr.
94. *Engineering Documentation Control Practices and Procedures*, Ray E. Monahan
95. *Refractory Linings: Thermomechanical Design and Applications*, Charles A. Schacht
96. *Geometric Dimensioning and Tolerancing: Applications and Techniques for Use in Design, Manufacturing, and Inspection*, James D. Meadows
97. *An Introduction to the Design and Behavior of Bolted Joints: Third Edition, Revised and Expanded*, John H. Bickford
98. *Shaft Alignment Handbook: Second Edition, Revised and Expanded*, John Piotrowski
99. *Computer-Aided Design of Polymer-Matrix Composite Structures*, edited by S. V. Hoa
100. *Friction Science and Technology*, Peter J. Blau
101. *Introduction to Plastics and Composites: Mechanical Properties and Engineering Applications*, Edward Miller
102. *Practical Fracture Mechanics in Design*, Alexander Blake
103. *Pump Characteristics and Applications*, Michael W. Volk
104. *Optical Principles and Technology for Engineers*, James E. Stewart
105. *Optimizing the Shape of Mechanical Elements and Structures*, A. A. Seireg and Jorge Rodriguez
106. *Kinematics and Dynamics of Machinery*, Vladimír Stejskal and Michael Valá□ek
107. *Shaft Seals for Dynamic Applications*, Les Horve
108. *Reliability-Based Mechanical Design*, edited by Thomas A. Cruse
109. *Mechanical Fastening, Joining, and Assembly*, James A. Speck
110. *Turbomachinery Fluid Dynamics and Heat Transfer*, edited by Chunill Hah
111. *High-Vacuum Technology: A Practical Guide, Second Edition, Revised and Expanded*, Marsbed H. Hablanian
112. *Geometric Dimensioning and Tolerancing: Workbook and Answerbook*, James D. Meadows
113. *Handbook of Materials Selection for Engineering Applications*, edited by G. T. Murray
114. *Handbook of Thermoplastic Piping System Design*, Thomas Sixsmith and Reinhard Hanselka
115. *Practical Guide to Finite Elements: A Solid Mechanics Approach*, Steven M. Lepi
116. *Applied Computational Fluid Dynamics*, edited by Vijay K. Garg
117. *Fluid Sealing Technology*, Heinz K. Muller and Bernard S. Nau
118. *Friction and Lubrication in Mechanical Design*, A. A. Seireg
119. *Influence Functions and Matrices*, Yuri A. Melnikov
120. *Mechanical Analysis of Electronic Packaging Systems*, Stephen A. McKeown
121. *Couplings and Joints: Design, Selection, and Application, Second Edition, Revised and Expanded,* Jon R. Mancuso
122. *Thermodynamics: Processes and Applications*, Earl Logan, Jr.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

123. *Gear Noise and Vibration*, J. Derek Smith
124. *Practical Fluid Mechanics for Engineering Applications,* John J. Bloomer
125. *Handbook of Hydraulic Fluid Technology*, edited by George E. Totten
126. *Heat Exchanger Design Handbook*, T. Kuppan
127. *Designing for Product Sound Quality*, Richard H. Lyon
128. *Probability Applications in Mechanical Design*, Franklin E. Fisher and Joy R. Fisher
129. *Nickel Alloys,* edited by Ulrich Heubner
130. *Rotating Machinery Vibration: Problem Analysis and Troubleshooting*, Maurice L. Adams, Jr.
131. *Formulas for Dynamic Analysis*, Ronald Huston and C. Q. Liu
132. *Handbook of Machinery Dynamics*, Lynn L. Faulkner and Earl Logan, Jr.
133. *Rapid Prototyping Technology: Selection and Application*, Ken Cooper
134. *Reciprocating Machinery Dynamics: Design and Analysis*, Abdulla S. Rangwala
135. *Maintenance Excellence: Optimizing Equipment Life-Cycle Decisions,* edited by John D. Campbell and Andrew K. S. Jardine
136. *Practical Guide to Industrial Boiler Systems*, Ralph L. Vandagriff
137. *Lubrication Fundamentals: Second Edition, Revised and Expanded,* D. M. Pirro and A. A. Wessol
138. *Mechanical Life Cycle Handbook*: *Good Environmental Design and Manufacturing,* edited by Mahendra S. Hundal
139. *Micromachining of Engineering Materials,* edited by Joseph McGeough
140. *Control Strategies for Dynamic Systems: Design and Implementation,* John H. Lumkes, Jr.
141. *Practical Guide to Pressure Vessel Manufacturing,* Sunil Pullarcot
142. *Nondestructive Evaluation: Theory, Techniques, and Applications,* edited by Peter J. Shull
143. *Diesel Engine Engineering: Dynamics, Design, and Control,* Andrei Makartchouk
144. *Handbook of Machine Tool Analysis,* Ioan D. Marinescu, Constantin Ispas, and Dan Boboc
145. *Implementing Concurrent Engineering in Small Companies,* Susan Carlson Skalak
146. *Practical Guide to the Packaging of Electronics: Thermal and Mechanical Design and Analysis*, Ali Jamnia
147. *Bearing Design in Machinery: Engineering Tribology and Lubrication,* Avraham Harnoy
148. *Mechanical Reliability Improvement: Probability and Statistics for Experi-mental Testing,* R. E. Little
149. *Industrial Boilers and Heat Recovery Steam Generators: Design, Applications, and Calculations*, V. Ganapathy
150. *The CAD Guidebook: A Basic Manual for Understanding and Improving Computer-Aided Design,* Stephen J. Schoonmaker
151. *Industrial Noise Control and Acoustics,* Randall F. Barron
152. *Mechanical Properties of Engineering Materials,* Wolé Soboyejo
153. *Reliability Verification, Testing, and Analysis in Engineering Design*, Gary S. Wasserman
154. *Fundamental Mechanics of Fluids: Third Edition*, I. G. Currie

Additional Volumes in Preparation

HVAC Water Chillers and Cooling Towers: Fundamentals, Application, and Operations, Herbert W. Stanford III

Handbook of Turbomachinery: Second Edition, Revised and Expanded, Earl Logan, Jr., and Ramendra Roy

Progressing Cavity Pumps, Downhole Pumps, and Mudmotors, Lev Nelik

Gear Noise and Vibration: Second Edition, Revised and Expanded, J. Derek Smith

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Piping and Pipeline Engineering: *Design, Construction, Maintenance, Integrity, and Repair*, George A. Antaki

Turbomachinery: *Design and Theory*: Rama S. Gorla and Aijaz Ahmed Khan

Mechanical Engineering Software

Spring Design with an IBM PC, Al Dietrich

Mechanical Design Failure Analysis: With Failure Analysis System Software for the IBM PC, David G. Ullman

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016



To Kathleen, whose secret patience is boundless

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016



To those who have not quite mastered the finer details of Nondestructive Methods

Daniel Yurchak





Gabriel Pio-Jesus



Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

To Robert E. Green
For his lifelong contributions and dedication to the field of Nondestructive Testing and his tireless efforts on behalf of his students. Thank you.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016



Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Preface

The rapidly expanding role of nondestructive evaluation (NDE) methods in manufacturing, power, construction, and maintenance industries, as well as in basic research and development, has generated a large demand for practitioners, engineers, and scientists with knowledge of the subject. This text is intended to help respond to this demand by presenting NDE methods in a format appropriate for a broad range of audiences—undergraduate and first-year graduate engineering students, practicing engineers, technicians, and researchers. *Nondestructive Evaluation* presents current practices, common methods and equipment, applications, and the potential and limitations of current NDE methods, in addition to the fundamental physical principles underlying NDE.

Properly employed, NDE methods can have a dramatic effect on the cost and reliability of products. The methods can be used to evaluate prototype designs during product development, to provide feedback for process control

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

during manufacturing, and to inspect the final product prior to service. Additionally, NDE methods permit products to be inspected throughout their serviceable life to determine when to repair or replace a particular part. In today's economy, the concepts "repair or retire for cause" and "risk-informed inspection" are becoming increasingly important, as the United States and other countries are faced with aging infrastructure, aging airline fleets, aging petrochemical and power plants, and a multitude of other industries that try to use equipment well beyond its design lifetime. NDE offers a margin of safety for this equipment and gives users the means with which to determine when equipment must be repaired or retired.

This text allows the reader to understand both (a) the physical interaction between the material and the measurement method employed and (b) current practices, equipment, and techniques, by discussing both conventional and emerging NDE technologies.

TARGET AUDIENCES AND TEXT FORMAT

This book has been carefully crafted to function as a textbook appropriate for undergraduate (or first-year graduate) engineering courses in NDE methods and as a resource book to accomodate the varying needs of NDE users—from current practitioners working in the field, to practicing engineers or technicians, to researchers who want to try new and unconventional implementation of NDE methods.

Each chapter, written by experts in the field, contains

- An overview of an NDE method, citing brief examples of how the technology can be applied
- A complete discussion of the physics behind the method, beginning with the fundamental physical laws
- An explanation of inspection techniques and typical equipment
- A description of current applications of the method by top practitioners in the field

Introductory sections review background material typically presented in an undergraduate physics course. Advanced topics and optional sections have been added for readers who may be involved in work requiring detailed knowledge of a method. If you do not need to know the advanced material, you can skip the optional sections without any loss of continuity.

TEXTBOOK RESOURCE WEB SITE

As the saying goes, "At some point we need to shoot the engineers and make the thing." That is, at some point we needed to stop writing and publish this textbook.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

But because changes and improvements are always possible (and always welcome), I have developed a Web site (www.aa.psu.edu/NDE) to accommodate improvements dynamically. To this end, I encourage you to submit comments, questions, end-of-chapter questions, ideas, case studies, tricks of the trade, experiments, and topics to include (or delete), as well as a general evaluation of the book. Please also let me know if you find any mistakes.

ACKNOWLEDGMENTS

This book was motivated by a need for a general NDE textbook written by many experts, each of whom works every day with a specific NDE method. The resulting level of collaboration is in contrast to many NDE books written by one or two authors—experts in several NDE methods, perhaps, but not in all. In order to achieve such a multiplicity of perspectives, an enormous effort went into coordinating all of the many contributors' works to be consistent in format, breadth, and depth of the various NDE methods. As editor of this text, I would like to sincerely thank all the contributors for their efforts (which went significantly beyond what I believe many had initially expected)—it resulted in the high quality of this work.

Additionally, I would like to thank the many experts in the field of NDE who donated their valuable time to review the chapters of this text.

Peter J. Shull

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Reviewers

Brian Berhosky Foerster Instrument Inc., Pittsburgh, Pennsylvania

Jim Cox Zetec, Inc., Issaquah, Washington

Jeff Draper Krautkramer Branson, Lewistown, Pennsylvania

Dale Fitting Stress Sensors Inc., Golden, Colorado

Alice Flarend The Pennsylvania State University, Altoona, Pennsylvania

Robert E. Green CNDE, Johns Hopkins University, Baltimore, Maryland

Steve Groeninger Magnaflux Corporation, Glenview, Illinois

Donald Hagemaier The Boeing Company, Long Beach, California

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Vilma Holmgren Magnaflux Corporation, Glenview, Illinois

David Hurley INEEL, Idaho Falls, Idaho

Nathan Ida University of Akron, Akron, Ohio

Frank A. Iddings* Nuclear Science Center, Louisiana State University, Baton Rouge, Louisiana

Shridhar Nath GE Corporate R&D, Schenectady, New York

Jim Rieger Magnaflux Corporation, Glenview, Illinois

Sam Robinson Sherwin Corporation, Burlington, Kentucky

Stan Rokhlin The Ohio State University, Columbus, Ohio

Joe Rose The Pennsylvania State University, University Park, Pennsylvania

Ward Rummel D & W Enterprises, Littleton, Colorado

Dave Russell Russell NDE Systems, Edmonton, Alberta, Canada

Kermit Skeie Kermit Skeie Associates, San Dimas, California

Henry Stephens, Jr. EPRI NDE Center, Charlotte, North Carolina

Chris Stockhausen Magnaflux Corporation, Glenview, Illinois

Bernhard R. Tittmann The Pennsylvania State University, University Park, Pennsylvania

Jim Yukes Russell NDE Systems, Edmonton, Alberta, Canada

* Professor Emeritus

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Contents

Preface vii
Reviewers xi
Contributors xxiii

1 Introduction to NDE ***Peter J. Shull*** **1**

1.1 WHAT IS NONDESTRUCTIVE EVALUATION? 1
1.2 NDE: WHAT'S IN A NAME? 2
1.3 HOW IS NDE APPLIED? 3
1.4 UNDERSTANDING THE NDE CHOICES 7
1.5 HOW MUCH DO WE INSPECT? 8
1.5.1 Statistics 8

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

1.5.2 Consequences of Part Failure (Non-Safety-Critical Parts) 9
1.5.3 Larger Systems or Safety-Critical Parts 11
1.5.4 Retirement for Cause 11
1.5.5 Risk-Informed Inspection 11
1.6 RELIABILITY OF NDE 13
REFERENCES 15

2 Liquid Penetrant ***Frank A. Iddings and Peter J. Shull*** **17**

2.1 INTRODUCTION 17
2.1.1 Basic Technique 18
2.1.2 History 19
2.1.3 PT's Potential 20
2.1.4 Advantages/Disadvantages 21
2.2 FUNDAMENTALS 21
2.2.1 Fluid Flow 21
2.2.2 Illumination and Detection: The Eye's Response to Visible Spectrum Illumination 28
2.3 TECHNIQUES 31
2.3.1 Basic Method 31
2.3.2 Cleaning 32
2.3.3 Types of Penetrants 37
2.3.4 Temperature 39
2.3.5 Dwell Time 39
2.3.6 Removing Excess Penetrant 40
2.3.7 Types of Developers 45
2.3.8 Examination and Interpretation 47
2.3.9 Illumination 48
2.3.10 Final Cleaning 49
2.3.11 Specifications and Standards 49
2.4 APPLICATIONS 51
2.4.1 Fabrication Industries 51
2.4.2 Aerospace Industries 52
2.4.3 Petrochemical Plants 53
2.4.4 Automotive and Marine Manufacture and Maintenance 55
2.4.5 Electrical Power Industry 55
2.4.6 Other Applications 56
2.5 SUMMARY 57
PROBLEMS 58

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

GLOSSARY 59
VARIABLES 60
REFERENCES 60

3 Ultrasound *Peter J. Shull and Bernard R. Tittmann* **63**

3.1 INTRODUCTION 63
3.1.1 What Ultrasonic Waves Are, and How They Propagate 64
3.1.2 Technique Overview 65
3.1.3 History 66
3.1.4 Applications 68
3.1.5 Advantages and Disadvantages 69
3.2 THEORY 69
3.2.1 Introduction to Wave Propagation 70
3.2.2 Wave Motion and the Wave Equation 74
3.2.3 Specific Acoustic Impedance and Pressure 89
3.2.4 Reflection and Refraction at an Interface 92
3.2.5 Attenuation 107
3.2.6 Guided Waves 114
3.3 TRANSDUCERS FOR GENERATION AND DETECTION OF ULTRASONIC WAVES 120
3.3.1 Piezoelectric Transducers 120
3.3.2 Electromagnetic Acoustic Transducers (EMATs) 129
3.3.3 Laser (Optical) Generation and Detection of Ultrasound 133
3.3.4 Transducer Characteristics 138
3.4 INSPECTION PRINCIPLES 147
3.4.1 Measurements 147
3.4.2 Wave Generation 149
3.4.3 Transducer Configurations 153
3.4.4 Excitation Pulsers 156
3.4.5 Receiving the Signal 158
3.5 APPLICATIONS 161
3.5.1 Roll-by Inspection of Railroad Wheels 161
3.5.2 Ultrasonic Testing of Elevator Shafts and Axles 167
3.6 ADVANCED TOPICS (OPTIONAL) 170
3.6.1 Interference 170
3.6.2 Group Velocity and Phase Velocity 171
3.6.3 Method of Potentials 173
3.6.4 Derivation of Snell's Law: Slowness Curves 178

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

PROBLEMS 180
GLOSSARY 185
VARIABLES 189
REFERENCES 190

4 Magnetic Particle *Arthur Lindgren, Peter J. Shull, Kathy Joseph, and Donald Hagemaier* **193**

4.1 INTRODUCTION 193
4.1.1 Technique Overview 194
4.1.2 History 194
4.1.3 Advantages/Disadvantages 195
4.2 THEORY 195
4.2.1 Basic Magnetism 195
4.2.2 Magnetic Field 197
4.2.3 Magnetic Fields in Materials 200
4.2.4 Types of Magnetization 202
4.2.5 Magnetic Hysteresis in Ferromagnetic Materials 203
4.2.6 Leakage Field 207
4.3 TECHNIQUES/EQUIPMENT 207
4.3.1 Direction of Magnetization 209
4.3.2 Type of Magnetization Current 215
4.3.3 Magnetic Field Generation Equipment 222
4.3.4 Magnetic Particles 230
4.3.5 Demagnetizing the Sample 235
4.3.6 Condition of Parts 236
4.4 INSPECTION AIDS 237
4.4.1 Controlling Particle Suspension 237
4.4.2 Controlling Magnetization and System Performance 241
4.4.3 Avoiding Nonrelevant Indications 245
4.5 APPLICATIONS OF MPI 247
4.5.1 Steel Coil Springs 247
4.5.2 Welds 249
4.5.3 Railroad Wheels 250
4.6 SUMMARY 251
PROBLEMS 253
GLOSSARY 255
VARIABLES 258
REFERENCES 259

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

5 Eddy Current *Peter J. Shull* **261**

5.1 INTRODUCTION 261
5.1.1 Technical Overview 262
5.1.2 History 264
5.1.3 Potential of the Method 266
5.1.4 Advantages and Disadvantages 267
5.2 BASIC ELECTROMAGNETIC PRINCIPLES APPLIED TO EDDY CURRENT 268
5.2.1 Magnetic Induction (Self and Mutual) 269
5.2.2 An Eddy Current Example 274
5.2.3 Coil Impedance 279
5.2.4 Phasor Notation and Impedance 280
5.2.5 Eddy Current Density and Skin Depth 285
5.2.6 Maxwell's Equations and Skin Depth (Optional) 287
5.2.7 Impedance Plane Diagrams 291
5.2.8 Impedance Plane Analysis for Tubes and Rods (Optional) 301
5.3 TRANSDUCERS AND MEASUREMENT EQUIPMENT 307
5.3.1 EC Transducers (Probes) 307
5.3.2 Measurement Equipment 316
5.3.3 Actual EC Equipment 323
5.4 INSPECTION PRINCIPLES 325
5.4.1 Impedance Plane Measurement 325
5.4.2 Remote Field Eddy Current (Optional) 342
5.4.3 EC Calibration 350
5.5 APPLICATIONS 350
5.5.1 Remote Field Testing 350
5.5.2 Aerospace Applications 355
PROBLEMS 358
GLOSSARY 360
VARIABLES 365
REFERENCES 366

6. Acoustic Emission *William H. Prosser* **369**

6.1 INTRODUCTION 369
6.1.1 Technical Overview 369
6.1.2 Historical Perspective 372
6.1.3 Potential of Technique 375
6.2 FUNDAMENTALS 377

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

	6.2.1	Sources	377
	6.2.2	Wave Propagation	383
	6.2.3	Measurement	395
6.3	ANALYSIS TECHNIQUES		406
	6.3.1	AE "Activity"	406
	6.3.2	Feature Set Analysis	410
	6.3.3	Waveform Analysis	419
6.4	SPECIFIC APPLICATIONS		422
	6.4.1	Acoustic Emission Testing of Metal Pressure Vessels	423
	6.4.2	Impact Damage in Graphite/Epoxy Composite Pressure Vessels	425
	6.4.3	Materials Testing—TMC Study in Composites	429
	6.4.4	Fatigue Crack Detection in Aerospace Structures	435
6.5	SUMMARY		440
	PROBLEMS		441
	GLOSSARY		442
	VARIABLES		443
	REFERENCES		444

7	**Radiology**	***Harry E. Martz, Jr., Clinton M. Logan, and Peter J. Shull***	**447**
7.1	INTRODUCTION TO X-RAYS		448
	7.1.1	History of X-Rays	451
	7.1.2	Advantages and Disadvantages	457
7.2	RADIATION FUNDAMENTALS		459
	7.2.1	Introduction	459
	7.2.2	Fundamental Sources of Radiation	464
	7.2.3	Photon Radiation in NDE	466
	7.2.4	Non-Photons Radiation Used for NDE	479
	7.2.5	Radiography	480
	7.2.6	Computed Tomography (CT)	481
	7.2.7	Radiation Transport Modeling—Monte Carlo Method	484
	7.2.8	Dosimetry and Health Physics	484
7.3	EQUIPMENT FOR RADIATION TESTING		486
	7.3.1	Introduction	486
	7.3.2	Radiation Sources	487
	7.3.3	Radiation Detectors	498
	7.3.4	Collimators	513
	7.3.5	Stages and Manipulators	515

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

7.3.6 Radiological Systems 515
7.4 RADIOGRAPHIC TECHNIQUES 516
7.4.1 Practical Considerations 519
7.4.2 General Measurement (Detector) Limitations 520
7.4.3 Film 525
7.4.4 Digital Radiography and Computed Tomography 528
7.4.5 Special Techniques 550
7.5 SELECTED APPLICATIONS 554
7.5.1 Traditional Film (Projection) Radiography 555
7.5.2 Digital (Projection) Radiography and Computed Tomography 558
7.6 RADIATION SAFETY 571
7.6.1 Why Is Radiation Safety Required? 571
7.6.2 Radiation Safety Procedures 574
7.6.3 Responsibilities for Safety 577
PROBLEMS 578
GLOSSARY 579
VARIABLES 586
REFERENCES 589

8 Active Thermography ***Jane Maclachlan Spicer and Robert Osiander*** **597**

8.1 INTRODUCTION/BACKGROUND 597
8.1.1 Technique Overview 597
8.1.2 Historical Perspective 599
8.1.3 Potential of Active Thermography Techniques 602
8.2 BASICS OF HEAT DIFFUSION 604
8.2.1 Steady State Heat Flow 604
8.2.2 Conduction of Heat in One Dimension 607
8.2.3 Periodic Solutions to the Heat Conduction Equation 609
8.2.4 Pulsed Excitation 615
8.2.5 Step Heating 617
8.2.6 Other Extensions 619
8.3 TECHNIQUES 619
8.3.1 Introduction 619
8.3.2 Infrared Radiometry 622
8.3.3 Heating Sources 628
8.3.4 Making Active Thermography Measurements 631
8.4 SPECIFIC APPLICATIONS 636
8.4.1 Imaging Entrapped Water Under an Epoxy Coating 636
8.4.2 Detection of Carbon Fiber Contaminants 638

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

8.4.3 Using Induction Heating for NDE of Rebar in Concrete 638
PROBLEMS 641
VARIABLES 642
REFERENCES 642

9 Microwave ***Alfred J. Bahr, Reza Zoughi, and Nasser Qaddoumi*** **645**

9.1 INTRODUCTION 645
9.1.1 Technical Overview 646
9.1.2 Historical Perspective 647
9.1.3 Potential of the Technique 647
9.1.4 Advantages and Disadvantages 649
9.2 BACKGROUND 650
9.2.1 Material Parameters 651
9.2.2 Basic Electromagnetic Wave Concepts 656
9.3 MICROWAVE EQUIPMENT 664
9.4 MICROWAVE SENSORS/TECHNIQUES 672
9.4.1 Transmission Sensors 672
9.4.2 Reflection and Radar Sensors 673
9.4.3 Resonator Sensors 674
9.4.4 Radiometer Sensors 675
9.4.5 Imaging Sensors 677
9.4.6 General Remarks on Sensors 677
9.5 APPLICATIONS 678
9.5.1 Dielectric Material Characterization Using Filled Waveguides 678
9.5.2 Dielectric Material Characterization Using Transmission Measurements 692
9.5.3 Inspection of Layered Dielectric Composites Using Reflection Measurements 695
9.5.4 Microwave Inspection Using Near-Field Probes 703
9.5.5 Summary and Future Trends 712
PROBLEMS 713
SYMBOLS 715
REFERENCES 717

10 Optical Methods ***Donald D. Duncan, John L. Champion, Kevin C. Baldwin, and David W. Blodgett*** **721**

10.1 INTRODUCTION 721

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

	10.1.1 Optical Techniques Overview	721
	10.1.2 Historical Perspective	723
	10.1.3 About this Chapter	724
10.2	THEORY	725
	10.2.1 Basic Properties of Light	725
	10.2.2 Interference	729
	10.2.3 Imaging Systems	736
	10.2.4 Holography	740
10.3	OPTICAL TECHNIQUES	744
	10.3.1 Holographic Interferometry	744
	10.3.2 Speckle Techniques	757
	10.3.3 Structured Light	772
	10.3.4 Photoelastic Techniques	775
	10.3.5 Equipment	781
10.4	SUMMARY	786
	PROBLEMS	787
	GLOSSARY	788
	SYMBOLS	792
	APPENDIX A: COHERENCE CONCEPTS	793
	APPENDIX B: POSTSCRIPT PROGRAMMING	803
	REFERENCES	804

Index *807*

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Contributors

Alfred J. Bahr SRI International, Menlo Park, California

Kevin C. Baldwin The Applied Physics Laboratory, Johns Hopkins University, Laurel, Maryland

David W. Blodgett The Applied Physics Laboratory, Johns Hopkins University, Laurel, Maryland

Harold Berger Industrial Quality, Inc., Gaithersburg, Maryland

John L. Champion Research and Technology Development Center, The Applied Physics Laboratory, Johns Hopkins University, Laurel, Maryland

Jim Cox Zetec, Inc., Issaquah, Washington

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Donald D. Duncan The Applied Physics Laboratory, Johns Hopkins University, Laurel, Maryland

Donald Hagemaier* The Boeing Company, Long Beach, California

Frank A. Iddings† Nuclear Science Center, Louisiana State University, Baton Rouge, Louisiana

Tomas S. Jones Industrial Quality, Inc., Gaithersburg, Maryland

Kathy Joseph The Pennsylvania State University, University Park, Pennsylvania

Arthur Lindgren* Magnaflux, Glenview, Illinois

Clinton M. Logan* Lawrence Livermore National Laboratory, University of California, Livermore, California

David Mackintosh Russell NDE Systems, Inc., Edmonton, Alberta, Canada

Harry E. Martz, Jr. Lawrence Livermore National Laboratory, University of California, Livermore, California

Robert Osiander Research and Technology Development Center, The Applied Physics Laboratory, Johns Hopkins University, Laurel, Maryland

William H. Prosser Nondestructive Evaluation Sciences Branch, NASA Langley Research Center, Hampton, Virginia

Nasser Qaddoumi American University of Sharjah, United Arab Emirates

Peter J. Shull The Pennsylvania State University, Altoona, Pennsylvania

Jane Maclachlan Spicer Research and Technology Development Center, The Applied Physics Laboratory, Johns Hopkins University, Laurel, Maryland

Bernard R. Tittmann The Pennsylvania State University, University Park, Pennsylvania

Reza Zoughi Electrical and Computer Engineering Department, University of Missouri–Rolla, Rolla, Missouri

* Retired
† Professor Emeritus

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

1

Introduction to NDE

Peter J. Shull
The Pennsylvania State University, Altoona, Pennsylvania

1.1 WHAT IS NONDESTRUCTIVE EVALUATION?

> Nondestructive evaluation (NDE) is the examination of an object with technology that does not affect the object's future usefulness.
>
> —American Society of Nondestructive Testing (ASNT) (1)

We find both the general concepts and uses of NDE not only in industry but in our everyday lives. For example, imagine you're at the fruit counter at the local market. You glance over the selection of melons, pick one, and look it over for flaws—that's *visual NDE*. Then you test the melon's ripeness by tapping the surface and listening for the hollow response—that's *acoustic NDE*. Finally, depending on the response to these tests relative to your preset NDE criteria (Does the melon look OK? Is it ripe enough?), you either look for another melon, or accept the product and head to the checkout counter.

Almost all of the melons at the market, even the flawed ones, are eventually sold, and this reveals a common problem with NDE: inspectors' standards often change. Ideally, once the accept-or-reject criteria are set, the inspection would be divorced from human judgment and always produce the same result (2).

The basic principle of NDE is simple. To determine the quality or integrity of an item nondestructively, simply find a physical phenomenon (the *interrogat-*

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

ing parameter) that will interact with and be influenced by your test specimen (the *interrogated parameter*) without altering the specimen's function.

For example, some questions can be answered noninvasively by passing X-rays (interrogating parameter) through an object, and monitoring the X-ray absorption (the interaction of the X-rays and the object is the interrogated parameter). You could answer the questions "Is my arm broken?" and "Is there a mouse in my beer can?" using this method.

Such questions do not query the X-ray absorptive qualities of the objects, but the answers are implied by the absorptive qualities. Although in many cases the interrogated parameter is specifically queried (for example, ultrasonically testing the elastic properties of a material), more often than not one must indirectly infer the condition of the test specimen from the results of NDE inspection.

Of course, we could simply slice the watermelon and taste it, or open the can and check for mice, but doing so would clearly alter future usefulness. Similarly, if we want to know the true strength of an old, heavily corroded bridge (common in many countries, including the U.S.), we could simply load it until it fails. But although this method tells us about the strength of the bridge, the principles of NDE have been violated; NDE, by definition, is nondestructive and does not change the physical properties of the test object. So if you squeeze a plum to test it for ripeness, that's NDE, provided you do not squeeze a little too hard and bruise it!

On the other hand, many NDE tests do in fact permanently alter the test specimen and are still considered NDE, provided the test specimen's function is unaltered and the specimen can be returned to service. For example, acoustic emission is commonly used to proof-test gas and liquid containers, pressurizing the tanks while sensors listen for acoustic emissions from submicron cracking caused by the pressure. The tank passes inspection if the amount of acoustic emission is below a predetermined threshold. Although the tank structure clearly has been altered on some level, the damage to the tank by the acoustic emission test does not affect its usefulness as a pressure vessel.

1.2 NDE: WHAT'S IN A NAME?

Within the testing community, NDE is not the only term used to describe the method in question. People have varied the terminology based on their specific application and information derived from the test. The terms are similar but with subtle differences:

- *Nondestructive Testing* (NDT) generally refers to the actual test only.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

- *Nondestructive Evaluation* (NDE) implies not only the test but also its role in the process. This is the most general definition among all the terms.
- *Nondestructive Examination* (NDE) is typically used in areas other than process control.
- *Nondestructive Inspection* (NDI) is, like nondestructive examination, typically used in areas other than process control.
- *Nondestructive Characterization* (NDC) refers to the specific characterization of material properties. NDC, of all the references, probably has the most restricted use.
- *Nondestructive Sensing* (NDS) generically refers to the use of sensors to derive information nondestructively. Here there is no specific implication that the object will be returned to service, although it could be.

Although these various terms are designed to create distinctions, most are used interchangeably. NDT, NDE, and NDI are the most common and are generally nonspecific. In this text, the use of NDE and NDT is interchangeable and implies the broadest of definitions.

1.3 HOW IS NDE APPLIED?

The process of NDE is often dictated by the specific case. The applications may impose requirements on the method and procedure, such as portability (for field use), speed of data acquisition and reduction (for high-speed process control), and environmental hardening (for operation in harsh environments). Likewise, the motivation for inspection (safety, cost) influences the when-and-where of applying NDE. Process or quality control, maintenance, medical, and research represent the most common areas where NDE is employed.

We separate process or quality control into feedback control and accept/reject criteria. In feedback control, the NDE sensor (a) monitors the process and (b) feeds the sensor response back to the process controller, after which (c) the controller, based on this feedback information, controls the process variables to maintain the product within predetermined limits. For example, in the manufacturing of aluminum beverage cans, aluminum billets are hot-rolled into very long thin sheets before processing into cans. If the temperature is too low, the rolling tends to elongate the individual crystals preferentially within the sheet, creating directional texture that adversely affects the deep drawability of the sheet necessary to produce the can. Figure 1.1 shows the "earing" caused by a nonuniform texture in the deep drawing process. To counter the texture, the rolling process includes a heat treatment of the sheet. Ultrasonic NDE sensors monitor the degree of texture within the sheet and feed the information back to

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016



FIGURE 1.1 Earing caused by improper texture in rolled aluminum sheet during a deep draw process for can manufacturing. In this process the NDE inspection data is used as feedback data to adjust the processing of the sheet aluminum.

the heat treatment temperature controller. To be effective, the ultrasonic test must be capable of acquiring data from a sheet moving at about 5–15 cm/s (~1–3 ft/s).

When used as an accept/reject quality control method, NDE examines the *finished* product. If the NDE inspection criteria accept the product, it is returned to perform its original service. If rejected, the product is either processed for recycling, reprocessed, or scrapped. Figure 1.2 shows several products that a manufacturer would like the NDE sensor to tag to reject: a tea bag wrapper without the tea bag, a fortune cookie wrapper without the cookie, and a crooked label on a bottle.*

Typically, the process/quality control engineer chooses a quality control NDE method based on economics. If the product is expensive to manufacture, the additional expense and complications of NDE feedback control may be warranted. (In the production of aluminum cans, if the texture of a given sheet produces earing, the *entire* sheet must be reprocessed or the ears removed by machining; both are expensive processes.) If, on the other hand, neither the raw

*All of these items where purchased off-the-self.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016



FIGURE 1.2 Examples of high-volume, low-cost, products that can be inspected for quality control and are either accepted or rejected.

materials nor the processing pose a significant economic expense, then NDE accept/reject criteria may be appropriate. Below a certain level, however, economic expense may be too insignificant to warrant a fully developed method of quality control: if one of our fortune cookie wrappers contains no fortune cookie, why bother? Then again, if we produce *enough* defective products, customers will look for a different supplier.

The extensive use of NDE as a diagnostic maintenance tool is driven by either economics and/or safety. Planned periodic maintenance of components is far more economical than an unexpected failure. In industries such as power production, NDE inspection/maintenance is considered in the design of the structure—that is, the structure is *designed for inspectability*. Systems designed for inspectability often do not have to be completely taken off-line during an inspection.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

In many industries, a failure of a single component can cause catastrophic results—for example, failure of a landing gear on a commercial aircraft. (Figure 1.3 shows penetrant testing applied to a landing gear.) Such components are generally referred to as *safety critical*. The airline industry, the power industry, and the government (for civil structures such as bridges and dams) all have large NDE programs. In many cases, inspection is also mandated by federal regulations.

The medical industry primarily uses NDE as a diagnostic tool. Common examples of medical NDE methods are standard X-ray, computer-aided tomography (CAT) scan, magnetic resonance imaging (MRI), and ultrasound. These methods are not always used as NDE tools; for example, at one time high-power ultrasound was used to fracture gallstones—a destructive process and therefore not considered NDE, even though the same technology is used.



FIGURE 1.3 Penetrant testing applied to a safety critical part on an aircraft—the landing gear. (Courtesy Magnaflux.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

1.4 UNDERSTANDING THE NDE CHOICES

Once we have decided to employ NDE methods to improve product quality or ensure the integrity of a part, there are several levels to choosing an appropriate NDE method for a specific application. We must have a sense of what NDE methods would maximize the likelihood of detecting the flaw or material property of interest, while also considering economic, regulatory, and other factors.

The basic levels of choosing an NDE method are:

1. Understanding the physical nature of the material property or discontinuity to be inspected
2. Understanding the underlying physical processes that govern NDE methods
3. Understanding the physical nature of the interaction of the probing field (or material) with the test material
4. Understanding the potential and limitations of available technology
5. Considering economic, environmental, regulatory, and other factors

To employ any NDE method, we need to have a reasonable knowledge of what we are looking for—material properties, a discontinuity such as a void or crack, sheet or coating thickness, etc. For material properties, we might be interested in the mechanical properties (elastic constants) or electromagnetic properties (conductivity, permittivity, or magnetic permeability). For discontinuities, we must not only be aware of their character, but also understand its relationship to the parent material, e.g., a crack in a ceramic is very different than a crack in a reinforced polymer composite.

We must also have a basic knowledge of how the various NDE methods work. For example, eddy currents use a magnetic field to create induced currents in the test part as the interrogating field. Thus, eddy-current methods require the test part to be electrically conductive. In addition, we must understand the interaction between the operating principles of the NDE method and those of the property or characteristic of interest in the test part. This is required (a) to determine if a given method will work, (b) to select a method in case of multiple choices, and (c) to determine the compatibility of the method and the part; for example, most ultrasonic methods require a coupling fluid or gel that might contaminate or cause corrosion of sensitive parts.

In selecting an NDE method, we must be aware of the potentials and limitations of the existing technology. Just because there is consistency between the physical principles of the NDE method and those of the test part does not mean that equipment is available or has the sensitivity to measure the desired feature. We must also be aware of the actual sensitivity of the equipment relative to the environment in which it is used. Although manufacturers often quote

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

laboratory-measured sensitivities, these may differ greatly from what you can expect in industrial inspections.

Many other factors can influence our choice of NDE method. Is it cost effective to employ NDE? Is speed of inspection a factor? What regulations exist that may dictate not only the NDE method but also the procedure? What environmental issues constrain our choices? Is the test to be performed in the field or in a specialized testing facility?

With a basic understand of all five levels, we can make an informed choice that optimizes NDE method with the application constraints.

1.5 HOW MUCH DO WE INSPECT?

As with most NDE choices, the question of whether to inspect 100% of a given part, and whether to inspect all parts or a representative subset, is a matter of economics or regulatory oversight. Ideally, we would need only to inspect parts that in fact have flaws. But because the existence of flaws is an unknown, the decision of how much to inspect is part of the overall NDE testing program. Generally, statistical methods are applied to a specific application and the results are used to compare the probability of the existence of a discontinuity to a preset criterion of acceptable flaw limits (type, size, frequency, etc.). These statistical results are used to develop NDE programs designed to ensure the part's or the system's integrity.

In general, a manufacturer or a maintenance facility would prefer to minimize the amount and degree of inspection, while still maintaining confidence in the quality and integrity of the product. To determine how much testing is necessary, statistical methods are applied. For example, when inspecting heat exchanger tube bundles, only a representative number are tested; one such method is the star pattern (Figure 1.4), that, by assuming that neighboring parts see similar environmental degradation (corrosion or erosion), predicts the overall performance of the entire tube bundle. If extensive damage is found in a specific area, the neighboring tubes are inspected more closely.

Other products demand more rigorous inspection. For example, safety-critical parts on a nuclear reactor or liquid petroleum gas (LPG) container are required by regulation to be completely inspected routinely. Newly installed gas pipelines generally have 100% inspection of all welds, though only routine partial inspection thereafter.

1.5.1 Statistics

Successful NDE programs use statistics extensively, to determine the scope of testing required. The importance of statistical analysis can be reasoned through a simple example. Assume that a study by a bicycle frame manufacturer determines

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016



FIGURE 1.4 Partial inspection of a heat exchanger tube bundle. (Courtesy Russell NDE Systems.)

that a weld on the frame has a 1-in-1000 chance of failing in the first year of normal use. This may seem a reasonable failure rate, until we realize that there are 20 such welds on the frame. Now the probability of frame failure in the first year is 1 in 50. If we now consider how many critical components there are on a complex system, such as a jet engine, we begin to understand the scope of ensuring the quality of the individual components.

Statistical analysis of a part is based on the fracture mechanics, which combines loading conditions, material properties, and mechanics of materials. For NDE we are interested in ascertaining what conditions cause a part to fail, and where and how the failure occurs. Once this information is determined, we can decide how much of a part, and which sections, should be inspected. For example, in the manufacture of ball joints for automotive steering columns, only the high-wear and high-stress regions receive 100 % inspection; the rest of the part may not be inspected at all (Figure 1.5). Often, the extent of testing is recommended by the design or manufacturing engineer.

1.5.2 Consequences of Part Failure (Non-Safety-Critical Parts)

Many parts are inspected not because of possible safety issues, but for the sake of customer satisfaction. Customer satisfaction in itself contains many different levels that can affect a company's reputation. Consider the following three flawed products:

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

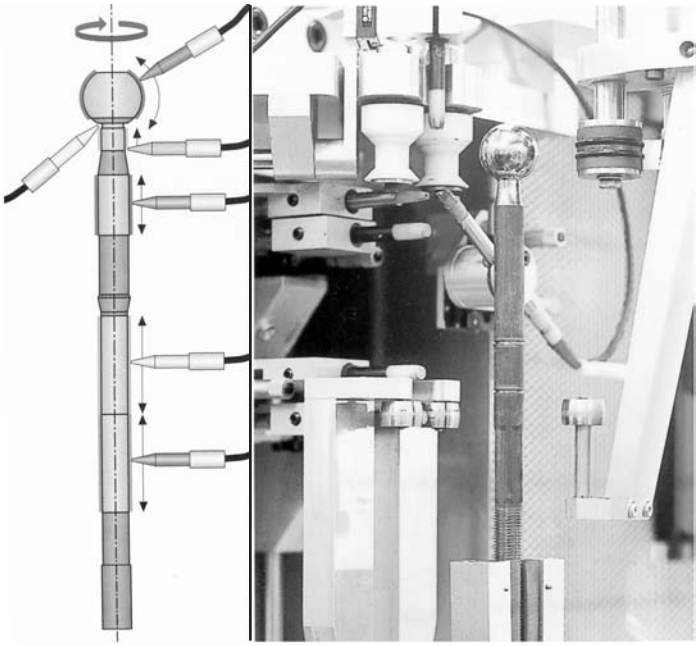


FIGURE 1.5 Partial inspection of automotive steering column ball joints using an automated, multiple head, eddy current system. This system operates on an accept/reject criterion. (Courtesy Foerster Instruments.)

1. A fortune cookie wrapper that does not contain a cookie
2. A bottle of expensive scotch with a crooked label
3. A defective fan belt on an automobile

The missing fortune cookie is of little consequence; one usually buys these items in quantity, and if only one is missing the buyer thinks little of it—as long as the percentage of defective products is low, the average customer does not feel cheated. On the other hand, the crooked label on an expensive bottle of scotch may trigger a complaint. Even though the quality of the scotch is not affected, the customer may question the quality of the product simply because of the label orientation. Scotch requires very careful processing and attention to detail.

A defective fan belt, of course, is even more serious. If you have ever been stranded in an automobile, you understand the customer's possibly long-term prejudice they may hold towards the manufacturer. Considering the number of components on a car that would disable it if they failed, the probability of failure of each part must be extremely small to ensure customer satisfaction.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

1.5.3 Larger Systems or Safety-Critical Parts

Large, expensive systems and safety-critical components receive more rigorous quality control than do parts whose failure poses little economic or personal risk. For large systems such as power plants and aircraft, elaborate NDE procedures have been developed to ensure longevity and safety. *Retirement-for-cause* NDE methods specifically address the issue of increasing the useful life of a system. *Risk-informed* methods focus inspection on components of a system that pose the greatest threat, based on the probability of a part failing *and* the consequence in the event of failure.

1.5.4 Retirement for Cause

The effort to extend structures beyond their design lifetime has been the impetus for developing retirement-for-cause NDE methods. Instead of being retired after a predetermined lifetime, a part or structure (such as an aircraft) is now often retired only if found to be defective. For example, a significant portion of the U.S. commercial and military air fleet has remained in operation well beyond lifetime dictates; in fact, some aircraft have been in service more than double their design lifetime.*

On structures with many safety-critical components, parts used to be replaced on a schedule based on earliest-expected failure; effectively, the majority of components were retired well before exhausting their useful life. More recently, such components have undergone rigorous inspection schedules instead, and are retired only when a cause is found.

Retirement for cause, however, increases the risk of component failure. Therefore, rigorous inspection schedule and procedures must be maintained.

1.5.5 Risk-Informed Inspection

Inspection practices for large and/or complex structures have shifted from prescriptive procedures to more risk-informed and performance-based inspection. For example, power plants contain thousands of welds requiring periodic inspection. Instead of simply inspecting these welds indiscriminately, inspections are based on a combination of the likelihood of failure, and the risk associated with failure, of a specific part (all of the welds might have similar likelihood of failure, but the consequences can be drastically different). Such planning allows resources to be concentrated on the most critical system components.

In risk-informed inspection, the probability of failure of a specific component is weighed against the consequence of its failure to develop a quality factor

*Note that airlines in the U.S. are rapidly replacing their aging fleets.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

TABLE 1.1 Risk-based Determination of NDE Testing Procedure and Schedule for Individual Components of a System

Risk Categories High Cat. 1, 2, and 3 Medium Cat. 4 and 5 Low Cat. 6 and 7		Consequence Category			
		None	Low	Medium	High
Degradation	Large	Cat. 7	Cat. 5	Cat. 3	Cat. 1
Mechanism	Small	Cat. 7	Cat. 6	Cat. 5	Cat. 2
Category	None	Cat. 7	Cat. 7	Cat. 6	Cat. 4

of its overall risk. Based on this risk assessment, the component is assigned a value indicating the frequency and extent of inspection required. Further, the inspection-for-cause method is applied to each inspected component to make certain the appropriate NDE method, procedures, and acceptance criteria are applied to address the specific failure expected (corrosion, cracking, etc.).

Table 1.1 depicts a risk-based matrix used to rank components in order of inspection needs. The matrix maps a component's statistical chance of failure (rows) with the consequence in case of failure (columns). The shaded regions rank the importance of testing the component as low-, medium-, and high-risk. Low-risk components require little or no inspection; if a low-risk part fails, even frequently, it can simply be repaired or replaced, with no major consequences to personal safety or downtime of the entire process. Medium-risk components are analyzed further, to determine if immediate or delayed NDE inspection is necessary. The high-risk category implies that a failure of this part would be very costly, and so the part is to be given the highest inspection priority.

Risk-informed NDE inspection methods require significant planning to develop a risk matrix, also known as a *Failure Mode and Effect Analysis* (FMEA). The method requires the analysis of considerable information about the operating facility, system components, history of the process and individual components, and previous failure and inspection records. Implementing a risk-informed method requires iteration over several inspection periods. Initially, many component risk values are simply a best guess using available data.

Because of the significant initial investment, risk-informed methods are typically used in large systems or safety-critical structures. According to ASME Boiler and Pressure Vessel Code Case N-560, applying a risk-informed method to Class 1 piping reduced inspections from 25 % to 10 %. For a typical power plant, this reduction translates to a cost reduction of about $1 million per inspection cycle, while also decreasing the overall risk at the plant.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

1.6 RELIABILITY OF NDE

Throughout most of history, the integrity of a part or structure was simply tested by time: you built a structure, and waited to see if it fell down. In modern times, we have developed many methods to query the integrity of parts or structures, using both destructive and nondestructive methods. Once we have a method, we have to know how reliable it is. There are many facets to this question, most of which are based on statistical analysis (see Section 1.5). This section focuses on the human factors that influence the reliability of an NDE inspection.

Inspector influence on reliability can be affected by a variety of factors—personal, environmental, and external. Personal influences include physical and mental attributes. All NDE methods require a certain level of physical competence, such as visual acuity and manual dexterity. Even though a qualified inspector must possess these competencies, they can be diminished or subverted by fatigue, unreliable vision (e.g. persistence of vision in fluorescent particle testing), adaptation to light, and general distractions from the inspections. Mental attributes cover the inspector's attitude towards the work. Of course, attitude is difficult to control; for this reason, inspection procedures are designed to track the operator's steps. Nonetheless many external pressures, such as production schedules, costs, fatigue, and temporary personal issues, can subvert an operator's integrity. Consider the consequences of a personal tragedy (such as a death in the family or a divorce) on an inspector's ability to concentrate on a relatively repetitive task.

The working environment also plays a significant role in both the sensitivity and reliability of NDE methods. The ideal testing environment is one designed specifically and solely for its purpose. Such an inspection environment would control outside influences such as lighting and cleanliness, allow for efficient handling of parts, use equipment not limited by portability or power, provide a controlled environment for the comfort of the operator, and allow the operator to go home at the end of the shift. Fieldwork, by contrast, is often performed in relatively inhospitable and remote locations such as pipeline, bridges, fuel containers, construction sites. In such environments, operators must compensate for the inability to control ambient conditions such as lighting and temperature. For example, they could use fluorescent magnetic particles, which are far more visible than colored particles. But since fluorescent particles require a darkened environment for viewing, and such an environment is generally unavailable in the field, the less sensitive colored particles are used anyway. Additionally, the operator must adapt the inspection to an existing structure, such as a tower or pipeline. To imagine the challenges, consider inspecting the entire circumference of an above-ground pipeline weld in cold weather (Figure 1.6).

External factors that contribute to the reliability of NDE inspections are reflected in the general working environment. For example, an inspector may be

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016



FIGURE 1.6 Magnetic particle inspection of a pipeline weld. (Courtesy Magnaflux.)

highly qualified and motivated, but the reliability of the results may be subverted in the reporting process, particularly if costs or production schedules are affected. It is particularly important that a positive environment exist in which the inspector can freely report findings without the pressure to produce specific results.

To give you some idea of how human factors can affect the reliability of NDE inspection, consider Silk's analogy to accidents during driving (3). Like NDE, driving is a skilled task with significant consequences if performed improperly, and it is subject to personal, environmental, and external influences. Assume that there would be no accidents if drivers strictly adhered to the procedures (the rules and regulations) that govern traffic control. (For the sake of the exercise, don't be concerned that this is a fallacy.) Therefore, we will attribute all accidents to a lapse in following these rules. Silk (3) quotes the number of accidents per 10^8 miles traveled as 324,840 due to any cause. This

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

translates to an operator making a mistake (with consequences) 3.25 times per 1000 miles traveled. If we remove all nonprocedural factors, such as heart attacks or pedestrians dropping pumpkins from overpasses, the rate of human error while driving drops to about 1 mistake per 1000 miles. This value is approxi-mately consistent with studies on the reliability of human-operated NDE inspections.

REFERENCES

1. What is NDE? American Society of Nondestructive Testing, 2000. http://www.asnt.org/whatndt/
2. R. Hochschild. Electromagnetic Methods of Testing Metals. In: EG Stanford, JH Fearon, eds. Progress in Non-destructive Testing. London: Heywood & Company, 1958, p 60.
3. MG Silk, AM Stoneham, and JAG Temple. The Reliability of Non-destructive Inspection. Bristol: Adam Hilger, 1987, pp 110–115.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

2

Liquid Penetrant

Frank A. Iddings*
Louisiana State University, Baton Rouge, Louisiana

Peter J. Shull
The Pennsylvania State University, Altoona, Pennsylvania

2.1 INTRODUCTION

Penetrant testing (PT) is a rapid, simple, inexpensive, and sensitive *nondestructive testing* (NDT) method. It allows the inspection of a large variety of materials, component parts, and systems for discontinuities that are open to the surface; these discontinuities may be inherent in the original materials, result from some fabrication process, or develop from usage and/or attack by the environment. Because PT is portable, it is often used in remote locations.

Using PT effectively requires:

Discontinuities open to the *surface* of the part (subsurface discontinuities or surface discontinuities not open to the surface aren't detected)
Special *cleaning* of parts
Good eyesight

*Professor Emeritus

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

This chapter discusses the fundamentals, techniques, and applications of penetrant testing.

2.1.1 Basic Technique (1–4)

Almost everyone has seen cracks in a ceramic coffee cup, or in the glaze on the cup, made visible by the coffee held in the cracks (Figure 2.1). Or you may have seen previously invisible cracks in concrete sidewalk paving, made visible by the dirty water leaking out of the cracks after a rain. If you have worked on an old automobile engine, you have probably seen cracks appear when dirty oil leaks out of them.

These cracks are *surface discontinuities* made visible by a *penetrant*—a material that seeps into and out of a surface discontinuity. To be seen, the penetrant must be a strikingly different color (contrast) from the surface and must exit the discontinuity under appropriate conditions. A blotter-like material on the surface, may draw the penetrant from the discontinuity, revealing the outline of the discontinuity. Such activity is penetrant testing in its simplest form.

PT requires less training and mechanical skills as compared to some other NDT methods, but you need to pay careful attention to cleanliness, procedures, and processing time, and you need comprehensive knowledge of types of discontinuities that may occur in the parts to be tested.

A detailed visual examination of the surface(s) to be tested should be conducted prior to beginning any NDT. This will identify any conditions that



FIGURE 2.1 Crack in a coffee cup made visible by the coffee penetrating into the crack. The high contrast between the penetrant (the coffee) and the white cup makes the fine crack clearly visible.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

would interfere with the test. Cleaning is crucial at several stages. First, cleaning is necessary so that nothing on the surface of the specimen prevents the discontinuity from being open to the surface—all rust, dirt, protective coatings, or any other interfering materials must be removed or cleaned from the surface. Then the penetrant can be applied to the surface and allowed to enter the discontinuity.

Procedures determine (a) the type of penetrant and (b) the *dwell time*, or length of time for the penetrant to enter the defects. After the penetrant has entered the defects, excess (surface) penetrant must be removed—without removing the penetrant that has entered the defects.

After this second cleaning, another material, called the developer, is placed on the surface. The developer draws some of the penetrant from the defects, and provides a contrasting background to make the penetrant easier to see. After the indications from the penetrant/developer have been interpreted and perhaps recorded, the surface is cleaned a third time, to remove the developer and any remaining penetrant.

2.1.2 History

PT may have been used in the early 19th century, by metal fabricators who noticed quench or cleaning liquids seeping out of cracks and other defects not obvious to the eye. But the earliest recognized application of PT was the "oil and whiting" inspection of railroad axles, wheels, couplers, and locomotive parts in the late-19th century. Railroad inspectors used "oil" (a penetrant made of a dark lubricating oil diluted with kerosene) to find surface defects. The parts of interest were coated with the oil, and after an appropriate dwell time (to permit the mixture to enter the defects), the excess oil was removed.

After the oil, a "whiting" (a developer made of a mixture of chalk and alcohol) was then applied to the parts. Evaporation of the alcohol left a thin, white powder coating of the chalk on the surface. When the part was struck with mallets, the oil trapped in the discontinuities was forced out, making a blackish stain on the chalk coating. Thus it was easy to see where the discontinuities were. Railroad inspectors used the oil-and-whiting technique until around 1940, when it was largely replaced by magnetic particle testing of ferromagnetic parts.

Two improvements to the oil-and-whiting technique allowed it to continue in limited use after 1940. The first change was to use an ultraviolet (UV or black light) lamp to illuminate the specimen after cleaning the excess oil from the surface. Since oil under a strong ultraviolet light fluoresces with a pale bluish glow, the traces of oil seeping out of small defects could be seen much more clearly. The second change was to heat the part and the oil, to improve penetration into the discontinuities. The thin, hot oil could enter the expanded discontinuities more easily, and the cooling of the part helped pull the penetrant into surface

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

discontinuities. But even with these changes, the "hot oil method" is much less sensitive than penetrant systems in use today (1, 2, 4).

During World War II, the aircraft industry needed better penetrant testing for nonferromagnetic parts, which could not be tested with magnetic methods or required greater sensitivity to ferromagnetic parts with discontinuities open to the surface. Robert C. and Joseph L. Switzer, working in 1941, made such significant improvements in fluorescent and color dye penetrant methods that they are still used today. Dyes, that were either strongly fluorescent or were brightly colored were introduced into special formulated oils with greater penetrating ability (hence recognition of many penetrant methods as *dye penetrant* methods). Compared to the oil-and-whiting or hot-oil methods, the fluorescent and visible (usually red) dyes provided much greater sensitivity. The fluorescence dye, when exposed to ultraviolet light, emits a bright yellow-green light, very close to the most sensitive range of the human eye. Red dye on a background of white developer provided the contrast needed for sensitive inspections. Many industries rapidly accepted these new penetrant techniques, as several commercial manufacturers made inspection systems available (1–4).

Advances in materials and methods continued. Some improvements have been driven by industry and government safety requirements, such as reduced toxicity of materials and reduced fire hazard from organic solvents made more environmentally friendly by removal of fluorocarbons and certain hydrocarbon compounds. Emulsifying agents (materials that make organic materials miscible with water) have been added and then removed from the penetrants with post emulsification (after the dwell time) of the penetrant being preferred now. The emulsifying agents permit removal of the penetrant with water. Specialized penetrants have been developed for use in situations (such as in the aerospace and nuclear industries) where halogens (present in early formulations of penetrants), which might induce corrosion in some materials, are now limited to lower levels. Penetrants have also been developed for systems using liquid oxygen (LOX) or other powerful oxidizing agents.

Other developments in penetrant formulations and documentation (Polaroid® photography, video recording, and plastic films) continue to expand PT's usefulness.

2.1.3 PT's Potential

Almost every industry has found a use for PT. Industries include metal production; metal fabrication; automotive, marine, and aerospace manufacture and maintenance; petrochemical industry, electrical power generation, electronics manufacture; and composite materials manufacture and maintenance (4).

PT can be and has been applied to almost every conceivable material. PT finds three classes of discontinuities: (a) inherent discontinuities in raw materials,

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

(b) fabrication discontinuities in manufactured components, and (c) service induced discontinuities.

2.1.4 Advantages/Disadvantages

PT's greatest advantage is its wide applicability. Even some porous materials can be inspected using a special penetrant system.

In addition, PT is a relatively simple NDT method as compared to the other methods and very economical. Testing kits are available that can be easily carried by one person to remote test sites. Large or small specimens in a variety of shapes and materials can be tested. The method can provide reliable results and gives few false indications in the hands of experienced inspectors. PT can be automated to provide rapid, large-area inspection of simple geometry specimens or large quantities of similar specimens (3).

The major disadvantage of PT is that it can detect only defects that are open to the surface of a specimen. In addition, results depend upon the experience and skill of the inspector. Porous, dirty, or rough surfaces on the specimen can hide discontinuities. Specimens at high (above 120°F/49°C) or low (below 40°F/4°C) temperatures cannot be tested with the standard penetrant materials, although some special penetrant materials can be used at much higher (350°F/175°C) or lower (10°F/−12°C) temperatures. Very tight discontinuities that restrict the entry of the penetrant may not be detected. Material and discontinuity types, and test conditions such as temperature and humidity, may also affect the sensitivity of the method (2, 4–6).

Although the cleaning materials used in PT are gradually being changed to comply with environmental and safety requirements, it is still important to consider their impact and use them responsibly (7–9). Such care may increase costs and inconvenience. The materials are also a bit messy and can pose a safety hazard in poorly ventilated areas.

Table 2.1 reviews the advantages and disadvantages of PT.

2.2 FUNDAMENTALS

This section reviews the basic principles of penetrant testing: fluid flow and illumination and detection.

2.2.1 Fluid Flow

Fundamental to PT is the ability of the penetrant fluid to coat (wet) the specimen surface completely, and then to penetrate the depths of the discontinuities open to the surface. Surface tension, contact angle and surface wetting, capillarity, and dwell time describe these functions (10).

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

TABLE 2.1 Advantages and Disadvantages of Penetrant Nondestructive Testing

Advantages	Disadvantages
Inexpensive	Surface connected defects only
Easy to apply	Poor on hot, dirty, rough surfaces
Use on most materials	Poor on porous materials
Rapid	Messy
Portable	Environmental and safety concerns
Small learning curve	Temperature range limitations
Accommodates wide range of sizes and shapes of test objects (works well on complex shapes)	Removal of coatings such as paint or other protective coatings
Does not require power source	Highly operator-dependent
Yields flaw location, orientation, approximate size and shape	Mechanical operations such as shot peening, grinding, machining, buffing, etc. tend to close the crack opening by smearing the surface material
Volume processing—batches of parts can be immersed in the penetrant fluids	Final inspection is visual
100% surface inspection	

Surface Tension

The physical phenomenon of *surface tension* explains the dynamics of blowing a soap bubble; it also explains why rain beads on a freshly waxed car, yet forms little pools on a car with a dull finish.

To examine the nature of surface tension, let's begin with the concept of surface energy. Imagine a droplet of water floating on the space shuttle (ignore the force of gravity). A particle of water within the bulk of the water droplet (away from the surface) experiences omnidirectional attractive molecular forces (*cohesive forces*) from the surrounding water particles. The net force acting on the particle is zero.

But as we examine particles near or at the surface of the droplet, the total cohesive forces become hemispherical in direction, i.e., there are no water particles outside the surface of the water droplet. Consequently, local forces acting on the surface particles do not sum to zero, and the particles experience a net force directed perpendicular to the surface and inward towards the liquid.

Eventually, this compressive force is balanced by outwardly directed elastic forces. In fact, a water molecule's very existence at or near the surface requires work to overcome this inwardly directed force. The energy associated with this work is stored in the surface layer in the form of an *elastic potential* energy, called

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

surface energy. The laws of thermodynamics, as well as Newton's laws of motion, require the fluid surface to move to a minimum energy state. This movement is analogous to that of a ball falling from a height (i.e., a given potential energy) to the surface of the earth (i.e., the minimum potential energy) (11). The surface energy of the water droplet is minimized when the surface area is at a minimum, i.e., when the droplet is a sphere.

If surface energy is reduced through a decrease in surface area, there must be an associated elastic force that acts tangential to the fluid surface to reduce the area. This force is called *surface tension*. Figure 2.2 shows the setup for the classic experiment to determine the surface tension for a given fluid. The U-shaped wire frame, with the slider at the base of the U, is partially submerged in the fluid. The slider is pulled slowly out of the liquid along the frame, creating a liquid film. The experiment measures the force required to increase the area of the film by a given amount. (Interestingly, this force is *independent* of the total film area. The independence contrasts with our understanding of an elastic membrane [such as a rubber balloon], where the force required to increase the membrane's area is directly proportional to the existing area. But because the molecules of the fluid film are orders of magnitude smaller than the thickness of the film, the film is primarily a bulk fluid. As the film area is increased, molecules of the liquid



FIGURE 2.2 Classic experiment to demonstrate the details of surface tension of a fluid.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

simply move from the bulk to the surface. In other words, fluids do *not* have to overcome shear stresses at slow rates of film movement (12).

The force F acting on the moveable wire of length l increases the surface area of both the front and back of the film. The surface tension, γ, for this example is given as

$$\gamma = \frac{F}{2l} \tag{2.1}$$

The surface tension for a soap bubble is related to the difference between the bubble's exterior and interior pressure ($p_o - p_i$) and its radius of curvature, r, as

$$p_o - p_i = \frac{4\gamma}{r} \tag{2.2}$$

whereas the surface tension for a droplet of liquid is

$$p_o - p_i = \frac{2\gamma}{r} \tag{2.3}$$

Surface tension can also be viewed as the forces that account for the differences in molecular attraction among like molecules (liquid–liquid and gas–gas) and unlike molecules (gas–liquid) that constitute the surface (boundary) between a gas and a liquid. Therefore γ, the surface tension between a liquid and gas, is more aptly written as $\gamma_{\text{liquid}-\text{gas}}$ or γ_{lg}.

Contact Angle and Surface Wetting

Surface tension is not unique to fluids. In fact, surface energy and surface tension *define* the surface or boundary of any condensed matter, solid or liquid. This section discusses the effects of the balance of the different surface forces at the interface of a solid, a liquid, and a gas.

We know from experience that a fluid in contact with a solid can manifest itself in many ways. The fluid might be contracted into little beads, like rain on a newly waxed car finish; or form little pools, like rain on a car with a dull finish; or completely wet (coat) a surface, like oil poured onto a cooking pan.* To understand these various responses, imagine a liquid in contact with a solid (Figure 2.3). We must also include the gas, which contacts both the fluid and the

*Oil is a popular lubricant partly because it coats a surface so well.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

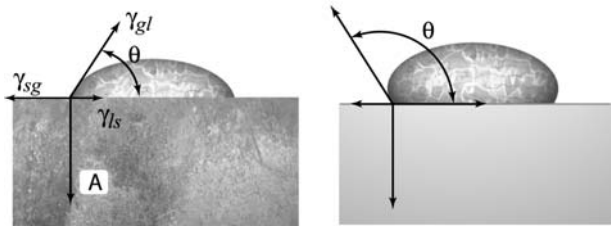


FIGURE 2.3 The forces acting on the interface between a liquid, a solid, and a gas are the surface tensions between the respective materials (indicated by subscripts). *A* represents the *attractive force* of the fluid-to-solid as a result of the interface.

solid. The following surface tensions act on the various interfaces that dictate the response of the liquid on the solid:

γ_{gl} Surface tension of the interface between the gas and the liquid

γ_{ls} Surface tension of the interface between the liquid and the solid

γ_{sg} Surface tension of the interface between the solid and the gas

Let us examine the interface where all three (gas, solid, and liquid) meet. The forces act tangentially to their respective surfaces. Because the solid surface does not flow (by definition), the directions of γ_{sg} and γ_{ls} lie in the plane of the solid. But the direction of γ_{gl}, which lies tangent to the liquid–gas surface, is determined by a balance of *all* the forces acting at the interface—the surface tensions of the three surfaces, and an *adhesive* force that balances the interaction of the different forces of the three different materials. In the plane of the solid (x-directed) the sum of the forces is

$$\sum F_x = \gamma_{ls} - \gamma_{sg} - \gamma_{gl}\cos\theta = 0 \tag{2.4}$$

and perpendicular to the solid surface (y-directed)

$$\sum F_y = \gamma_{gl}\sin(\theta) - A = 0 \tag{2.5}$$

where A represents the *adhesive* force. Notice that A acts *normal* to and *inward* towards the solid surface, thus *adhering* the liquid surface to the solid. The *contact angle* θ defines the angle at which the fluid surface meets the solid surface. This contact angle, which is a function of the material properties of the contacting materials, indicates the degree to which the fluid will wet (coat) the surface. The four cases are listed:

Copyright © Marcel Dekker, Inc. All rights reserved.

$\theta = 0$	The fluid will *completely* wet the surface of the solid (note: the gas will no longer touch the solid).
$0 < \theta < 90^\circ$	The fluid will tend to spread out and wet the solid surface.
$90^\circ < \theta < 180^\circ$	The fluid will tend to contract and form beads of liquid on the solid surface.
$\theta = 180^\circ$	The fluid will contract and bead to the point where the fluid will not contact the solid surface and it detaches.

Fluids whose contact angles are greater than 90° are said *not* to wet the surface, while fluids at angles less than 90° wet the surface. Typical penetrant testing fluids have contact angles on the order of 10°.

Capillarity

Wetting the surface of the test specimen is the first requirement of the penetrant fluid. Once the penetrant has coated the surface, it must be drawn into the depths of the surface-breaking discontinuities. *Capillarity* is the driving force that draws fluid into the narrow openings, regardless of their orientation relative to gravity.

Water in a glass container slightly climbs the walls to a level above the nominal water level, along a curve of radius r (Figure 2.4a). On the other hand, if the glass were coated with wax, the water would appear to pull *away* from the container walls (Figure 2.4b). These different possible contact angles are explained by the balance of the constituent surface tensions and the adhesion force A. When the distance between the walls of a container is less than twice the curvature r, the forces become unbalanced and the fluid rises or falls an additional distance (Figures 2.4c and 2.4d). Such a container is called a *capillary*, and the fluid movement away from its natural level is called *capillary action*. The curvature of the fluid is called the *meniscus* ("little moon").

Now, imagine a capillary that is open into a reservoir of fluid. If the contact angle is less than 90°, the fluid wets the surface and rises in the capillary above the reservoir fluid level (Figure 2.4c). If the fluid contact angle is less than 90°, the liquid in the capillary will depress below the reservoir level (Figure 2.4d).

The height of liquid ascension or depression is a function of the contact angle of the liquid on the solid *and* the surface tension of the liquid. (The contact angle is a function of the *difference* of the surface tensions.) This height, h, can be calculated by summing the forces acting on the liquid in the capillary: (a) surface tension and (b) gravity force. The liquid surface layer exerts an attractive force (surface tension) on the capillary wall at the line of contact—$2\pi r$ for a cylindrical tube. Therefore, the force pulling the liquid along the length of the capillary is

$$(2\pi r)\gamma_{gl}\cos(\theta) \tag{2.6}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

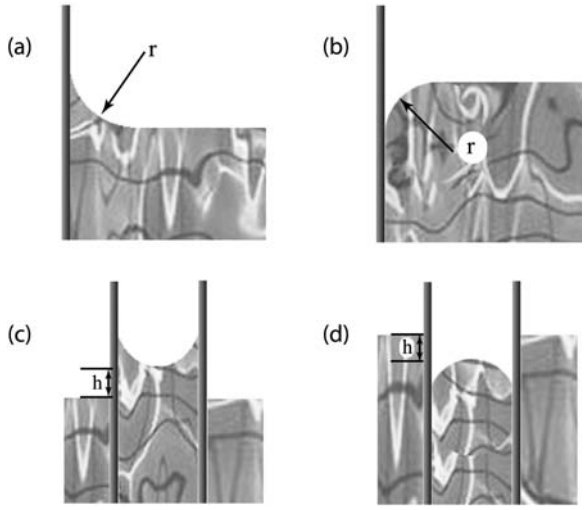


FIGURE 2.4 The ascension (a) or depression (b) of a fluid at the wall of a large container. (c, d) Capillary action of a fluid where the container walls are less than twice the radius *r*.

whereas the gravity force opposing this movement is

$$(\pi r^2 h)\rho g \tag{2.7}$$

where ρ is the fluid density and g is the acceleration of gravity. The final fluid height occurs when these two forces balance, i.e.,

$$(2\pi r)\gamma_{gl}\cos(\theta) = \pi r^2 \rho g h;$$

$$h = \frac{2\gamma_{gl}\cos(\theta)}{r\rho g} \tag{2.8}$$

Fluid penetration into a real crack will generally be different from that estimated by Eq. (1.8), for three reasons. First, crack width is not a constant: a crack typically narrows with depth. Second, parts of the crack wall may be in contact, i.e., portions of the crack may be closed. Third, trapped gas or contaminants within the crack limits fluid penetration. Nonetheless, penetration depth is still proportional to penetrant surface tension, and inversely proportional to crack width and penetrant density.

One might expect the fluid *viscosity* to affect both surface wetting and crack penetration. (Consider viscosity as the analog to friction for a solid in motion.) The higher the viscosity (friction) coefficient, the more resistance there is to fluid (solid) movement. Although the penetration time is greatly influenced by the viscosity—higher viscosity, slower penetration rate—viscosity does not appear to

Copyright © Marcel Dekker, Inc. All rights reserved.

significantly affect sensitivity or the ability of the fluid to penetrate fine cracks, at least within the range of viscosities used for penetrant systems.

Dwell Time

To detect a crack of a given size, the penetrant fluid requires a lapse of time for capillary action to draw the penetrant into the flaws. This time between applying the penetrant and removing the excess fluid is called the *dwell time*. Although a number of mathematical models exist for predicting dwell time, in practice, inspectors generally refer to tables of empirically determined times. The most effective manner of determining the minimum required dwell time is to perform a procedure qualification on parts containing known discontinuities. All of the essential test parameters that affect the results, such as, temperature, specific materials used, times, etc., should be documented.

2.2.2 Illumination and Detection: The Eye's Response to Visible Spectrum Illumination

To detect defects using PT, the indicator material must be visible. This visibility normally requires *illumination*. The type and intensity of illumination of indicators depends on:

- Where the inspection is to be performed—in a specialized testing facility (where the lighting is easily controlled) or in the field (where background lighting cannot be controlled)
- The size and shape of the part
- Whether the inspection is automated (using laser scanning or full field imaging) or performed by a human viewer (This section deals only with human viewing.)

One reason that PT is simpler and more versatile than other NDE methods is that a human operator can simply *see* defects. True, automated systems (using laser scanning or full-field imaging) are sometimes used where high repeatability is required, but these systems are expensive, complicated, and not very versatile. Consequently, the most commonly used detector in PT is the unaided human eye.

The spectral response of the human eye is fairly constant throughout the species. (Only about 8 % of the population has some form of color blindness.) But over the so-called *visible* spectrum ($780 > \lambda > 390$ nm or $384 < f < 769$ THz [Table 2.2][13]), the response of the human eye varies. Figure 2.5 plots the normalized frequency response of the *nominal* human eye in bright light (photopic: cone response) and in dim light (scotopic: rod response). The eye's sensitivity shifts from the yellow–red spectrum (in bright light) toward the green–blue spectrum (in dim light). This shift is known as the *Purkinje shift* (14).

Copyright © Marcel Dekker, Inc. All rights reserved.

TABLE 2.2 The Colors of the Visible Spectrum and Their Approximate Frequencies and Wavelengths in Free Space

Color	λ nm	f THz
Red	780–622	384–482
Orange	622–597	482–503
Yellow	597–577	503–520
Green	577–492	520–610
Blue	492–455	610–659
Violet	455–390	659–769

nm $= 10^{-9}$ m and THz $= 10^{12}$ Hz

Penetrant dyes are chosen to make detection easier under both bright and dim light. In bright light (natural or artificial), *color-contrast* (visible) dye, typically red dye on a white background, is used; the high contrast maximizes the eye's response to indications. The intensity and type of ambient lighting also affect detection.

The development of fluorescent penetrants (dim-light or *brightness-contrast* dyes) have significantly increased PT detection sensitivity. Although the illumination for fluorescent materials is in the ultraviolet (UV) range

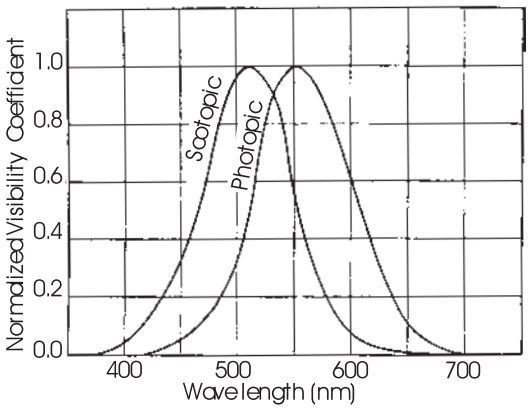


FIGURE 2.5 The frequency response of the human eye for in bright light conditions and in dim light conditions. Each curve is normalized to the maximum sensitivity. If plotted on an absolute intensity scale, the dim light response curve would fall far below the bright light curve. (From Ref. 14 by permission.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

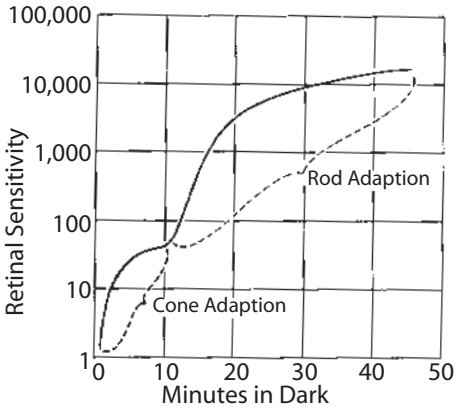


FIGURE 2.6 Dark adaptation of the human eye exposed to bright light conditions for more than 20 min. (From Ref. 16 by permission.)

(normally outside of the visible range of the human eye),* the fluorescence of illuminated material is visible as blue–green light. The atoms of fluorescent material absorb the incident photons of UV-illuminating light, causing the electrons to increase their energy state. These excited atoms lose energy to other molecules in vibrational energy transfer or radiationless internal conversion until they reach an excited state which allows de-excitation with the emission of visible light to their original or ground state, i.e., the energy of the remitted photon is less than that of the absorbed photon. The remitted photon appears in the visible spectrum at blue–green wavelengths (15).

Visibility of fluorescent dyes is proportional to the power of the UV source and its contrast to the background. Because UV light is outside the visible range, its reflection appears black to the naked eye—hence *black light*. Defects are easiest to detect when the blue–green light emanating from the fluorescent penetrant at a discontinuity is the only visible light present. To maintain this high-brightness contrast (blue–green light on a black background) inspectors normally work in specially darkened rooms.

To detect discontinuities using fluorescent penetrants, the inspector's eyes need to adjust to the darkened environment. This so-called *dark adaptation* is not instantaneous. Figure 2.6 plots the time for the human eye to adjust *from* bright-light adaptation *to* low- (dark-) light conditions. Within 1 min, the sensitivity of the eye has increased tenfold in the darkened environment, i.e., the eye can respond to $\frac{1}{10}^{\text{th}}$ the ambient light of the bright-light condition.

*In fact, the *tail* of the bandwidth of most UV light sources falls within the visible spectrum (violet-blue wavelengths).

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

At 10 min, when the eye's sensitivity to light has increased by 80 times, we note an inflection in the curve. The reason for the 10-min boundary is simple. Within the first 10 min of entering the darkened environment, the fast-responding light receptors (*cones*) have reached full sensitivity. Beyond 10 min, the slow-responding light receptors (*rods*) continue to adapt to the low-light intensity, with full sensitivity occurring at about 40–50 min (16).

2.3 TECHNIQUES

Despite advances in penetrant-testing science, the various PT techniques all follow the same basic method as the oil-and-whiting testing of 100 years ago. In part, the method's simplicity accounts for its widespread application. Although the method appears rather simple, qualified procedures must be followed properly and with extreme care, in order to achieve the sensitivity and reliability necessary for critical applications. This section presents details of the method and introduces specifications, codes, and written standards.

2.3.1 Basic Method

The steps for any penetrant test are as follows:

1. Performing a detailed visual examination of the specimen to be tested
2. Preparing specimen surface
3. Verifying the temperature of the specimen and the penetrant materials
4. Applying penetrant to specimen surface
5. Dwell time (time for penetrant to be drawn into the defects)
6. Removing excess penetrant from specimen surface
7. Applying developer (in order to draw penetrant from discontinuities to surface, and to enhance the visibility of the indication)
8. Locating and interpreting indications

Figure 2.7 illustrates these fundamentals. The penetrant is applied to a clean surface (Figure 2.7a). Penetrant enters the discontinuity if the liquid wets the surface of the specimen and capillary action pulls the liquid in during the dwell time (Figure 2.7b). After the appropriate dwell time, the excess penetrant is removed from the surface without disturbing the penetrant in the discontinuity (Figure 2.7c). A developer then coats the surface to draw the penetrant from the discontinuity and to provide a contrasting background for the dye in the penetrant (Figure 2.7d). The penetrant-developer indication can then be observed and interpreted under proper illumination (Figure 2.7e). Figure 2.8 shows a commercial stationary penetrant testing station.

Various PT techniques differ in their way of performing each of these steps. Figure 2.9 presents the process flow charts for four types of liquid PT

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

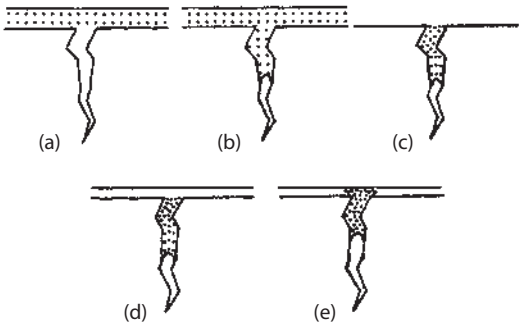


FIGURE 2.7 Schematics of penetrant testing: (a) application of penetrant to clean surface, (b) dwell time (penetrant enters defect), (c) removal of excess penetrant, (d) application of developer to surface, (e) observation and interpretation of indication(s).

systems: water-washable, lipophilic- and hydrophilic-postemulsified, and solvent-removable.

2.3.2 Cleaning

You may have heard that "cleanliness is next to godliness." That old saying is also the golden rule of penetrant testing. Cleaning the specimen before testing (often called "precleaning" [Figure 2.10]) is *vital* to the success of PT, and inadequate or improper cleaning is the most frequent cause of failure to locate flaws.

The penetrant must be able to enter the discontinuity, and so it must not be trapped on the surface by contaminants, foreign materials, or surface anomalies. Cleaning must remove oily films, dirt, grease, rust, paint, scale, slag, and welding flux—as well as the cleaning materials themselves—from the surface. The cleaning method(s) must match the type of material to be removed from the surface material without damaging the material by corrosion or by dissolving it. Common methods include detergent cleaning, solvent cleaning, vapor (and/or dipping) degreasing, descaling, ultrasonic cleaning, and abrasive cleaning (2–6, 9).

Detergent Cleaning

Detergent cleaning, i.e., washing with soap (or a soap-substitute) and water, is the most common method for cleaning superficial materials and contaminants from a surface. Nonabrasive materials such as rags and soft brushes clean most safely, without damaging the surface such that discontinuity openings are sealed. (Wire

Copyright © Marcel Dekker, Inc. All rights reserved.

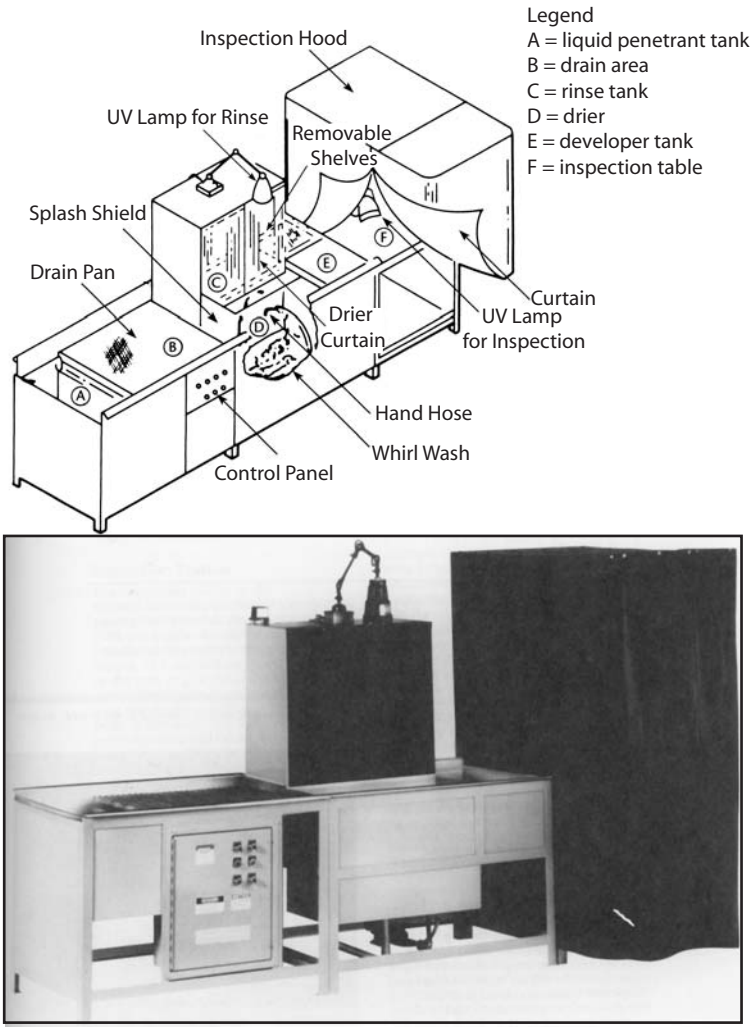


FIGURE 2.8 Schematic and photograph of a commercial liquid penetrant testing system. (Copyright 1999 © The American Society for Nondestructive Testing, Inc. Reprinted with permission from *The Nondestructive Testing Handbook, 2nd ed*. Principles of Liquid Penetrant Testing.)

Copyright © Marcel Dekker, Inc. All rights reserved.

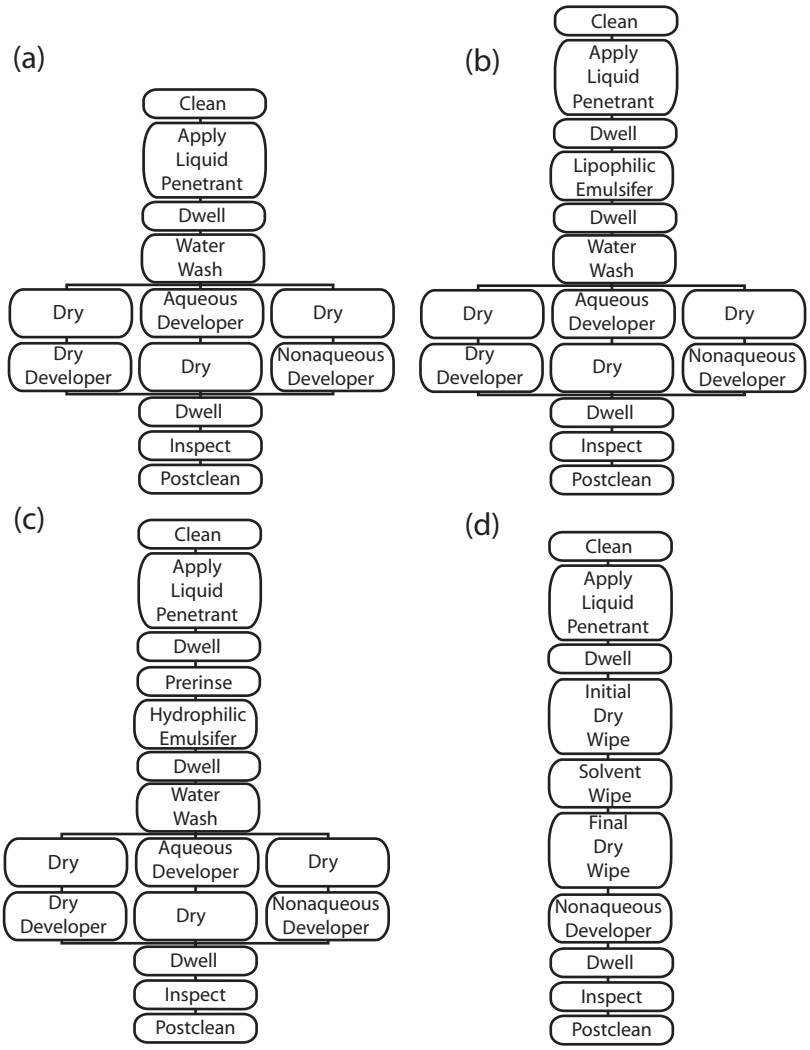


FIGURE 2.9 Penetrant testing flow charts for (a) water-washable, (b) lipophilic and (c) hydrophilic postemulsified, and (d) solvent-removable penetrant systems. (Copyright 1999 © The American Society for Nondestructive Testing, Inc. Reprinted with permission from *The Nondestructive Testing Handbook, 2nd ed*. Principles of Liquid Penetrant Testing.)

Copyright © Marcel Dekker, Inc. All rights reserved.



FIGURE 2.10 Picture of precleaning a specimen. (Courtesy Met-L-Chek.)

brushes may cause such damage and prevent detection of the defects.) Detergents must be compatible with the surface, so as not to cause deterioration of the specimen during or after testing.

Solvent Cleaning

Solvent cleaning, using materials such as alcohol or hydrocarbons, works well for removal of grease and oils; but solvents present environmental and safety issues, such as fire hazards and toxic fumes. Additionally, their chemical compatibility with the specimen must be known. Rags and soft brushes aid in the cleaning.

Vapor (and/or Dipping) Degreasing

Vapor degreasing is a common method of removing grease, oil, and other organic contaminants. In this method, parts are suspended above a hot liquid. Vapor from the liquid enters the defects, condenses, and then returns to its source.

If vapor alone doesn't do the job, you may have to dip the specimens into the hot solvent itself. Dry the surface carefully.

Safety concerns: if you use chlorinated solvents such as perchloroethylene or methyl chloroform, or certain organic liquids such as light hydrocarbons, ketones, or alcohols, for health and safety concerns you must use a sealed container large enough to hold the specimen, or many items at once. Vapor

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

degreasing cannot be used on field specimens, and generally is limited to large production shops. Tanks of special paint-removing solvents may be required for some specimens.

Descaling

Solutions containing hydrochloric, nitric, or hydrofluoric acids—and in some cases, strong alkaline solutions—can remove rust and other oxide scale on metal surfaces. It is important to use extreme care with such materials, and to rinse and dry thoroughly. Residual acid or alkaline solutions may react with the penetrant dyes and "kill" the fluorescence. As usual, descaling materials must be selected carefully, so as not to degrade the specimen. Grinding may not be an appropriate alternative to descaling particularly on softer materials that may smear the material, resulting in closing discontinuity surface openings (making them impossible to detect by PT). In no case should abrasive blasting or peening be performed for the removal of surface incrustations prior to penetrant application.

Ultrasonic Cleaning*

Ultrasonic cleaning may be used in conjunction with any of the above techniques. In this method, the specimen is submerged in a cleaning fluid. To enhance its cleansing ability, the fluid is agitated by ultra-high-frequency, low-amplitude (ultrasonic) waves. (Jewelers routinely use ultrasonically agitated baths to clean jewelry.) Generally, the ultrasonic agitation improves the cleansing quality of the cleaning fluid, and reduces the time required for cleaning. As with all the cleaning methods using solvents that may enter the discontinuities, and especially with ultrasonic cleaning, an adequate drying time to allow complete evaporation of the solvent from the discontinuity must be included in the procedure.

Abrasive Cleaning

> A basic rule of the use of mechanical (abrasive) cleaning is that when they *appear* to be the only way of cleaning surfaces—think again—when this has been done two or three times and the conclusion to use a mechanical method cannot be avoided, start to think of ways of overcoming the problems that mechanical methods cause.
>
> David Lovejoy on the use of abrasive cleaning for PT (17)

Lovejoy's caution embodies the basic problem of abrasive cleaning for PT. Abrasive cleaning abrades (tears) unwanted material (scale, corrosion, etc) from the specimen surface. Such deformation tends to close the discontinuity

*Although ultrasonic agitation of a cleaning fluid primarily takes place beyond the range of human hearing, a high-pitched squeal can be heard from most ultrasonic cleaners.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

openings to the surface, thus reducing the potential of PT detection. In many cases, the abrasive cleaning completely closes the discontinuity opening.

There are two approaches to abrasive cleaning: (a) using specialized abrasives (and expecting that defects will be harder to detect); or (b) after cleaning, chemically etching the surface to expose defects. *Avoid both of these techniques when practical*.

For most metals, lignocellulose abrasives have been used with good results, both with and without a chemical etch. Lignocellulose is a wood-based abrasive (grit) made from fruit stones such as peach, plum, and apricot. Although lignocellulose abrasives are softer than the specimen metal, they still deform the surface. But choosing the proper grit size can minimize deformation with a recognized loss in detection capability (17).

2.3.3 Types of Penetrants

Any good penetrant has two basic functions. It is (a) a highly visible medium that acts as a discontinuity indicator, and (b) a fluid carrier that distributes the indicator over the material surface and into the discontinuity depths. Two major classifications of penetrants are:

Type I: fluorescent penetrants (*brightness contrast*)
Type II: visible or dye penetrants (*color contrast*)

Further classification relates to the technique of penetrant removal:

Water-washable
Lipophilic postemulsified (oil-based)
Solvent wipe
Hydrophilic postemulsified (water-based)

These removal techniques can be applied by dipping, brushing, or spraying. Table 2.3 gives a general classification based on the penetrant type and technique of removing excess penetrant (18).

TABLE 2.3 Penetrant Types and Associated Methods

Type I Fluorescent Dye (Brightness Contrast)		Type II Visible Dye (Color Contrast)	
Process A	Water washable	Process A	Water washable
Process B	Lipophilic Postemulsified		—
Process C	Solvent removed	Process C	Solvent removed
Process D	Hydrophilic Postemulsified		—

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Visible dye or color-contrast penetrants are inspected under normal lighting. A bright red dye provides the colorant most often found in these penetrants, although other colors may be used when they provide a special benefit (usually a color contrast) for the particular specimen. Red dye is most often used due to its particle size, its ability to remain in solution in high concentrations in the penetrating oils used, as well as its relatively high contrast with the white background provided by most developer materials. Figure 2.11 shows visible penetrant being applied. (Note: Red dye should not be used on components that are subject to repeated service-life inspections due to difficulties in removing the red dye from the defects.)

Fluorescent penetrants require a special ultraviolet (UV, or "black light," wavelength 320 to 400 nm) light source. Subdued lighting or even a fairly dark room may be required to obtain the high sensitivity possible with fluorescent penetrants. The dye in fluorescent penetrants absorbs UV light and re-emits the energy as yellow–green light (wavelength 480–580 nm), i.e., as the color in the electromagnetic spectrum to which the human eye is most sensitive under subdued lighting.

Color-contrast or fluorescent penetrant is applied by dipping, brushing, or spraying, depending on the size, shape, number, and location of the specimen(s). Dipping is the preferred method for large specimens, or for large numbers of



FIGURE 2.11 Picture of visible penetrant being applied. (Courtesy Met-L-Chek.)

Copyright © Marcel Dekker, Inc. All rights reserved.

small specimens. The parts must be arranged so that the required surface is covered with penetrant, completely and without air pockets. Dipping is normally accomplished by placing the specimen(s) in a screen or rack that is lowered into a tank of penetrant. The parts are then removed. Note: the parts are left immersed during the penetrant dwell time. Ventilation is important, and care must be taken to prevent contamination of the penetrant bath.

Brushing and spraying are common in field applications. Soft-bristle brushes such as those used for painting work well—but be sure to use brushes *exclusively* for penetrant application. Spraying requires a paint-type spray gun or an aerosol can filled with penetrant. A short distance (6–8 inches) between the aerosol can and the specimen prevents overspraying. Adequate ventilation is important.

2.3.4 Temperature

Temperature of the specimen makes an important difference in the formulation of the penetrant and in the dwell time (see next section). If the surface of the specimen is too hot, the solvent may evaporate before the penetrant can enter the discontinuity. Some manufacturers provide penetrants for three categories:

Ambient to too-hot-to-touch (40°F/4°C) to (125°F/50°C)
Hot specimens (125°F/50°C to about 150°F/70°C)
High temperatures (125°F/50°C) up to about (350°F/180°C)

A warm specimen that cools slightly after penetrant is applied may provide increased sensitivity by sucking penetrant into the defects. On the other hand, a cold specimen may condense water vapor that will interfere with the penetrant. Water may prevent entry or dilute the penetrant in a defect with a resultant loss in detection sensitivity.

2.3.5 Dwell Time

Dwell time (or *penetration time*) has been defined as the time between applying the penetrant and removing the excess, i.e., the time necessary for the penetrant to wet the surface and penetrate the discontinuities. Dwell time is a function of temperature, specimen composition, surface characteristics, size and type of the discontinuity to be detected, and properties of the penetrant. Warm temperatures provide shorter dwell times than do cool temperatures. Most of the common formulations are for ambient temperatures between 50°–120°F (10°–49°C). If the temperature decreases, the dwell time must be extended. A test may have to be examined (or a procedure performance demonstration) to prove that it has the necessary sensitivity at the specific temperature; otherwise, the parts and

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

penetrant may have to be warmed. Dwell time may be lengthened if cleaning is poor or the surface is damaged.

During long dwell times or at high temperatures, the penetrant may dry. In this case, reapplication is necessary.

Metals such as stainless steel, titanium, tungsten, and other high-temperature, high-strength alloys generally require a longer dwell time than the more common ferrous materials. ASTM E-1417 standard suggests a minimum of 10 minutes for normal ambient temperature on common alloys. Special high-strength, high-temperature alloys typically require significantly longer dwell times. Table 2.4 suggests dwell times for penetrants according to (a) process for removing excess penetrant from the specific material, (b) process for manufacturing the material, and (c) type of discontinuity. Some discontinuities, such as intergranular stress corrosion cracking or mechanical fatigue cracks in compression, may require very long penetrant dwell times (>24 hours).

2.3.6 Removing Excess Penetrant

Removing excess penetrant from the specimen surface *without* removing it from the discontinuity may be done in four different ways:

Water washing
Postemulsified lipophilic (oil-based)
Solvent wipe
Postemulsified hydrophilic (water-based)

A client or an industrial standard often specifies which method is to be used, for reasons of specimen materials, geometry, or experience. In each case, a detailed procedure demonstration should be performed to assure the essential parameters, i.e., temperature, water pressure, distance to the surface, wash time, etc., are appropriately established for the items to be tested.

Water-Washable Penetrants

Water-washable penetrants are formulated for easy removal by rinsing with a gentle water spray. Specimens must have relatively smooth surfaces for this method to work well. The water spray may be rather coarse or aerated with water supplied at tap pressures of 20–40 psi. Rough surfaces or improper rinsing will leave some penetrant on the surface, resulting in a pink to red shading of the developer background that makes discontinuity detection and interpretation more difficult. It is an integral part of the procedure to inspect the quality of the wash job visually before continuing to the next step.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

TABLE 2.4 Suggested Liquid Penetrant Dwell (Penetration) Times

		Type I and II penetrants		
		Process (A) water-washable	Process (B or D) postemulsified	Process (C) solvent-wipe
Form	Type of discontinuity	Penetration dwell time[a] (minutes)		
Aluminum				
Castings	Porosity	5–10	5[b]	3
	Cold shuts	5–15	5[b]	3
Extrusions and forgings	Laps	NR[c]	10	7
Welds	Lack of fusion	30	5	3
	Porosity	30	5	3
All	Cracks	30	10	5
All	Fatigue cracks	NR[c]	30	5
Magnesium				
Castings	Porosity	15	5	3
	Cold shuts	15	5	3
Extrusions and forgings	Laps	NR[c]	10	7
Welds	Lack of fusion	30	10	5
	Porosity	30	10	5
All	Cracks	30	10	5
All	Fatigue cracks	NR[c]	30	7
Steel				
Castings	Porosity	30	10[b]	5
	Cold shuts	30	10[b]	7
Extrusions and forgings	Laps	NR[c]	10	7
Welds	Lack of fusion	60	20	7
	Porosity	60	20	7
All	Cracks	30	20	7
All	Fatigue cracks	NR[c]	30	10
Brass and bronze				
Castings	Porosity	10	5[a]	3
	Cold shuts	10	5[a]	3
Extrusions and forgings	Laps	NR[c]	10	7

(*continued*)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

TABLE 2.4 Continued

Form	Type of discontinuity	Type I and II penetrants		
		Process (A) water-washable	Process (B or D) postemulsified	Process (C) solvent-wipe
		Penetration dwell time[a] (minutes)		
Brazed parts	Lack of fusion	15	10	3
	Porosity	15	10	3
All	Cracks	30	10	3
Plastics				
All	Cracks	5–30	5	5
Glass				
All	Cracks	5–30	5	5
Carbide-tipped tools				
	Lack of fusion	30	5	3
	Porosity	30	5	3
	Cracks	30	20	5
Titanium and high-temp. alloys				
All		NR[c]	20–30	15
All metals				
All	Stress or intergranular corrosion	NR[c]	240	240

Source: Marshall Space Flight Center.
[a]For parts at temperature of 15.5°C (60°F) or higher.
[b]Precision castings only.
[c]NR = Not recommended.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Postemulsified Penetrant Wash (19, 20)

Some penetrants are impervious to a water wash and must be made water-washable by adding an emulsifying agent at the end of the penetrant dwell time. Emulsifying agents come in two types, lipophilic (oil soluble) and hydrophilic (water soluble). The type of penetrant will dictate the emulsifying agent. (The type of penetrant will be dictated by specifications, the part geometry, or the size and type of discontinuity to be found.) Because formulations vary among manufacturers, the penetrant-emulsifying agent should be purchased from the same manufacturer as the penetrant-remover system.

Lipophilic Emulsifiers. Lipophilic emulsifiers are dissolved in an oil carrier and used solely with penetrants designed for them. Applied at the end of the penetrant dwell time, lipophilic emulsifiers diffuse into the layer of excess penetrant. The rate of emulsification is governed by the emulsifier's rate of diffusion, and by its solubility in the oil-based liquid penetrant.

The specific dwell time for the emulsifier is critical. (Dwell times range from under 30 seconds for fast-acting low-viscosity emulsifiers to 2–4 minutes for slow-acting, high-viscosity emulsifiers.) The excess penetrant needs enough time to be emulsified for complete removal, including corners and recesses. Too much dwell time, however, and the penetrant will also emulsify within the discontinuities. This causes two adverse effects that reduce flaw sensitivity: (a) emulsified penetrant within the discontinuity opening may be removed during the wash, and (b) the emulsifier can contaminate the penetrant, reducing its ability to indicate a discontinuity.

Removing excess penetrant using lipophilic emulsifiers normally proceeds as follows:

Dipping the part into an emulsifying liquid bath
Removing the part from the bath
Allowing a *drain-off* period (part of the emulsification dwell time), during which the emulsifying agent drains from the part to remove the emulsified liquid penetrant;* and
Applying a water rinse to remove the emulsified penetrant (Figure 2.12a).

During drain-off, the emulsifier mixes mechanically with the penetrant. This mixing is crucial for the emulsifier to diffuse quickly into the penetrant: the mechanical action reduces the diffusion time from hours (static dip only) to a matter of minutes (dip plus draining). The relatively high viscosity of lipophilic emulsifier, combined with the fluid turbulence during drain-off, preferentially

*The drain-off period is equal to the dwell time. Considering both together constitutes the *drain–dwell technique*.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

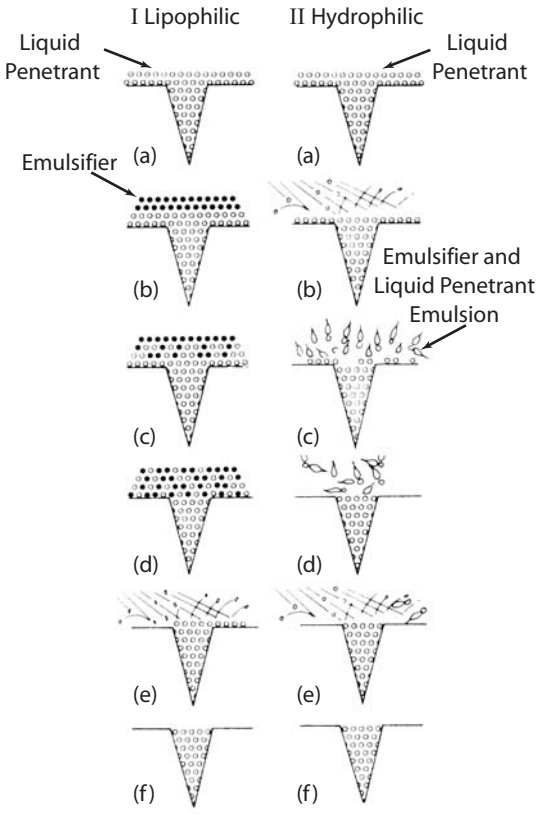


FIGURE 2.12 Penetrant removal process for (I) lipophilic and (II) hydrophilic postemulsifying washes. (Copyright 1999 © The American Society for Nondestructive Testing, Inc. Reprinted with permission from *The Nondestructive Testing Handbook, 2nd ed*. Principles of Liquid Penetrant Testing.)

emulsifies the surface excess penetrant, i.e., reduces the fluid's ability to enter the cracks. To arrest the emulsifying diffusion process, a water wash simply removes the excess (emulsified) penetrant.

Hydrophilic Emulsifiers. Hydrophilic emulsifiers, as the name implies, are water-soluble emulsifiers; they combine a solvent (detergent) and a dispersant (a wetting or surfactant agent) used to remove the penetrant. These emulsifiers remove the penetrant in two simultaneous steps: (a) the solvent dissolves the penetrant, while (b) the dispersant suspends the penetrant in the emulsion, preventing the penetrant from redepositing onto the part surface. Hydrophilic emulsifiers do not operate by diffusion.

Copyright © Marcel Dekker, Inc. All rights reserved.

When hydrophilic emulsifiers are used, the process typically begins with a recommended, but not required, prerinse (Figure 2.12b). Hydrophilic emulsifiers are then applied by either spraying (or immersing) the part. The spraying action (or mechanical agitation of the bath) is critical. Initially, the bulk of the liquid penetrant is removed as the dispersant preferentially wets the surface. The mechanical action of spraying or agitation peels off this bulk layer, allowing the emulsifier to dissolve and suspend the rest of the penetrant. Agitation also ensures that the suspended penetrant is not redeposited onto the part surface. A final rinse arrests the process.

Hydrophilic emulsifiers generally produce brighter flaw indications than lipophilic emulsifiers do. Unlike lipophilic emulsifiers, hydrophilic emulsifying agents tend not to penetrate into the entrapped penetrant in the flaws, thus neither removing nor contaminating it. For these reasons, hydrophilic emulsifiers are recommended for inspecting parts with fine critical discontinuities.

Solvent-aided Penetrant Removal. Removing excess penetrant with the aid of solvents is probably the most common method, especially in field work with portable kits (Figures 2.13 and 2.14). For this method, most of the excess penetrant is removed using clean, dry rags or toweling. The remaining penetrant is then removed by wiping the surface of the specimen with fresh, clean rags or toweling that has been dampened (not saturated) with a solvent appropriate for the penetrant.* It is important that the examiner develop the skills necessary to remove the excess penetrant properly. That is, to assure all the excess penetrant be removed from the surface without removing penetrant from discontinuities.

2.3.7 Types of Developers

After removing the excess penetrant, a thin film of developer is applied to the surface. The developer (a) acts as a blotter to draw entrapped residual penetrant from the flaw recesses; and (b) spreads the penetrant over the surface, increasing the visibility of the indication. The developer's primary ingredient is a powder, the function of which is to make visible the location of the flaws. This powder draws the penetrant from the flaw by a combination of absorption (where the penetrant is drawn into the powder particles) and adsorption (where the penetrant adheres to the surface of the particles).‡

The developer also provides a high-contrast background to increase the visibility of the penetrant. In fluorescent penetrant systems the developer appears blue–black, while the penetrant fluoresces as yellow–green. Visible dyes, on the other hand, are contrasted against a white developer background.

*The solvent may be water, if the penetrant itself has an emulsifier in it.
‡Carbon filters, with their extremely high surface area, operate on the principle of adsorption.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016



FIGURE 2.13 Picture of putting cleaner/penetrant remover on clean toweling. (Courtesy Met-L-Chek.)



FIGURE 2.14 Picture of cleaning excess penetrant from specimen surface. (Courtesy Met-L-Chek.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Developers are classified by the carrier mechanism that facilitates their application:

Dry powders
Water soluble
Water suspendable
Solvent suspendable

The dry powder developer (very light, fluffy, and low density) is normally applied in a cabinet, where the part is exposed to a dynamic cloud of developer created by a flock-type or electrostatic-powder spray gun, or by air blasts into a vat of the powder.

The other methods apply the developer as wet solutions. As you might expect, in water-soluble developers the particles are dissolved in the water carrier, and in water-suspendable developers they are suspended within the water carrier. In the water-suspendable method, agitation of the fluid is required to keep the particles from settling to the bottom of the tank. Wetting agents and corrosion inhibitors are commonly added to water-based developers, to improve consistency of the surface coating and to reduce surface damage.

Solvent- (or *nonaqueous*) suspendable developers suspend the particles in a volatile solution. The solvent aids the developer process by dissolving the penetrant, improving the blotting action of the powder base. As with the water-suspendable method, agitation of the mixture is required. The solvent also dries quickly, unlike aqueous developers, which commonly require active drying in hot air dryers to save time. Solvent developer is commonly applied using spray guns or aerosol cans for field work (Figure 2.15).

The *uniformity* and *thickness* of the developer coating are the two controlling parameters for applying developer. The uniformity provides consistent reading for a given-size flaw throughout the part. If the developer is too thick, the small flaws will not have enough penetrant to permeate through the thickness of the absorbent/adsorbent developer to be visible.

2.3.8 Examination and Interpretation

Examining and interpreting PT test results requires skill and experience. You need to understand the process of PT, and you need to know the types of defects, as well as false indications, you may encounter. Good illumination and good eyesight (testing of eyesight—near vision acuity and color-contrast discrimination is required for inspectors) are also important.

Watching indications develop as a function of time offers insight into the size and depth of discontinuities. Figure 2.16 is a picture of visible penetrant indication on a weld. Shallow discontinuities show up immediately, whereas deeper discontinuities produce indications that grow for a longer period of time.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016



FIGURE 2.15 Picture of developer being sprayed onto specimen. (Courtesy Met-L-Chek.)

The amount of penetrant that becomes visible in the developer depends on the *volume* of the discontinuity.

Test results are generally recorded photographically. Polaroid® pictures are preferred, in order to be certain that the results are really archived and not lost. Video recording is also becoming popular, because it offers low cost, real-time motion recording of the procedures, as well as easy duplication and viewing. For short-term records, the indication may be sprayed with a transparent plastic film, which holds the image for a short time after removal from the specimen. Transparent tape can also be used to "lift" the indication and provide a record.

2.3.9 Illumination

Good illumination is required for quality results. This is true for color contrast or for fluorescent PT. For color contrast daylight, incandescent lights, fluorescent tube lamps, and vapor arc lamps are all satisfactory, if the levels are at least 1000 lux (~98 ft candles). A bright, sunlit day gives about 10,000 lux.

Fluorescent PT requires a darkened area, dark-adapting of the inspectors' eyes, and a sufficiently intense UV lamp (about 8–10 watt/m^2 of UV light). Distance of the specimen from the UV light, and the age of the light, can change the sensitivity of the test. For most critical testing, UV light meters and/or test

Copyright © Marcel Dekker, Inc. All rights reserved.

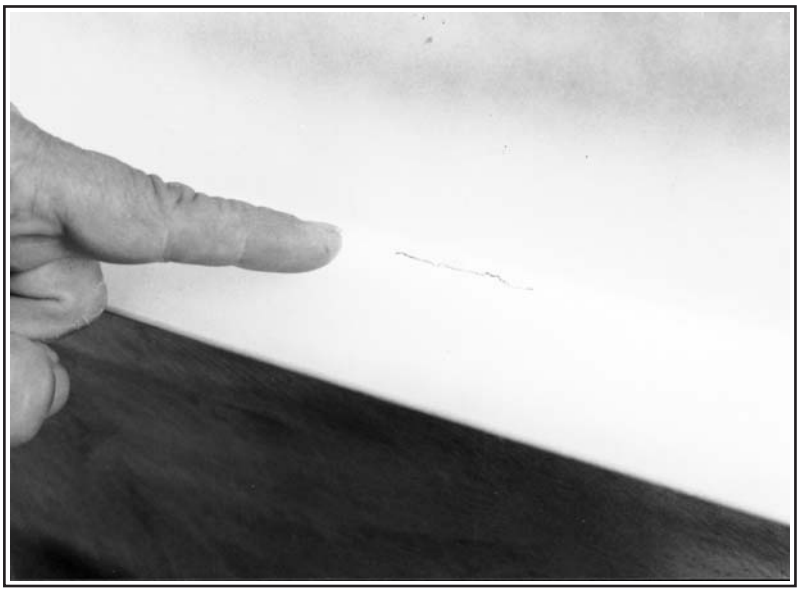


FIGURE 2.16 Picture of a visible penetrant indication of a crack in a weld. (Courtesy Met-L-Chek.)

comparators should be used to demonstrate consistency and sensitivity of the penetrant test. (Also see Section 2.3.3 for additional information on the spectra of illumination and fluorescence.)

2.3.10 Final Cleaning

When PT is complete, the specimen (unless it is to be scrapped) must be cleaned to remove any penetrant materials still on the surface or in crevices or discontinuities. The specimen may also have to be coated with oil, paint or some other protective coating if it is to be stored for a length of time. If the specimen is to be reworked and then retested, cleaning is also necessary to prevent the penetrant and developer from clogging the discontinuities.

2.3.11 Specifications and Standards

Many organizations provide specification and standards for PT and for personnel that perform the testing. Some are governmental agencies, some are technical societies, and some are industrial organizations. The following is a representative listing of some of the more common specifications and standards (3–5, 21).

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

American Society for Testing and Materials

ASTM E-165, E-1220, E-1417, and E-1418 are *standard test methods* for the various PT classifications and procedures. For example, a recent specification from E-1417 (which replaces Mil-Std-6866 for many companies doing work or providing supplies for military systems) requires daily comparisons of penetrant inspections with a *known defect standard* (KDS) (21). Figure 2.17a shows an example of twin KDS panels that have a series of star-shaped cracks along their edge. Figure 2.17b is a closeup photograph of two of the cracks enhanced by penetrant testing. KDS makes it easy to compare two different penetrant procedures or batches of materials.

Aerospace Materials Specification

AMS-2645 addresses materials and equipment for fluorescent penetrant inspection. AMS-2646 covers materials and equipment for visible-dye penetrant inspection. AMS-3155 to 3158 provide specifications for solvent-soluble, water-soluble, high-fluorescence, and water-based penetrant for liquid-oxygen specimen systems.

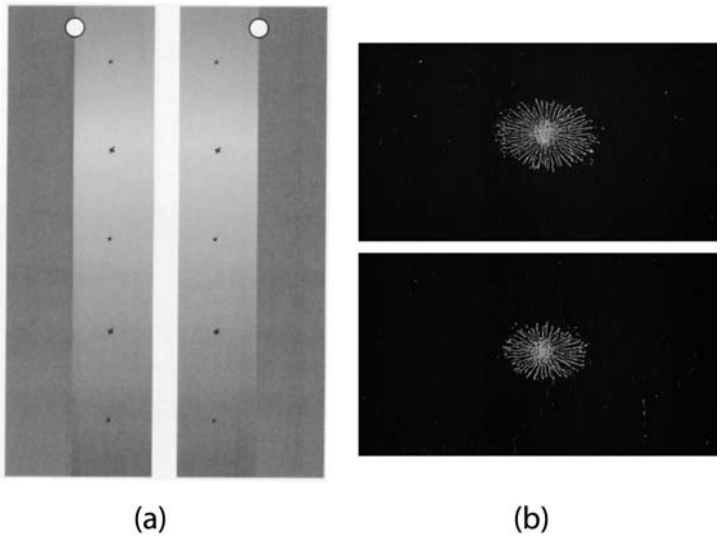


FIGURE 2.17 (a) Twin *Known Defect Standard* (KDS) panels with star-shaped cracks of different sizes; and (b) photographic closeup of two of the cracks shown after fluorescent penetrant inspection. (Courtesy Sherwin, Inc.)

Copyright © Marcel Dekker, Inc. All rights reserved.

American Society of Mechanical Engineers

The ASME Boiler & Pressure Vessel Code gives details on test method, procedure qualifications, limits of contaminants, and procedure requirements.

American Society for Nondestructive Testing

SNT-TC-1A provides recommendations for training, qualification, and certification of personnel for performing PT as well as setting up a testing program.

Department of Defense

In the past, DOD has provided requirements for testing of materials provided to and tested at DOD facilities. Examples: MIL-STD-410A (personnel) (now AIA 410), MIL-I-25135E (penetrant materials), MIL-6866B (method of penetrant inspection), NAVSEA 250-1500-1 (nuclear Navy materials and process specification).

Society of Automotive Engineers

SAE recommends codes (often AMS) for manufacturers and part suppliers inspecting automotive components.

Additional Standard Materials

A "penetrant comparator" metal block (Figure 2.18a) may be used to satisfy ASME Boiler and Pressure Vessel Codes Section III or V that require a comparison of performance of two different penetrant systems or one system at two different conditions such as two different temperatures. Figure 2.18b shows fluorescent penetrant results made at a temperature of 80°F (27°C) on one half of the block and 350°F (27°C) on the other (block has been cut into two pieces). Figure 2.19a is a diagram of a pair of twin-tapered test panels that compare an old batch of penetrant with a new or different batch of penetrant (Figure 2.19b). Such standards are essential to ensure that the materials, procedure, and technique perform as needed and as expected.

2.4 APPLICATIONS (2–5)

To understand PT further, consider these specific examples: the fabrication and aerospace industries, petrochemical plants, automotive and marine manufacture and maintenance, and the electrical power industry.

2.4.1 Fabrication Industries

Fabrication industry processes in which materials are welded, cut, bent, or machined commonly produce surface defects such as cracks and pits. PT provides a rapid, sensitive, and often economical way to inspect a material before it

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

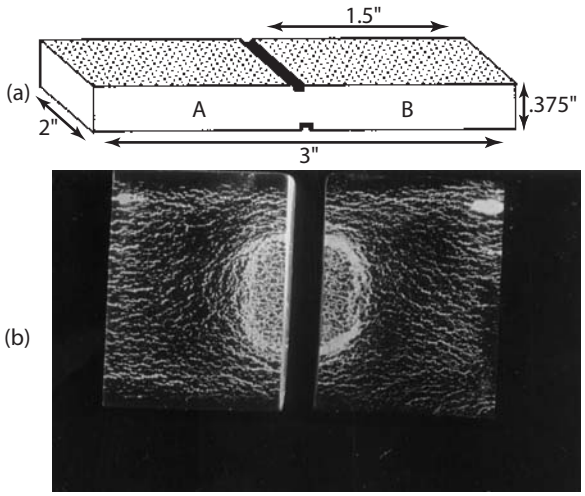


FIGURE 2.18 (a) Diagram of a *Penetrant Comparator* block; and (b) photograph of fluorescent penetrant results. (Courtesy Sherwin, Inc.)

becomes a part of a customer-owned system. Inspections are often preformed on the raw materials themselves (ingots, bars, rods, and sheets of steel or aluminum), as well as after fabrication. Stainless steels are frequently tested by penetrants, since many of them cannot be tested with magnetic particle or flux leakage tests. Figure 2.20 is a photograph of a fluorescent penetrant inspection of a stainless steel weld with crater cracks and pores. Remember that the fluorescent penetrant test is interpreted in a darkened area. Magnesium and magnesium alloys also benefit from penetrant testing.

An important area of PT is the inspection of machine-tool bits: regular grinding of the cutting tools may produce cracks that penetrants can detect, and carbide tools may receive defects from brazing, grinding or lapping during their production. Fluorescent PT (which has the sensitivity to test these tools) saves considerable money and time by detecting cracks and other defects in the tips of machine tools, and gives large savings in scrap costs, enabling machinists to eliminate defective cutting tools before machining begins on new stock.

2.4.2 Aerospace Industries

Much of the aerospace industry uses nonmagnetic, lightweight alloys, which preclude magnetic testing, so PT is greatly depended upon in the manufacture and service testing of aircraft and spacecraft. (The standards for the aerospace industry could make a large book if they were put together.) Failures in such aerospace systems often produce tragic and expensive accidents, so testing is,

Copyright © Marcel Dekker, Inc. All rights reserved.

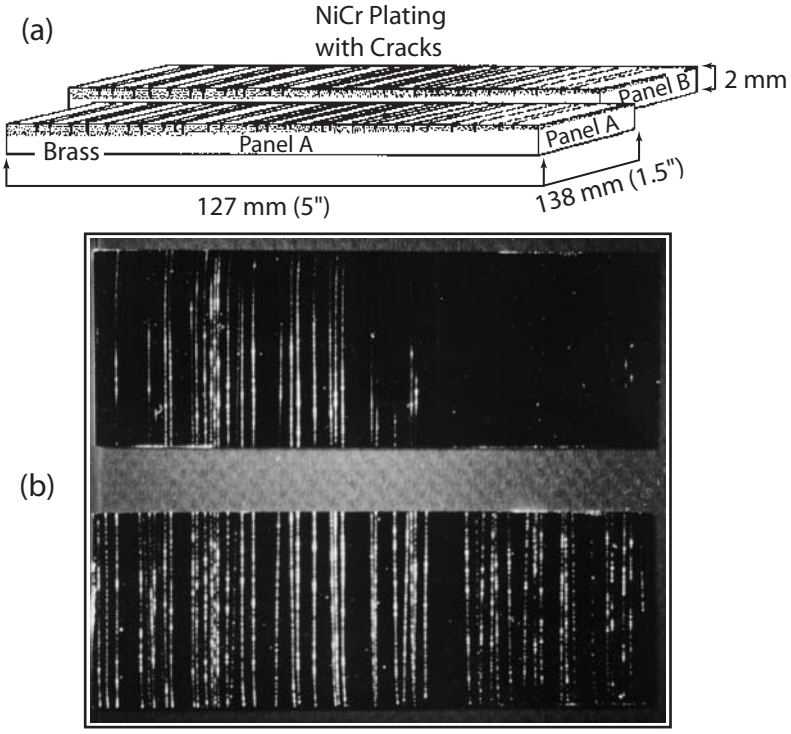


FIGURE 2.19 (a) Diagram of twin tapered test panels with (b) photograph of results using two different penetrants. (Courtesy Sherwin, Inc.)

obviously, essential. PT, detecting cracks from fabrication or from extended use, is one of the very important nondestructive tests used to prevent failures. There is hardly any major structural, highly stressed component (such engine components, landing gear, or the skin of the systems) that is not penetrant-tested regularly. Figure 2.21 shows porosity found in a magnesium casting.

2.4.3 Petrochemical Plants

The nature of the materials in petrochemical plants requires considerable testing of pipes and vessels before and during use. Ease of testing and interpretation are important in this industry; the portability of PT systems allows testing in remote areas at any time. Figure 2.22 is a photograph showing shrinkage cracks in cast piping couplings tested by fluorescent penetrant as viewed in a slightly darkened area.

Copyright © Marcel Dekker, Inc. All rights reserved.

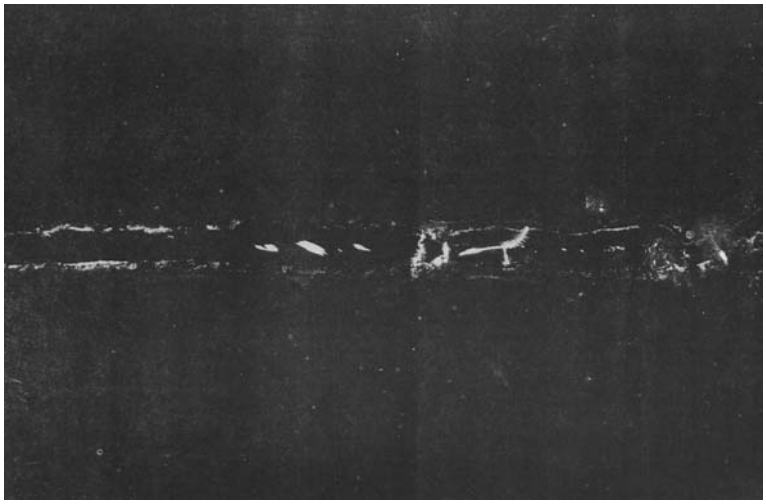


FIGURE 2.20 Photograph of a fluorescent penetrant inspection of a stainless steel weld (taken in a darkened area). (Courtesy Magnaflux.)

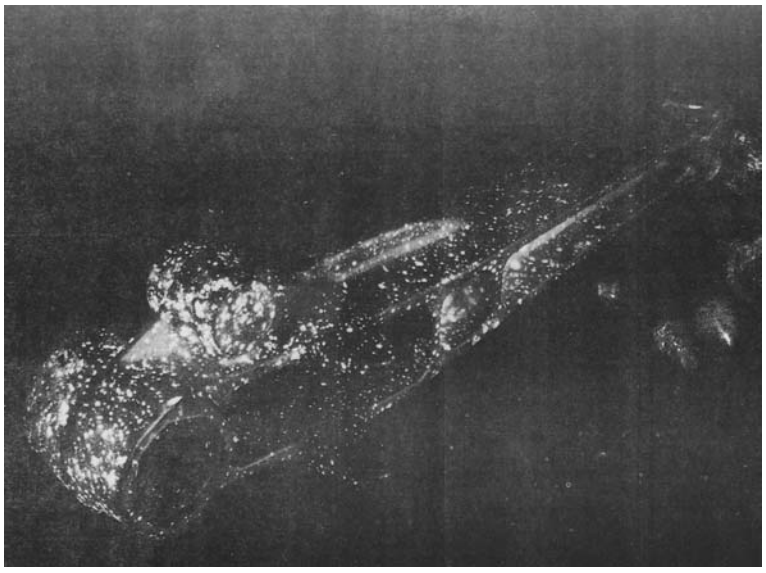


FIGURE 2.21 Photograph of magnesium casting with surface porosity. (Courtesy Magnaflux.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

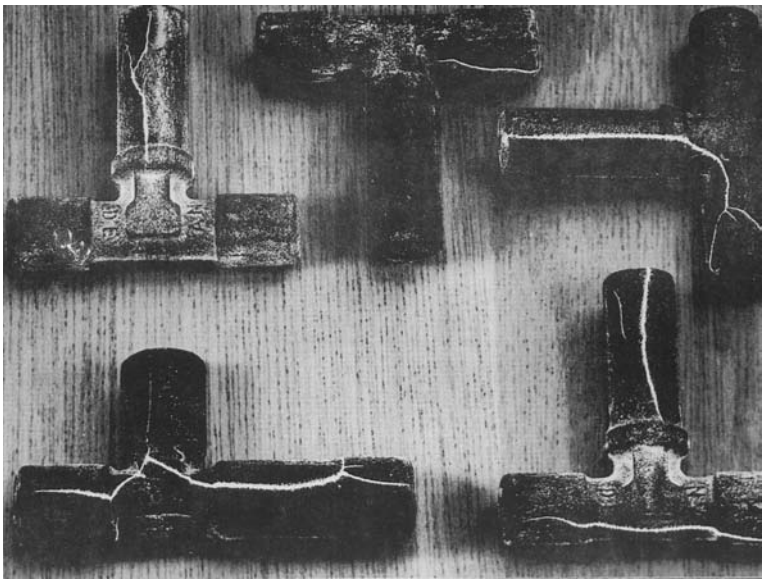


FIGURE 2.22 Photograph of cast piping couplings showing shrinkage cracks. (Courtesy Magnaflux.)

2.4.4 Automotive and Marine Manufacture and Maintenance

The huge numbers of parts for vehicles and ships require fast, easy testing. Completed systems, wherever they may be, require testing as well. PT can test both individual parts and complete systems. Welds in high-stress areas of parts (such as axles) and corrosion of external parts of ships offer considerable opportunities for PT.

2.4.5 Electrical Power Industry

The pressures and temperatures on vessels and piping in power plants necessitate testing both at fabrication and regularly during use. Testing is particularly well documented (not to mention required) for nuclear power plants, which have detailed testing plans. Specifications for these tests often include requirements for special penetrant materials that are free from contaminants, such as halides, that can produce stress-corrosion cracks.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

2.4.6 Other Applications

Ceramic materials that are not porous may be penetrant-tested to detect cracks. Figure 2.23 is a photograph of a fluorescent penetrant inspection of a ceramic coil with thermal cracks, porosity and seams. Expensive, special-use ceramics are generally the ones that require such testing. In ceramics, the time between beginning of a crack and complete failure is very short; if a ceramic part is a major component of a system (e.g., combustion tubes), locating even the tiniest cracks can be extremely important.

Unfired, porous ceramics may also be penetrant-tested, using a *filtered-particle* technique. A filtered-particle penetrant contains many very fine dyed particles. When the penetrant liquid (carrier) enters a discontinuity, the particles

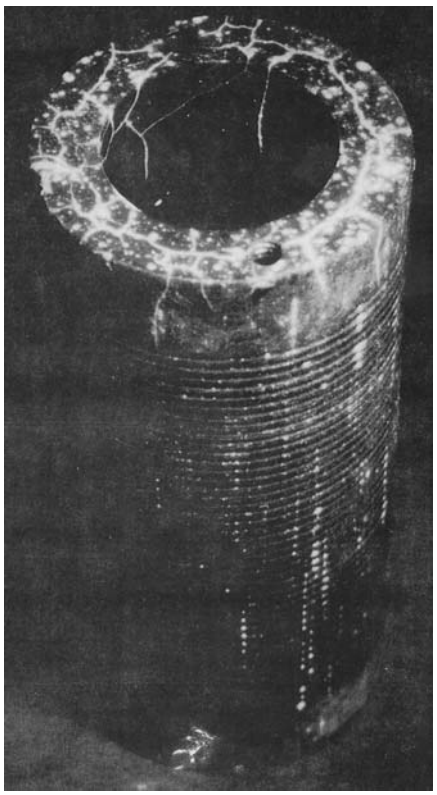


FIGURE 2.23 Photograph of a ceramic coil-form showing thermal cracks, porosity, and seams. (Courtesy Magnaflux.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

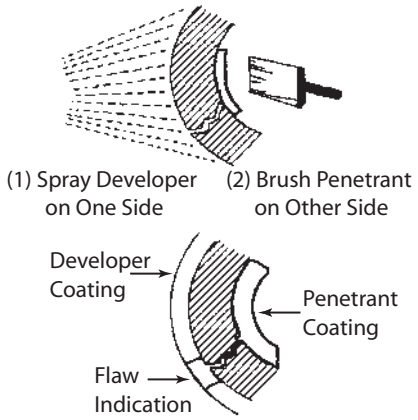


FIGURE 2.24 Diagram of leak testing with penetrant inspection. (Courtesy Met-L-Chek.)

bunch up at the entrance to the discontinuity as they are filtered out, and remain on the surface. Ceramic coils for electrical parts have been tested in this manner.

Leak testing is possible if both sides of the specimen are accessible. The penetrant is sprayed or brushed onto one side of the specimen and pressure is applied to that side, if possible. Developer is sprayed onto the other side of the specimen. The penetrant emerging through the leak is easily seen in the developer (Figure 2.24).

2.5 SUMMARY

Penetrant testing provides a rapid, simple, inexpensive, and sensitive nondestructive testing method for an extremely wide range of materials and systems. Because it is portable, the penetrant method finds use in remote places where testing is vital to safety and continued operations but difficult for application of other testing methods. A great deal of material is available in the open literature and from the manufacturers of the penetrant testing systems and materials to help a novice learn about this important nondestructive test method.

Requirements for achieving good inspection results include experience and good eyesight of the inspector, following good procedures and specifications related to the materials being tested, special cleaning of the specimen, and location of defects open to the surface. When these requirements are fulfilled, penetrant testing produces sensitive tests with outstanding simplicity of procedure.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

PROBLEMS

Introduction

1. What are the advantages and disadvantages of penetrant testing?
2. Describe the basic theory behind penetrant testing.
3. Describe the dye penetrant method as used in the turn of the century in the railroad industry.
4. Describe the two important advances in dye penetrant technology and why they are useful.

Fundamentals

5. What force pushes out in a water droplet and what force pushes in?
6. Why would a water droplet in space be spherical? Why is a raindrop not spherical?
7. What is the contact angle and why is it important for penetrant testing?
8. At what minimum contact angle will the fluid *completely* wet the surface?
9. Typical liquid penetrants have a contact angle of what? Why would you NOT want a large contact angle?
10. What force causes a liquid to enter a narrow opening? Why is this force independent of the orientation of the opening?
11. Name two reasons why Eq. (2.8) is not completely accurate in real world situations.
12. How could one increase penetration depth in a crack?
13. What is dwell time?
14. What factors can effect the dwell time?
15. Why would inspectors use color contrast dye instead of florescent dye in the field?
16. What is dark adaptation and why is it important for florescent penetrant testing?

Techniques

17. What are the eight steps for penetrant testing?
18. Why is cleaning a sample important in dye penetrant testing?
19. Describe in detail two common methods of material cleaning.
20. What are the three most common methods of applying dye penetrant and what are their advantages and disadvantages?
21. What are the four common methods to remove excess penetrant, and what are their advantages and disadvantages.
22. Name the three purposes of a developer in the dye penetrant process.
23. How and why are known defect standards (KDS) used?

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Applications

24. Describe how dye penetrant testing is used in two industries.
25. Why is dye penetrant used in the aerospace industries?
26. Describe the filtered-particle penetrant technique.

GLOSSARY

Black light: Light with a wavelength shorter than 780 nm, which cannot be seen by the human eye (ultraviolet light). Objects illuminated by black light appear to be black in subdued light. See also *UV Lamp*.

Capillarity: The driving force that draws fluids into narrow openings or tubes.

Discontinuity/defect: Any disruption in the uniformity of a material. It may or may not be objectionable, depending on size, location, and shape of the discontinuity and the use made of the material. See also *Flaw.*

Developer: Generally a powder-like material deposited upon a surface to draw a penetrant out of a surface discontinuity. It acts as a blotter and contrasting background.

Dwell time: The time necessary for a penetrant to be in contact with a surface to permit its adequate entry into surface defects.

Emulsifying agent: A substance that allows mixing of hydrophilic (water-like) and lipophilic (oil-like) materials.

Flaw: A discontinuity in a material that prohibits the use of that material for a specific purpose.

Fluorescence: The property of materials to absorb a high-energy, short-wavelength light such as ultraviolet light, and then re-emit the absorbed energy as a longer-wavelength, lower energy light (visible light).

Halogen: The family of elements including fluorine, chlorine, bromine, and iodine.

Hydrophilic: Chemically like water or soluble in water.

Inherent defects: Defects in raw materials as the result of the process of making them.

Lipophilic: Chemically like hydrocarbons or soluble in them.

Penetrant: A formulation that seeps into a surface discontinuity when placed on the surface and then exits the discontinuity into a developer, to form an easily seen indication when properly illuminated.

Stress-corrosion cracks: Cracks formed in austenitic steels exposed to sulfur- or halogen-containing materials and then subjected to stress.

Surface tension: The elastic force acting tangentially to the surface to reduce the surface area to a minimum. Acts to draw a volume of liquid into a sphere.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Ultraviolet (UV) lamp: A lamp that produces principally ultraviolet light that cannot be seen by the human eye. Some visible light often escapes the filter on the lamps and can be seen.

Viscosity: The resistance to flow exhibited by fluids. Lighter fluid has a very low viscosity; glass has an extremely large viscosity.

Wet (wetting ability): The capacity of a fluid to coat or cover a material in a thin layer. If the fluid does not wet the surface, it will tend to ball up on the surface. Wetting indicates strong attractive forces between the molecules in the fluid and the molecules on the surface of the material. Water does not "wet" clean glass or a waxed surface, but water containing a soap or detergent will wet those surfaces and spread out into a thin sheet

VARIABLES

γ	Surface tension
ρ	Fluid density
g	Acceleration of gravity
p	Pressure
F	Force
H	Column height

REFERENCES

1. RC McMaster. Nondestructive Testing Handbook. 1st ed. New York: The Ronald Press Company, 1959, pp 6.1–6.24, 8.1–8.24.
2. RL Pasley. Liquid Penetrants. In: Nondestructive Testing: A Survey. Washington, DC: prepared under contract for National Aeronautics and Space Administration by Southwest Research Institute (NASA SP–5113), 1973, pp 7—26.
3. FA Iddings. Penetrant Testing. In: Survey of NDE. Charlotte, NC: EPRI Nondestructive Evaluation Center, 1982, ch 7, pp 1–10 (IG 1–14).
4. RC McMaster, ed. Nondestructive Testing Handbook. 2nd ed. Vol. 3: Liquid Penetrant Tests. Columbus, Ohio: American Society for Nondestructive Testing, 1982.
5. Anon. The Penetrant Professor. Vol. 1–6. Santa Monica, California: The Met-L-Chek Company, 1993–1999.
6. Anon. Liquid Penetrant Testing: Classroom Training Handbook. Huntsville, AL: prepared by The Convair Division of General Dynamics Corporation for the National Aeronautics and Space Administration (RQA/M1–5330.15), 1967.
7. SJ Robinson, R Goff, and AG Sherwin. Water based penetrants: advantages and limitations. Materials Evaluation 57:9:893–897, 1999.
8. P Hessinger and ML White. Treatment alternatives for liquid penetrant rinse waters. Materials Evaluation 56:8:969–970, 1998.
9. W Rummel. Cautions on the use of commercial aqueous precleaners for penetrant inspection. Materials Evaluation 56:8:950-952, 1998.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

10. IV Savelyev. Physics: A General Course. Moscow: MIR, 1980, pp 376–387.
11. I Asimov. Understanding Physics, Vol. I. New York: Walker and Company, 1966, pp 124–125.
12. DA Porter and KE Easterling. Phase Transformations in Metals and Alloys. London: Chapman and Hall, 1990, p 111.
13. E Hecht. Optics. 2nd ed. Reading, MA: Addison-Wesley, 1990, p 72.
14. TC Ruch. Vision. In: TC Ruch, HD Patton, eds. Physiology and Biophysics. Philadelphia: WB Saunders, 1966, p 414.
15. JG Winans and EJ Seldin. Fluorescence and Phosphorescence. In: EU Condon, H Odishaw, eds. Handbook of Physics. New York: McGraw-Hill, 1958, pp 6–128.
16. AC Guyton. Textbook of Medical Physiology. 4th ed. Philadelphia: WB Saunders, 1971, p 620.
17. D Lovejoy. Penetrant Testing: a Practical Guide. London: Chapman & Hall, 1991, pp 67–68.
18. Nondestructive Testing, V03.03, Section 3. Annual Book of ASTM Standards. Philidelphia: American Society For Testing and Materials, 1998, p 696.
19. NA Tracy, JS Borucki, A Cedillos, B Graham, D Hagemaier, R Lord, S Ness, JT Schmidt, KS Skeie, A Sherwin. Principles of Liquid Penetrant Testing. In: NA Tracy, PO Moore, eds. Nondestructive Testing Handbook: Liquid Penetrant Testing. Society for Nondestructive Testing, 1999, pp 34–81.
20. VG Holmgren, BC Grahm, A Sherwin. Characteristics of Liquid Penetrant and Processing Materials. In: NA Tracy, PO Moore, eds. Nondestructive Testing Handbook: Liquid Penetrant Testing. Society for Nondestructive Testing, 1999, pp. 34–81.
21. SJ Robinson, AG Sherwin. ASTM E-1417 Penetrant System Check, New Requirements, New Test Pieces! Materials Evaluation 57:11:1137–1141, 1999.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

3

Ultrasound

Peter J. Shull
The Pennsylvania State University, Altoona, Pennsylvania

Bernard R. Tittmann
The Pennsylvania State University, University Park, Pennsylvania

3.1 INTRODUCTION

Ultrasonic NDE is one of the most widely used NDE methods today. This chapter discusses ultrasonic theory, transducers, inspection principles, applications, and advanced topics.

Although ultrasonic techniques and theory can be quite complex, the basic concepts behind ultrasonic NDE are simple. Ultrasonic waves (sound waves vibrating at a frequency too high to hear) can propagate in solids—that is, travel through them. (Of course, such waves can also travel through liquids or air.) As the waves travel, they interact with solids in ways that we can predict and represent mathematically. Armed with this kind of understanding, we can create our own ultrasonic waves with a transducer, and thereby use ultrasonics as a way of finding out the nature of a solid material—its thickness, its flaws, its elasticity, and more. As you can imagine, this knowledge has many applications in the aircraft, piping, semiconductor, fabrication, railroad, power, and other industries.

This section is written to give you a basic introduction to ultrasonic testing (UT). It discusses what ultrasonic waves are, and how they propagate; provides a

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

technique overview; reviews the history of ultrasonics; describes applications of UT; and lists advantages and disadvantages of UT.

3.1.1 What Ultrasonic Waves Are, and How They Propagate

If a tree falls in the forest and no one is around, does it make a sound?

The answer to the old riddle about a tree falling in the forest depends on how you define a *sound*. Is a sound (a) the *reception* of a sound wave by the ear, or (b) the *production* of a sound wave by a source (in this case, the falling tree)?

Let's say that sound is the production of a sound wave by a source. Now consider this: What if the falling tree produced sound waves vibrating at a frequency that the human ear could not hear?

That's what we mean by *ultrasonic* waves. Ultrasonic waves are high- ("ultra") frequency sound ("sonic") waves: they vibrate at a frequency above 20,000 vibrations per second, or 20,000 *Hertz* (Hz)—too fast to be audible to humans. Your ear can detect frequencies between about 20 Hz (for example, the sound from a bumblebee's wing) and 17 kHz (the sound from a whistle). Any sound at a higher frequency is ultrasonic—and inaudible. In NDE applications, ultrasonic frequencies typically range from 50 kHz to as high as several GHz.*

What do sounds propagate (travel) *through*? Usually, we think of sounds as propagating through air. For example, when you talk to someone, your vocal cords produce sound waves that travel through the air and are received by your listener's ear. But sound waves also propagate in fluids and solids. In fact, waves travel at higher *velocity*, and with lower *attenuation* (loss of energy), in fluids and solids than in air. That's why you can hear a train coming by listening to the noise in the railroad tracks much sooner than you could by listening for the sound of the train in the air. The velocity of the sound waves in the railroad tracks is more than 15 times that of the sound in air.

This behavior in sound waves—faster in solids, slower in air—is the opposite of what we know about electromagnetic (light) waves. Electromagnetic waves propagate optimally in a vacuum and minimally in solids. Sound waves, on the other hand, cannot propagate in a vacuum at all. So the exploding spaceship in a science fiction movie wouldn't, in reality, make any noise.

Ultrasonic wave propagation in solids is the fundamental phenomenon underlying ultrasonic NDE. We can use the features of an ultrasonic wave (velocity, attenuation) to characterize a material's composition, structure, elastic properties, density, and geometry.

---

*k stands for kilo (1,000), M for mega (1,000,000), and G for giga (1,000,000,000) so that 50 kHz is 50,000 Hz.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

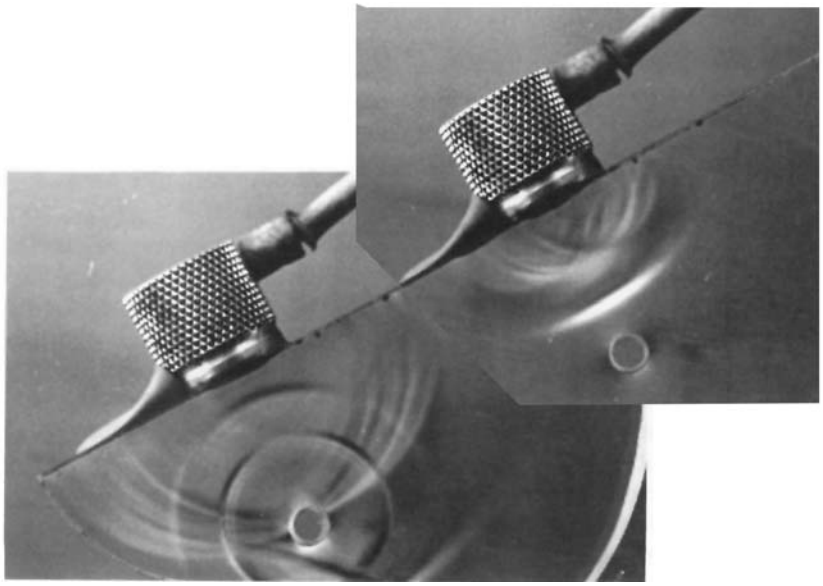


FIGURE 3.1 Ultrasonic reflection off of a circular defect. (KG Hall. Reprinted by permission from Materials Evaluation, 42:922–933. Copyright 1984 ©The American Society for Nondestructive Testing, Inc.)

In addition, we can use ultrasonics to detect and describe flaws in a material. Flaws cause scattering of ultrasonic waves, in the same way a stationary rock reflects a water wave (Figure 3.1). This scattering can be detected as an echo. Using the properties of the echo, we can determine the position, size, and shape of a flaw. In other words, thanks to ultrasonic detection, we know not only that a flaw exists, but also the severity of the damage.

3.1.2 Technique Overview

The basic technique of ultrasonic inspection is simple: *a transducer transforms a voltage pulse into an ultrasonic pulse* (wave). One places the transducer onto a specimen and transmits the pulse into the test object. The pulse travels through the object, responding to its geometry and mechanical properties. The signal is then either (a) transmitted to another transducer (*pitch–catch* method) or (b) reflected back to the original transducer (*pulse–echo* method). Either way, the signal is transformed back into an electrical pulse, which is observed on an oscilloscope.

The observed signal can give a detailed account of the specimen under investigation. Using either method, we can determine:

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

The ultrasonic wave velocity in or thickness of the specimen
The presence of a flaw, defect, or delamination, and its size, shape, position, and composition
Material properties, such as density and elastic constants (Young's Modulus, Poisson's ratio, etc.)
Part geometry or the nature of layered media

Finally, a transducer can be scanned across a surface to create a 2-D or even a 3-D image of a specimen.

3.1.3 History (1–6)

Like many engineering theories with industrial applications, ultrasonics underwent its greatest and most rapid development in the late 19th and 20th centuries. But our understanding of sound waves in general is considerably older. As early as 240 B.C., the Greek philosopher Chrysippus, observing waves in water, speculated that sound took the form of waves. But it was not until the late 16th and early 17th centuries that Galileo Galilei (considered the "Father of Acoustics") and Marin Mersenne developed the first laws governing sound. In 1686, Sir Isaac Newton developed the first mathematical theory of sound, interpreting it as a series of pressure pulses transmitted between particles. Newton's theory allowed for wave phenomena such as diffraction (Figure 3.2). Later expansion of Newton's theory by Euler, Lagrange, and d'Alembert led to the development of the wave equation.

The wave equation allowed sound waves to be represented mathematically. At the end of the 19th century, Lord Rayleigh (John William Strutt, 1842–1919; Figure 3.3), one of the most prominent scientists in ultrasonics (7, 8), discovered Rayleigh surface waves, which are commonly used in many ultrasonic NDE methods. Rayleigh also worked with Lamb, who discovered guided waves in plates, called Lamb waves (9). Almost all of the closed-form solutions for wave propagation were solved in the late 19th century. Since then, computers have

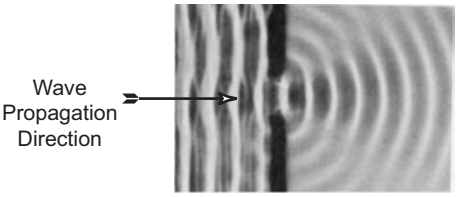


FIGURE 3.2 Diffraction in water tank. Note expansion of wave leaving the aperture. (E. Hecht. Optics. 2nd ed. Reading, MA: Addison-Wesley, 1990, p 393. Reprinted by permission of Pearson Education, Inc.)

Copyright © Marcel Dekker, Inc. All rights reserved.



FIGURE 3.3 Lord Rayleigh. (H. Benson. University Physics. New York: John Wiley, 1991, p 771. Reprinted by permission of John Wiley & Sons.)

allowed us to solve complicated problems, including reflections from defects of various shapes, and problems involving multiple layers.

Although our modern methods are sophisticated, sound waves have been used for NDE for centuries. A simple, time-honored test, tapping an object to see if it sounds "healthy," has given us such expressions as "sound as a bell" and "ring of truth." The test works on the principle that a crack or flaw in an object will change its natural frequency. If you go to a train station today, you can still see a worker with a long-handled hammer, tapping railcar wheels. The trouble with this testing method is that it is sensitive only to flaws large enough to have low (audible) frequencies—the tester has to be able to hear them. To reveal flaws that are smaller, but still critical, it is necessary to inspect with ultrasonic waves.

In other words, only when ultrasonic waves could be easily generated and detected could ultrasonics be accepted as a widely employed NDE method. In 1880, the brothers Curie discovered crystals that could convert ultrasonic energy to electrical energy. In 1881, Lippmann discovered the inverse effect known as the piezoelectric effect (Section 3.3.1), converting electrical energy into ultrasonic energy in the same crystals. After these discoveries, one of the earliest practical ultrasonic endeavors was iceberg detection at sea—developed in 1912 as a direct response to the sinking of the *Titanic*. The success of this application led to other underwater applications, including submarine detection during World War I.

After the war, ultrasonic applications developed quickly. Around 1930, Sokolov presented work using ultrasound to test materials (10). At about the same time, Mulhauser patented an ultrasonic flaw-detection device, using a pitch–catch method employing a separate transmitter and receiver. In 1940, the Iron and Steel Institute, eager to improve both the performance and consistency of their

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

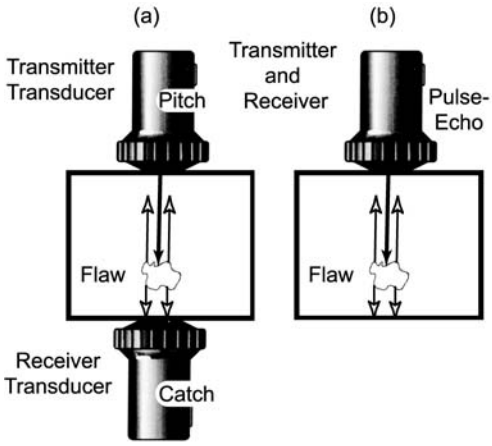


FIGURE 3.4 The two basic methods of transmitting and receiving ultrasound: (a) pitch–catch (separate transducers) and (b) pulse–echo (single transducer). (Courtesy Krautkramer Branson.)

products, developed ultrasonic testing methods for iron and steel. This method also used two transducers (Figure 3.4a).

In the early 1940s, Firestone and Simmons revolutionized ultrasonic testing (11, 12). Among other improvements, they introduced the concept of pulsed ultrasound. (Continuous-wave methods previously employed had extremely low signal-to-noise ratios and were difficult to interpret.) Firestone also introduced the single-transducer approach (Figure 3.4b), in which the transmitter doubled as the receiver (pulse–echo). His method was dubbed "echo-reflection" using a "supersonic reflectoscope" (13). The fundamentals of this work underlie modern ultrasonic testing methods today. Oddly, these ultrasonic methods were initially used only to supplement X-ray inspection. By 1960, however, ultrasonic testing had supplanted, rather than supplemented, many X-ray methods of inspection.

Since these original inventions, ultrasonic technology has advanced greatly. Modern technologies include immersion systems as well as many imaging techniques, including tomographic reconstructions and acoustic microscopes. Another important development is the tone–burst system, which allows the operator to take greater control of testing by specifying the signal to be applied to the transducer.

3.1.4 Applications

Ultrasonic NDE has nearly innumerable applications in the aircraft, piping, semiconductor, fabrication, railroad, power, and medical industries. For example,

Copyright © Marcel Dekker, Inc. All rights reserved.

in the transportation industry (where a catastrophic failure can lead to the deaths of hundreds of people), ultrasonic NDE is used to detect cracks and fatigue damage on safety-critical parts. The rotors of military helicopters, which used to be replaced at regular intervals whether they needed it or not, can now be tested with ultrasonics and replaced only when necessary (retirement for cause), saving both time and money. Ultrasonics can also detect ice formation on aircraft wings, to determine if de-icing is necessary before takeoff—or in midair.

In addition to detecting flaws and determining properties in materials, ultrasonics can also be used for imaging. In medicine, ultrasonic imaging methods, such as the B-scan and tomography, help evaluate fetal development and allow diagnostic imaging of soft body tissue. Ultrasonic imaging is also used during surgery to allow for less invasive surgical techniques (by determining the best place to cut). These imaging techniques are equally widespread in the other industries where NDE is used.

3.1.5 Advantages and Disadvantages

Ultrasonic NDE is a very flexible and robust technique, with applications in a wide range of industries. Only two NDE methods can reveal substantial subsurface flaws in materials: ultrasonic and X-ray methods. (Other techniques, such as magnetic particle and eddy current, can detect some subsurface features, but only near the surface of the material.) Unlike X-ray techniques, however, ultrasonic methods pose no environmental or health risks. In addition, ultrasonic methods offer contacting as well as noncontacting approaches. (Laser generated/detected ultrasound allows standoff distances of up to several meters.) Ultrasonic probes can also be designed to test complex geometries. Finally, ultrasonic detection can be used for all types of materials from biological to metal to ceramics.

But ultrasonic techniques do have some disadvantages. First, they require a highly experienced technician. In addition, although noncontacting methods exist, in the majority of cases the transducer must be in contact with the object, through a water- or gel-coupling layer. Also, ultrasonic waves typically cannot reveal planar flaws (cracks) whose length lies parallel to the direction of wave travel. Finally, ultrasonic methods can be expensive to operate. For a review of advantages and disadvantages, see Table 3.1.

3.2 THEORY

Fluids can move in the various fashions that solids can move. They can undergo translational motion, as when rivers flow or winds blow. They can undergo rotational motion, as in whirlpools and tornadoes. Finally, they can undergo vibrational motion. It is the last that concerns us now, for a vibration can produce a distortion in shape that will travel outward.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

TABLE 3.1 Advantages and Disadvantages of Ultrasonic Testing

Advantages	Disadvantages
High penetration depth	Significant operator training
High sensitivity (detection of minute discontinuities)	Often requires contact using couplant
High accuracy (quantitative flaw sizing and positioning)	Cannot detect planar flaws perpendicular to wavefront
Rapid testing (allows automation and area scans)	Intrinsically a point-by-point measurement
Can test complex geometries	Many geometries cannot be tested
Can measure density and material properties	Can be expensive
Can test on all materials	
Portable	
Safe	

> Such a moving shape-distortion is called a *wave*. While waves are produced in solids, they are most clearly visible and noticeable on the surface of a liquid. It is in connection with water surfaces, indeed, that early man first grew aware of waves.
>
> —Isaac Asimov (14)

To understand ultrasonics, we need to understand wave behavior. In this section you will learn about the different manifestations of waves, as well as vocabulary and mathematical representations.

3.2.1 Introduction to Wave Propagation

A wave is "a disturbance [of a medium] from a neutral or equilibrium condition that propagates without the transport of matter" (15). For example, a stone dropped into a pool of water produces waves whose particles of water are momentarily displaced transverse (up and down) relative to the wave propagating along the surface of the water. Only the *energy* or *momentum* of the wave propagates across the surface of the water. Just as a leaf floating on the surface of the water only moves up and down, and does not travel with a wave, so water particles do not travel with a wave; instead, they oscillate about a neutral point. The difference between particle motion (up and down) and energy propagation (in a direction along the surface) can be confusing at first, since the phenomenon isn't always obvious when we observe fluids. But it's easier to understand when we think of waves traveling through solids: energy propagates, but obviously the particles of the solid don't travel (propagate) through the material.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

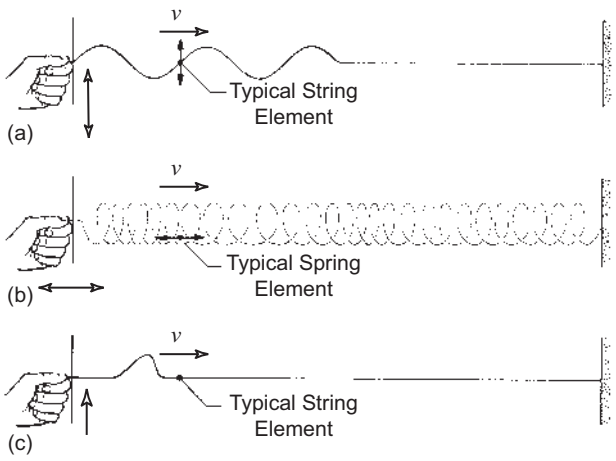


FIGURE 3.5 Example of waves traveling on a string: (a) Transverse wave, (b) longitudinal wave, (c) single transverse pulse. (R Resnick, D Halliday, K Krane. Physics. 4th ed. New York: John Wiley, 1992, p 418. Reprinted by permission of John Wiley & Sons, Inc.)

What is the shape of a wave? You might notice that in water waves, the crests and valleys (the *wave fronts*) are not straight lines (two-dimensional waves) or planes (three-dimensional waves), but instead are curved as they propagate away from their source. Even so, for convenience we commonly assume that wave fronts are straight lines or planar, at least over a short distance of the wave front. This *plane wave* approximation will be used in much of this chapter.

In ultrasonics, we distinguish between two types, or *modes*, of plane waves: (a) *transverse* and (b) *longitudinal*. In the transverse wave mode, particles move perpendicular to the direction of wave propagation. Because the transverse particle motion has an associated shear stress, the transverse wave mode is often called a *shear wave*. Low-viscosity fluids, such as water or air, do not support shear stresses, and thus cannot support shear waves.*

In the longitudinal wave mode, particles move parallel to the direction of wave propagation. The longitudinal wave mode is also called a *pressure wave* or *P-wave*, because the stress of the periodic compression and rarefaction (tension for a solid) of the particles is along the direction of propagation.

Figures 3.5a and 3.5b are diagrams of the periodic particle displacement for transverse and longitudinal plane wave modes, respectively. A traveling single cycle (Figure 3.5c) of either a transverse or a longitudinal mode can be easily

*Surface tension allows nearly transverse waves to propagate on the surface of fluids.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

demonstrated using a Slinky®.* You can create a transverse mode by placing the slinky on a flat surface, holding one end fixed, and with the slinky somewhat stretched rapidly vibrating the other end of the slinky sideways, i.e., the displacement is perpendicular to the direction of wave or energy travel. A longitudinal mode is created by holding one end fixed, the spring taut, and rapidly moving the other end of the spring along the axis of the Slinky® causing displacement in the direction of the wave or energy travel.

Assume a wave propagates in the x direction, then the amplitude of the wave, y, should be a function of x and t (time); or,

$$y = f(x,\ t) \tag{3.1}$$

Ideally, the pulse shape remains constant no matter how far the wave travels. If the pulse travels with velocity v, then at a time t the pulse has moved a distance

$$x' = vt \tag{3.2}$$

For the amplitude at any position on the pulse to remain constant no matter how far it has traveled, the argument of f must subtract the distance the pulse travels, $x' = vt$, from its current position, x:

$$\begin{aligned} y &= f(x,\ t) = f(x - x') \\ &= f(x - vt) \end{aligned} \tag{3.3}$$

Note that v describes the rate at which the pulse moves, and is often referred to as the *phase velocity*; the actual moving mass is described by the *particle velocity*, v_{particle}.

Most ultrasonic waves can be described in terms of harmonic (sinusodial) waves, simply because *any* pulse can be described as a sum of harmonic waves.† Figure 3.6 illustrates the progression of a single cycle of a traveling harmonic wave. To describe the propagation, mark a point on a cycle, e.g., the peak displacement, and follow it as it moves with phase velocity v. A harmonic wave has the following characteristics:

$\tau \equiv$ *period* = the period of time for one complete cycle to move past a stationary observer (Figure 3.6a–c);

$\lambda \equiv$ *wavelength* = the period of space or distance required for *one complete* cycle to occur, assuming *fixed time*. As shown in Figure 3.6a–c, no matter what value of fixed time, the wavelength λ is the same.

Combining these concepts of space and time periods, the wave velocity is:

$$v = \frac{\lambda}{\tau}\left(\frac{\text{m}}{\text{cycles}} \Big/ \frac{\text{s}}{\text{cycles}}\right) = \frac{\lambda}{\tau}\left(\frac{\text{m}}{\text{s}}\right) \tag{3.4}$$

*Slinky® is a common spring-like toy that can be found at most toy stores that was developed by a machinist in Holidaysburg, Pennsylvania.

†This is the subject of Fourier analysis.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

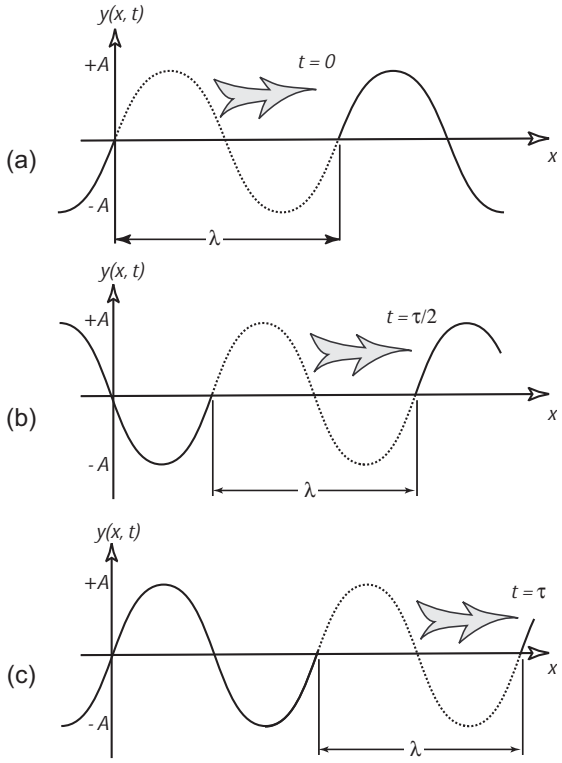


FIGURE 3.6 Progression of a cycle of a traveling wave through a distance of one wavelength. (Adapted from Optics 2nd ed by E Hecht Copyright ©1987, 1974 by Addison-Wesley Publishing Co. Inc. Reprinted by permission of Pearson Education, Inc.)

Several useful variables can be created from the period and the wavelength, including:

$$f \equiv \text{linear temporal frequency} \equiv 1/\tau\left(\frac{\text{cycles}}{\text{s}}\right);$$

$$\chi \equiv \text{linear spatial frequency} \equiv 1/\lambda\left(\frac{\text{cycles}}{\text{m}}\right)^{*};$$

$$\omega \equiv \text{angular temporal frequency} \equiv 2\pi f\left(\frac{\text{radians}}{\text{s}}\right);$$

$$k \equiv \text{angular spatial frequency} \equiv 2\pi\chi\left(\frac{\text{radians}}{\text{m}}\right).$$

*Spatial frequency, χ, can be visualized as the dashed lines of a passing zone on a highway. How many dash/no-dash cycles occur per meter of road? My measurements show $\chi = 0.2$ cycles/m.

Copyright © Marcel Dekker, Inc. All rights reserved.

The velocity of the wave can now be written as $v = \omega/k$(m/s).

A traveling harmonic wave of amplitude A moving with velocity v in the x direction can be given as

$$y(x,\ t) = A\sin[k(x - vt)] \tag{3.5}$$

or in the more common form of

$$y(x,\ t) = A\sin[kx - \omega t] \tag{3.6}$$

The argument $kx - \omega t$ is called the *phase* of the wave function, where $v = dx/dt = \omega/k$ is the *phase velocity.*

In the literature on ultrasonic waves, u instead of y is typically used to indicate particle displacement. For consistency, we will also use u from now on.

To extend our work so far from a single dimension to three dimensions, the particle displacement scalar u becomes the vector quantity

$$\vec{u} = u_x\hat{x} + u_y\hat{y} + u_z\hat{z} \tag{3.7}$$

and the position scalar x is replaced by the position vector

$$\vec{r} = x\hat{x} + y\hat{y} + z\hat{z},\ (\hat{x},\ \hat{y},\ \text{and}\ \hat{z}\ \text{are unit direction vectors}) \tag{3.8}$$

Likewise, the spatial frequency extends to a vector quantity $\vec{k}$ in three dimensions:

$$\vec{k} = k_x\hat{x} + k_y\hat{y} + k_z\hat{z} \tag{3.9}$$

Therefore, the product of the spatial frequency times the position, kx, in three dimensions becomes the vector dot product:

$$\vec{k}\cdot\vec{r} = k_x x + k_y y + k_z z. \tag{3.10}$$

$\vec{k}$, which is *precisely a spatial frequency*, is typically called the *wave vector*. $\vec{k}$ always points in the direction of the phase velocity, i.e., perpendicular to the phase front (a plane of constant phase for a three-dimensional wave). k_x, k_y, and k_z which are the *spatial frequencies* in the directions of $\hat{x}$, $\hat{y}$ and $\hat{z}$, respectively, are typically called *wave numbers*.*

3.2.2 Wave Motion and the Wave Equation

> Ordinarily we emphasize a result rather than a particular derivation of it. In this chapter we take the opposite view. The point here, in a sense, is the derivation itself.... The mathematical physicist has two problems:

*To my knowledge, spatial frequency is not widely recognized as the space-domain counterpart to the (time-domain) temporal frequency. But spatial frequency is a very useful construct for understanding wave propagation, particularly when Fourier analysis is employed.

Copyright © Marcel Dekker, Inc. All rights reserved.

> one is to find solutions, given the equation, and the other is to find the equations which describe a new phenomenon. The derivation here is an example of the second kind of problem.
>
> Richard Feynman (16)

We can describe waves in a number of complementary ways. You have already seen (in Section 3.2.1) a wave described as an undulation traveling through a medium by transferring momentum from one particle to adjacent particles. We described the *physical appearance* of this undulation of particles, through the displacement of the particles u, the wavelength λ, the period τ, the phase velocity v, etc.

Another way to describe waves is to detail the *physical events* of the forces applied to the material and the subsequent response according to the material properties. Richard Feynman succinctly describes the physical events of wave propagation:

> The physics of the phenomenon of sound waves involves three features:
>
> I. The gas (liquid or solid) moves and changes the density.
> II. The change in density corresponds to a change in pressure.
> III. Pressure inequalities generate gas (liquid or solid) motion (17).

The pattern Feynman describes is cyclic—the end of step III is the beginning of step I. A pressure (force) inequality, applied by an external agent, initiates this cyclic pattern. This section develops the *wave equation* that predicts this wave behavior, by showing that the *material properties* dictate the vibrational response to the *applied forces* that drive the wave motion.

Waves in One Dimension

The simplest method of developing the equation that governs the wave motion is (a) to know the complete solution to the equation, and (b) to find what equation the solution satisfies. Although this initial approach does not directly include the material parameters, it readily yields the form of the wave equation. Then, having the form of the wave equation, we will derive the wave equation from Newton's second law of motion, which incorporates the relevant material properties, for specific vibrational modes, i.e., longitudinal and transverse modes.

The wave function for a traveling harmonic wave is*:

$$u(x\,t) = u_0 \cos(\omega t - kx) \tag{3.11}$$

*Hecht [18] explains why $(\omega t - kx)$ is commonly used instead of $(kx - \omega t)$.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

or

$$u(x\,t) = u_0 e^{j(\omega t - kx)} \tag{3.12}$$

Take the second partial differential with respect to space and time of either of these solutions—say, the exponential format

$$\frac{\partial^2 u}{\partial x^2} = -k^2 u_0 e^{j(\omega t - kx)} \quad \text{or} \quad \frac{1}{-k^2}\frac{\partial^2 u}{\partial x^2} = u(x, t) \tag{3.13}$$

$$\frac{\partial^2 u}{\partial t^2} = -\omega^2 u_0 e^{j(\omega t - kx)} \quad \text{or} \quad \frac{-1}{\omega^2}\frac{\partial^2 u}{\partial t^2} = u(x, t) \tag{3.14}$$

Rearranging and equating the two differential equations gives:

$$\frac{\partial^2 u}{\partial x^2} = \frac{k^2}{\omega^2}\frac{\partial^2 u}{\partial t^2} = \frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} = k^2 u(x, t)$$

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} = 0 \tag{3.15}$$

Eq. (3.15) is the *linear wave equation* for one dimension (x), where $v = \omega/k$ is the phase velocity and k is the angular spatial frequency (wavenumber). It is left to the reader to prove that the above equation is also satisfied by a harmonic wave whose form is given by:

$$u(x,\ t) = A\cos(kx + \omega t)$$

If the sign in front of the ω is negative, then this implies that the wave is moving in the negative x direction.

For this chapter, the material will always be assumed to have a constant density (i.e., to be *homogeneous*) and equal elastic constants along every axis (i.e., to be *isotropic*).

Longitudinal Mode (one-dimensional). According to Feynman, pressure causes displacement or deformation of the material. Associated with this deformation must be a stress σ. Consequently, ultrasonic waves are commonly referred to as *stress waves*. For most ultrasonic applications, we are interested in linear elastic stress.*

Assume an ultrasonic wave is propagating along the axis of a long, thin, rod (Figure 3.7). Also, assume the periodic particle displacements (compression and rarefaction) are in the same direction as the direction of propagation, i.e., a longitudinal wave. In the enlarged element of the rod, the solid lines indicate the rod in mechanical equilibrium (no wave). The dotted lines indicate the displace-

*The small effect associated with nonlinear (second-order) elastic constants is an important topic in ultrasonic NDE, but is not addressed in this text.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

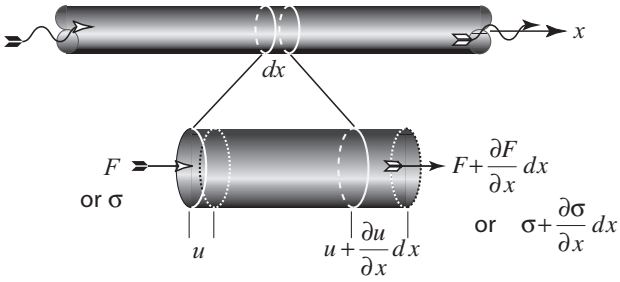


FIGURE 3.7 Ultrasonic longitudinal (compressional) wave propagating in a long thin rod.

ment due to the passage of an ultrasonic wave at time t_0. Additionally, the figure shows the forces (stresses) acting on the deformed element.

The particle movement is caused by the imbalance of forces (stresses) acting on the volume ($A\ \mathrm{d}x = V$):

$$\left(\sigma + \frac{\partial \sigma}{\partial x}\mathrm{d}x\right) - \sigma = \frac{\partial \sigma}{\partial x}\mathrm{d}x \text{ (net stress)} \tag{3.16}$$

$$F_{\text{imbalance}} = \left(\frac{\partial \sigma}{\partial x}\mathrm{d}x\right)A \tag{3.17}$$

Newton's second law of motion, $F = m(\partial v_{\text{particle}}/\partial t)$, states that an imbalance in force will *cause* motion. Substitution of the force-imbalance into Newton's second law of motion yields

$$\begin{aligned}\left(\frac{\partial \sigma}{\partial x}\mathrm{d}x\right)A &= m\frac{\partial v_{\text{particle}}}{\partial t} \\ &= \rho(\mathrm{d}x\,A)\frac{\partial}{\partial t}\frac{\partial u_x}{\partial t} \\ \frac{\partial \sigma}{\partial x} &= \rho\frac{\partial^2 u_x}{\partial t^2}(u_x \text{ implies } x\text{-directed displacement})\end{aligned} \tag{3.18}$$

where ρ is the density of the material. Because we contrived the problem using a long thin rod with a stress-free surface, there are no Poisson's ratio effects. Therefore,

$$\sigma = E\varepsilon, \tag{3.19}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

where

$$\varepsilon = \frac{\Delta l(\text{change in length})}{l_0(\text{original length})} = \frac{\left(u_x + \dfrac{\partial u_x}{\partial x}\mathrm{d}x\right) - u_x}{\mathrm{d}x} = \frac{\partial u_x}{\partial x}(\textit{uniaxial}\ \text{strain})$$

E − Young′s Modulus (engineering elastic constant for uniaxial stress−strain).

Substituting this stress–strain relationship into Eq. (3.18), we have

$$\begin{aligned} \frac{\partial \sigma}{\partial x} &= E\frac{\partial \varepsilon}{\partial x} = E\frac{\partial}{\partial x}\frac{\partial u_x}{\partial x} = \rho\frac{\partial^2 u_x}{\partial t^2} \\ \frac{\partial^2 u_x}{\partial x^2} &= \frac{\rho}{E}\frac{\partial^2 u_x}{\partial t^2} \end{aligned} \tag{3.20}$$

Note that the above equation is of exactly the same form as the wave equation (Eq. [3.15]) except that there is a ρ/E term instead of a $1/v^2$ term. Therefore, the phase velocity of a 1-D longitudinal wave in a homogeneous, isotropic medium (compare Eq. [3.20] to Eq. [3.15]) is:

$$v_l = \sqrt{\frac{E}{\rho}} \tag{3.21}$$

Thus, the phase velocity is determined *solely* from the density and the elastic properties of the material.

Transverse Mode (one-dimensional). A transverse wave is excited by applying time-varying forces to produce linear, elastic, *shear stresses*. Consequently, all vibrational motion is *perpendicular* to the direction of travel of the wave (phase front). Intuitively, one expects the phase velocity for a transverse wave v_t to be related to the elastic shear modulus μ.

Figure 3.8 depicts a shear wave traveling in the x-direction. The forces on the faces of the enlarged element act only in the $\pm y$-direction while the transfer of momentum (energy) and the phase velocity are in the x-direction.* This configuration of forces on the element is called *simple shear* (simple shear is the sum of pure shear plus a local rigid rotation: the particles simply deform and rotate) (19).

*Momentum transfer is not always in the same direction as phase velocity. If the material has different elastic constants for different directions (i.e., is *anisotropic*), then the wave momentum may be steered in a direction other than the direction of constant phase.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

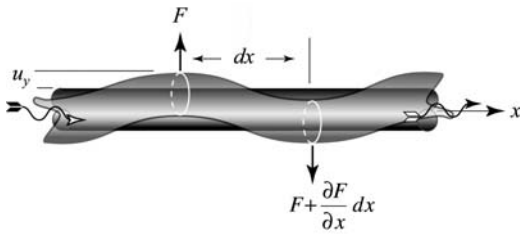


FIGURE 3.8 Ultrasonic transverse (shear) wave propagating in a long thin rod.

The imbalance in the shear stresses (forces),

$$\left(\sigma_{\text{shear}} + \frac{\partial \sigma_{\text{shear}}}{\partial x}\mathrm{d}x\right) - \sigma_{\text{shear}} = \frac{\partial \sigma_{\text{shear}}}{\partial x}\mathrm{d}x \text{ (net shear stress)} \tag{3.22}$$

$$F_{\text{imbalance}} = \left(\frac{\partial \sigma_{\text{shear}}}{\partial x}\mathrm{d}x\right)A \tag{3.23}$$

causes the y-directed displacements u_y.

Applying Newton's second law of motion:

$$\begin{aligned}\left(\frac{\partial \sigma_{\text{shear}}}{\partial x}\mathrm{d}x\right)A &= m\frac{\partial v_{\text{particle}}}{\partial t} = \rho(\mathrm{d}x\,A)\frac{\partial^{\partial u_y/\partial t}}{\partial t^2} \\ \frac{\partial \sigma_{\text{shear}}}{\partial x} &= \rho\frac{\partial^2 u_y}{\partial t^2}\end{aligned} \tag{3.24}$$

The relationship between shear stress and shear strain for linearly elastic material is:

$$\sigma_{\text{shear}} = \mu\gamma_{\text{shear}} \tag{3.25}$$

where

$$\begin{aligned}\gamma_{\text{shear}} &= \frac{\partial u_y}{\partial x}(\textit{engineering}\text{ shear strain})^* \\ \mu &= \text{shear modulus for isotropic, homogeneous material}\end{aligned} \tag{3.26}$$

Substituting this stress–strain relationship into Eq. (3.24):

$$\frac{\partial \sigma_{\text{shear}}}{\partial x} = \mu\frac{\partial^{\partial u_y/\partial x}}{\partial x} = \rho\frac{\partial^2 u_y}{\partial t^2}$$

$$\frac{\partial^2 u_y}{\partial x^2} = \frac{\rho}{\mu}\frac{\partial^2 u_y}{\partial t^2}$$

*The difference between *engineering* shear strain and *shear* strain is presented in the following section.

Copyright © Marcel Dekker, Inc. All rights reserved.

Note that Eq. (3.27) is of exactly the same form as the wave equation (Eq. (3.15)), except that it includes a ρ/μ term instead of a $1/v^2$ term. Therefore, the phase velocity of a 1-D transverse wave in a homogeneous, isotropic medium (compare Eq. (3.27) to Eq. (3.15)) is:

$$v_t = \sqrt{\frac{\mu}{\rho}} \tag{3.28}$$

For materials with a negligible shear modulus, such as water, the shear velocity is approximately zero.

Waves in Three Dimensions (bulk waves)

Now let's expand the problem from one-dimension to a wave propagating in three dimensions in a material without boundaries—a bulk wave traveling in a bulk material. Again, we assume a planar wave front and an isotropic, homogeneous material. The major difference between bulk waves and one-dimensional waves is that in bulk waves, due to Poisson's effect, when the substance compresses in one direction (say the x-direction) the material expands in the other directions (y- and z-directions). Poisson's ratio, v, relates the *applied* strain to the *induced* strain*

$$v = -\frac{\varepsilon_y}{\varepsilon_x} = -\frac{\varepsilon_z}{\varepsilon_x} \tag{3.29}$$

Consequently, the one-dimensional, longitudinal wave velocity $v_l = \sqrt{E/\rho}$ must be modified to include the volumetric changes that occur in bulk material:

$$v_l = \sqrt{\frac{E}{\rho}\left(\frac{1-v}{(1+v)(1-2v)}\right)} \tag{3.30}$$

This bulk longitudinal wave changes the volume of the volume elements. Therefore it is often called a *dilatational wave*: a wave that dilates the volume.

In contrast, the shear deformation in bulk material causes no new elastic effects, so the volume does not change. Thus there is no difference between the shear wave velocity that propagates in bulk material and a long thin rod:

$$v_t = \sqrt{\frac{\mu}{\rho}} \tag{3.31}$$

Table 3.2 lists the longitudinal and shear velocities for common material. Note that the longitudinal velocity in bulk material (dilatational) is always faster than the longitudinal velocity in a thin rod. Poisson's ratio for an ideal material is

*Poisson's ratios for the two lateral directions are equal for isotropic material. In general, however, they are not equal.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

TABLE 3.2 List of Density, Velocity, and Specific Acoustic Impedance for Selected Materials

Material	Density g/cm^3	Velocity of Sound: v_l (longitudinal) km/s or mm/μs	Velocity of Sound: v_t (transverse) km/s or mm/μs	Acoustic Impedance $Z = \rho v_l$ 10^6 kg/m^{2}s
A. Metals				
Aluminium	2.7	6.32	3.13	17
Al 2117 T4	2.8	6.5	3.1	18.2
Berylium	1.78	12.9	8.9	22.9
Bismuth	9.8	2.18	1.10	21
Brass (Naval)	8.42	4.43	2.1	1.95
Brass (58)	8.4	4.40	2.20	37
Bronze	8.86	5.53	2.2	2
Cadmium	8.6	2.78	1.50	24
Cast iron	6.9–7.3	3.5–5.8	2.2–3.2	25–42
Constantan	8.8	5.24	2.64	46
Copper	8.9	4.70	2.26	42
German silver	8.4	4.76	2.16	40
Gold	19.3	3.24	1.20	63
Inconel	8.25	5.72	3	2.79
Indium	7.3	2.22	–	16.2
Stellite	11–15	6.8–7.3	4.0–4.7	77–102
Iron (steel)	7.7	5.90	3.23	45
Steel Stainless	7.89	5.79	3.1	45.7
Lead	11.4	2.16	0.70	25
Magnesium	1.7	5.77	3.05	10
Manganin	8.4	4.66	2.35	39
Mercury	13.6	1.45	–	20
Monel	8.83	62	2.7	1.96
Nickel	8.8	5.63	2.96	50
Platinum	21.4	3.96	1.67	85
Silver	10.5	3.60	1.59	38
Tin	7.3	3.32	1.67	24
Tungsten	19.1	5.46	2.62	104
Zinc	7.1	4.17	2.41	30
B. Nonmetals				
Aluminium oxide	3.6–3.95	9–11	5.5–6.	32–43
Brick	3.6	3.65	2.6	15.3
Concrete	2.6	3.1	–	8.1

(*continued*)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

TABLE 3.2 (*continued*)

Material	Density g/cm^3	Velocity of Sound: v_l (longitudinal) km/s or mm/μs	Velocity of Sound: v_t (transverse) km/s or mm/μs	Acoustic Impedance $Z = \rho v_l$ 10^6 kg/m^{2}s
Delrin	1.42	2.52	–	3.58
Epoxy resin	1.1–1.25	2.4–2.9	1.1	2.7–3.6
Glass, flint	3.6	4.26	2.56	15
Glass, crown	2.5	5.66	3.42	14
Glass, pyrex	2.24	5.64	3.3	12.6
Human tissue	–	1.47	–	
Ice	0.9	3.98	1.99	3.6
Paraffin wax	0.83	2.2	–	1.8
Acrylic resin (Perspex)	1.18	2.73	1.43	3.2
Polyamide (nylon, perlon)	1.1–1.2	2.2–2.6	1.1–1.2	2.4–3.1
Polystyrene	1.06	2.35	1.15	2.5
Porcelain	2.4	5.6–6.2	3.5–3.7	13
Quartz glass (silica)	2.6	5.57	3.52	14.5
Rubber, soft	0.9	1.48	–	1.4
Rubber, vulcanized	1.2	2.3	–	2.8
Polytetrafluoroethylene (Teflon)	2.2	1.35	0.55	3.0
Scotch tape	1.16	1.9	–	2.2
Silicon rubber	1.18	1.05	–	1.77
C. Liquids				
Diesel oil	0.80	1.25	–	1.0
Glycerine	1.26	1.92	–	2.5
Methylene iodide	3.23	0.98	–	3.2
Motor car oil (SAE 20 a. 30)	0.87	1.74	–	1.5
Nitrogen	0.8	0.86	–	0.69
Olive oil	–	1.43	–	–
Turpentine	0.87	1.25	–	1.24
Water (at 20°C)	1.0	1.483	–	1.5

Source: From Ref. 50.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

0.5, with typical measured values around 0.3. Also, note that the shear waves travel at about one-half the velocity of a longitudinal wave.

Wave velocities for isotropic materials are commonly described in terms of the measured engineering parameters ρ, E, and v. Mathematically, however, it is often easier to manipulate the elastic wave equations if the stress–strain relationships are written using the *Lamé constants*, λ^* and μ:

$$\lambda = \frac{vE}{(1+v)(1-2v)} \tag{3.32}$$

$$\mu = \frac{1}{2}\frac{E}{1+v} \tag{3.33}$$

Therefore, one can rewrite the longitudinal and shear velocities as:

$$v_l = \sqrt{\frac{\lambda + 2\mu}{\rho}} \tag{3.34}$$

$$v_t = \sqrt{\frac{\mu}{\rho}} \tag{3.35}$$

Displacement/Deformation in Bulk Material (optional). Let's consider bulk waves in more precise mathematical detail. To understand deformation and displacement, consider an automobile accident. A car fender can be deformed both plastically (permanently) and elastically (temporarily) under the force of an opposing vehicle or tree. The car also experiences rigid body displacement, as either rotation or linear displacement. The total displacement of any body is the sum of deformation and rigid body displacements. For our problem of ultrasonic waves in nondestructive applications, we are interested only in elastic deformation and rigid body rotational displacements (19). Rigid body rotation in ultrasonics takes the form of *local* rigid rotations. *Linear* rigid body displacement is of no importance in ultrasonics.

Imagine a volume element of material. The *normal strain* (relative displacement) in the x-direction on the x-face is:

$$e_{xx} = \frac{\partial u_x}{\partial x} \tag{3.36}$$

and the normal strains on the other two faces are

$$e_{yy} = \frac{\partial u_y}{\partial y} \tag{3.37}$$

$$e_{zz} = \frac{\partial u_z}{\partial z} \tag{3.38}$$

*This λ is distinct from the wavelength λ.

Copyright © Marcel Dekker, Inc. All rights reserved.

The relative angular displacement (combined shear deformation and rigid rotational displacements) on the three faces are:

$$\begin{aligned}
e_{xy} &= \frac{\partial u_x}{\partial y}(x\text{-directed force, } y\text{-face}) \\
e_{xz} &= \frac{\partial u_x}{\partial z}(x\text{-directed force, } z\text{-face}) \\
e_{yx} &= \frac{\partial u_y}{\partial x}(y\text{-directed force, } x\text{-face}) \\
e_{yz} &= \frac{\partial u_y}{\partial z}(y\text{-directed force, } z\text{-face}) \\
e_{zx} &= \frac{\partial u_z}{\partial x}(z\text{-directed force, } x\text{-face}) \\
e_{zy} &= \frac{\partial u_z}{\partial y}(z\text{-directed force, } y\text{-face})
\end{aligned} \tag{3.39}$$

Note that you cannot differentiate how much of the angular displacement is due to shear deformation and how much is due to local rigid rotation. Fortunately, mathematically, e_{ij} can be separated into shear (deformation) strain and rigid rotation:

$$\begin{aligned}
e_{ij} &= \frac{1}{2}(e_{ij} + e_{ji}) + \frac{1}{2}(e_{ij} - e_{ji}) \\
e_{ij} &= \text{shear strain } + \text{ rigid rotation}
\end{aligned} \tag{3.40}$$

(The i and j are dummy indices for the x-, y-, and z-direction coordinates.) Short hand notations for these are

$$\text{shear strain: } \varepsilon_{ij} = \frac{1}{2}(e_{ij} + e_{ji}) \tag{3.41}$$

$$\text{rigid rotation: } \varpi_{ij} = \frac{1}{2}(e_{ij} - e_{ji}) \tag{3.42}$$

These quantities that represent the displacement of a material can be succinctly written in matrix form:

$$\begin{aligned}
\varepsilon_{ij} &= \begin{vmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{yx} & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_{zz} \end{vmatrix} \\
&= \begin{vmatrix} \dfrac{du_x}{dx} & \dfrac{1}{2}\left(\dfrac{\partial u_x}{\partial y} + \dfrac{\partial u_y}{\partial x}\right) & \dfrac{1}{2}\left(\dfrac{\partial u_x}{\partial z} + \dfrac{\partial u_z}{\partial x}\right) \\ \dfrac{1}{2}\left(\dfrac{\partial u_y}{\partial x} + \dfrac{\partial u_x}{\partial y}\right) & \dfrac{du_y}{dy} & \dfrac{1}{2}\left(\dfrac{\partial u_y}{\partial z} + \dfrac{\partial u_z}{\partial y}\right) \\ \dfrac{1}{2}\left(\dfrac{\partial u_z}{\partial x} + \dfrac{\partial u_x}{\partial z}\right) & \dfrac{1}{2}\left(\dfrac{\partial u_z}{\partial y} + \dfrac{\partial u_y}{\partial z}\right) & \dfrac{du_z}{dz} \end{vmatrix}
\end{aligned} \tag{3.43}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

and

$$\varpi_{ij} = \begin{vmatrix} \varpi_{xx} & \varpi_{xy} & \varpi_{xz} \\ \varpi_{yx} & \varpi_{yy} & \varpi_{yz} \\ \varpi_{zx} & \varpi_{zy} & \varpi_{zz} \end{vmatrix}$$

$$= \begin{vmatrix} 0 & \frac{1}{2}\left(\frac{\partial u_x}{\partial y} - \frac{\partial u_y}{\partial x}\right) & \frac{1}{2}\left(\frac{\partial u_x}{\partial z} - \frac{\partial u_z}{\partial x}\right) \\ \frac{1}{2}\left(\frac{\partial u_y}{\partial x} - \frac{\partial u_x}{\partial y}\right) & 0 & \frac{1}{2}\left(\frac{\partial u_y}{\partial z} - \frac{\partial u_z}{\partial y}\right) \\ \frac{1}{2}\left(\frac{\partial u_z}{\partial x} - \frac{\partial u_x}{\partial z}\right) & \frac{1}{2}\left(\frac{\partial u_z}{\partial y} - \frac{\partial u_y}{\partial z}\right) & 0 \end{vmatrix} \tag{3.44}$$

(The first index denotes the direction of the force, and the second indicates the face subjected to the force.) Note that $\varepsilon_{ij} = \varepsilon_{ji}$ and $\varpi_{ij} = -\varpi_{ji}$.

This definition of the normal strains ε_{xx}, ε_{yy}, and ε_{zz} is the same as that used to calculate the one-dimensional longitudinal wave velocity. For the transverse wave velocity, however, we used *engineering shear strain* γ, not shear strain ε_{ij}. Engineering shear strain represents experimental results of total shear displacements and does not detail the form of the displacement. Imagine that an element undergoes an arbitrary amount of angular strain (Figure 3.9a, depicting simple shear). The total shear strain is $e_{xy} + e_{yx}$. (Note that $e_{yx} = 0$ for simple shear. Nonetheless the total strain is $e_{xy} + e_{yx}$). This total is the engineering strain γ.

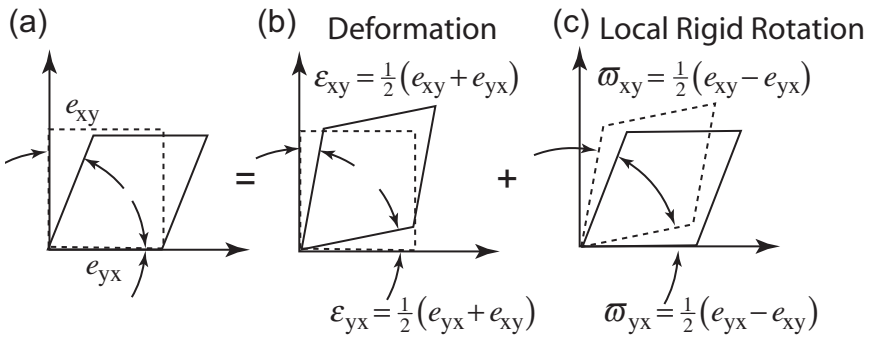


FIGURE 3.9 (a) General deformation of an elemental volume equals (b) a pure strain plus (c) a rigid rotation.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Figures 3.9b and 3.9c show how this arbitrary shear deformation can be represented as equal shear strains ε_{xy} and ε_{yx} plus rigid rotations ϖ_{xy} and ϖ_{yx}.

$$\begin{aligned}\gamma \equiv e_{xy} + e_{yx} &= (\varepsilon_{xy} + \tilde{\omega}_{xy}) + (\varepsilon_{yx} + \tilde{\omega}_{yx}) \\ &= (\varepsilon_{xy} + \varepsilon_{yx}) + (\tilde{\omega}_{xy} + \tilde{\omega}_{yx})\end{aligned} \tag{3.45}$$

Recall $\varepsilon_{ij} = \varepsilon_{ji}$ and $\tilde{\omega}_{ij} = -\tilde{\omega}_{ji}$.

$$\begin{aligned}\gamma &= 2\varepsilon_{xy} \\ \gamma &= 2\varepsilon_{ij}\end{aligned} \tag{3.46}$$

Therefore,

$$\sigma_{ij} = \mu\gamma = 2\mu\varepsilon_{ij} \tag{3.47}$$

Now, let's put it all together—stress, strain, and Poisson effects. Normal stress–strain:

$$\begin{aligned}&\varepsilon_{ii} \leftrightarrow \text{1-D} + \text{Poisson Ratio Effects} \\ &\varepsilon_{xx} = \frac{1}{E}\sigma_{xx} - \frac{1}{E}(v\sigma_{yy} + v\sigma_{zz}) = \frac{1+v}{E}\sigma_{xx} - \frac{v}{E}(\sigma_{xx} + \sigma_{yy} + \sigma_{zz}) \\ &\varepsilon_{yy} = \frac{1}{E}\sigma_{yy} - \frac{1}{E}(v\sigma_{zz} + v\sigma_{xx}) \\ &\varepsilon_{zz} = \frac{1}{E}\sigma_{zz} - \frac{1}{E}(v\sigma_{xx} + v\sigma_{yy})\end{aligned} \tag{3.48}$$

Shear stress–strain:

$$\begin{aligned}\varepsilon_{xy} = \varepsilon_{yx} &= \frac{1}{2\mu}\sigma_{xy} \\ \varepsilon_{yz} = \varepsilon_{zy} &= \frac{1}{2\mu}\sigma_{yz} \\ \varepsilon_{zx} = \varepsilon_{xz} &= \frac{1}{2\mu}\sigma_{zx}\end{aligned} \tag{3.49}$$

Adding the normal strains together, we have

$$\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz} = \frac{1-2v}{E}(\sigma_{xx} + \sigma_{yy} + \sigma_{zz}) \tag{3.50}$$

Writing the stress in terms of strain, we have

$$\sigma_{xx} = \frac{E}{1+v}\varepsilon_{xx} + \frac{vE}{1+v(1-2v)}(\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}) \tag{3.51}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

The coefficients (comprised of material properties only) are called the Lamé constants:

$$2\mu = \frac{E}{(1+v)} \tag{3.52}$$

$$\lambda = \frac{vE}{(1+v)(1-2v)} \tag{3.53}$$

where μ is the shear constant* and is also considered a Lamé constant.

One can now rewrite the stress in terms of the Lamé constant, i.e.:

$$\begin{aligned}
\sigma_{xx} &= (2\mu+\lambda)\varepsilon_{xx} + \lambda\varepsilon_{yy} + \lambda\varepsilon_{zz} \\
\sigma_{yy} &= \lambda\varepsilon_{xx} + (2\mu+\lambda)\varepsilon_{yy} + \lambda\varepsilon_{zz} \\
\sigma_{zz} &= \lambda\varepsilon_{xx} + \lambda\varepsilon_{yy} + (2\mu+\lambda)\varepsilon_{zz} \\
\sigma_{xy} &= 2\mu\varepsilon_{xy} \\
\sigma_{yz} &= 2\mu\varepsilon_{yz} \\
\sigma_{zx} &= 2\mu\varepsilon_{zx}
\end{aligned} \tag{3.54}$$

These stress–strain relationships are succinctly written in tensor form as

$$\sigma_{ij} = 2\mu\varepsilon_{ij} + \lambda\varepsilon_{kk}\delta_{ij}\,^{\dagger} \tag{3.55}$$

or

$$\sigma_{ij} = 2\mu\varepsilon_{ij} + \lambda\Delta\delta_{ij} \tag{3.56}$$

where $\Delta = \varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}$ and the indices i, j, k equal x, y, or z.

We are now prepared to develop the wave equation for a bulk material. Assume an ultrasonic (stress) wave propagates in a bulk material. Figure 3.10 depicts the x-directed stresses acting on a volume element within the bulk material. As before and as always, Newton's second law relates the acceleration of the particle (elemental volume) to the unbalanced forces acting on the particle. The sum of the forces in the x-direction acts to accelerate the mass (particle) in the x-direction:

$$\begin{aligned}
&\left(\sigma_{xx} + \frac{\partial\sigma_{xx}}{\partial x}\mathrm{d}x\right)\mathrm{d}y\,\mathrm{d}z - \sigma_{xx}\,\mathrm{d}y\,\mathrm{d}z + \\
&\left(\sigma_{xy} + \frac{\partial\sigma_{xy}}{\partial y}\mathrm{d}y\right)\mathrm{d}x\,\mathrm{d}z - \sigma_{xy}\,\mathrm{d}x\,\mathrm{d}z \\
&\quad + \left(\sigma_{xz} + \frac{\partial\sigma_{xz}}{\partial z}\right)\mathrm{d}x\,\mathrm{d}y - \sigma_{xz}\,\mathrm{d}x\,\mathrm{d}y = (\rho\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z)\frac{\partial^2 u_x}{\partial t^2}
\end{aligned}$$

or

$$\left(\frac{\partial\sigma_{xx}}{\partial x} + \frac{\partial\sigma_{xy}}{\partial y} + \frac{\partial\sigma_{xz}}{\partial z}\right)\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = (\rho\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z)\frac{\partial^2 u_x}{\partial t^2} \tag{3.57}$$

*$\mu = E/2(1+v)$ relationship is detailed in most mechanics of materials text books [20].

†δ_{ij} is called the Kroneker delta function; it is simply a shorthand notation for $\delta_{ij} = 1$ (for $i = j$) and $\delta_{ij} = 0$ (for $i \neq j$).

Copyright © Marcel Dekker, Inc. All rights reserved.

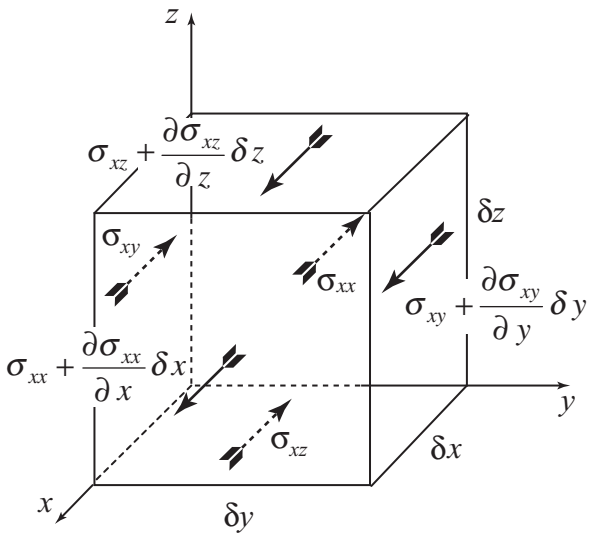


FIGURE 3.10 The *x*-directed stresses acting on a volume element in homogeneous bulk (unbounded) medium.

Therefore, the *x*, *y*, and *z* vector components of the three-dimensional wave equation (in *any* medium) are

$$\begin{aligned}
\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} + \frac{\partial \sigma_{xz}}{\partial z} &= \rho \frac{\partial^2 u_x}{\partial t^2} \\
\frac{\partial \sigma_{yx}}{\partial x} + \frac{\partial \sigma_{yy}}{\partial y} + \frac{\partial \sigma_{yz}}{\partial z} &= \rho \frac{\partial^2 u_y}{\partial t^2} \\
\frac{\partial \sigma_{zx}}{\partial x} + \frac{\partial \sigma_{zy}}{\partial y} + \frac{\partial \sigma_{zz}}{\partial z} &= \rho \frac{\partial^2 u_z}{\partial t^2}
\end{aligned} \tag{3.58}$$

For an isotropic, homogeneous material, *direct* substitution of $\sigma_{ij} = 2\mu\varepsilon_{ij} + \lambda\Delta$ and the definition of ε_{ij} (Eq. 3.43), yields

$$\rho \frac{\partial^2 u_x}{\partial t^2} = (\mu + \lambda) \frac{\partial \Delta}{\partial x} + \mu \nabla^2 u_x \tag{3.59}$$

for the *x* direction, where:

$$\nabla^2 = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \tag{3.60}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

similarly, for the y and z directions:

$$\rho \frac{\partial^2 u_y}{\partial t^2} = (\mu + \lambda) \frac{\partial \Delta}{\partial y} + \mu \nabla^2 u_y \tag{3.61}$$

$$\rho \frac{\partial^2 u_z}{\partial t^2} = (\mu + \lambda) \frac{\partial \Delta}{\partial z} + \mu \nabla^2 u_z \tag{3.62}$$

Now, assume the energy of the wave is propagating with phase velocity v in the x direction, i.e., $f(\omega t - kx)$. Equation 3.59 yields a phase velocity of

$$v_l = \left(\frac{2\mu + \lambda}{\rho} \right)^{1/2} \text{(pressure or longitudinal wave)} \tag{3.63}$$

and both Eqs. (3.61) and (3.62) yield the same phase velocity of

$$v_t = \left(\frac{\mu}{\rho} \right)^{1/2} \text{(shear or transverse wave)} \tag{3.64}$$

Previously, these velocities were referred to as dilatational (longitudinal) and torsional (transverse).

What physically happens to the material with the propagation of a *dilatational* (longitudinal) wave—a wave that dilates the volume? We can understand this process by thinking about it in two parts. First, think of a longitudinal wave—a propagating pressure disturbance (compression and rarefaction) whose force is in the same direction as the propagation of energy (phase). Second, add the Poisson effects, v; when a volume is elongated or compressed along one axis by an applied normal force, normal forces appear on the four lateral faces to satisfy $\varepsilon_{xx} = v\varepsilon_{yy} = v\varepsilon_{zz}$. Note that the dilatational wave produces normal stresses only (no shear stresses). Therefore, it travels as a volume change through the material with no rotation of the particles.

In contrast, the *torsional* (transverse) wave propagates by transmitting its momentum from one particle to the next through simple shear. The motion of simple shear contains pure shear deformation plus local, rigid rotation. Hence the term torsional wave. Because the volume of the particle does not change, we say it is *equivolumetric*. For an isotropic, homogeneous medium, we have found that dilatational (longitudinal) and torsional (transverse) waves propagate independently; they do not interact. Therefore, the mass-acceleration side of the complete wave equation can be separated into an *irrotational* and an *equivolumetric* component. This is common practice. See Section 3.6.3 for details.

3.2.3 Specific Acoustic Impedance and Pressure

In any physical movement, electrical, mechanical, or chemical, a *potential* (difference) must act as a driving force, and a fundamental *resistance* or

Copyright © Marcel Dekker, Inc. All rights reserved.

impedance must restrict movement to a finite velocity. The relationship among the potential (driving force), particle velocity, and impedance is:

$$\text{Impedance} = \frac{\text{Driving potential}}{\text{Particle velocity}}$$

For example, the impedance Z of an electrical circuit is

$$Z = \frac{V(\text{Voltage potential})}{I(\text{Average velocity of an electron})}$$

In the case of an acoustic wave,

$$Z = \frac{p}{v_{\text{particle}}} \tag{3.65}$$

where the driving potential p is the instantaneous acoustic pressure on the particles within the medium, v_{particle} is the instantaneous velocity of the oscillating particles, and Z is the impedance to movement of the particles within the medium.

To understand the concept of *acoustic pressure*, imagine two parallel planes, I and II, separated by a distance of dx within a medium at rest or at mechanical equilibrium (solid lines in Figure 3.11). A planar, longitudinal, pressure-driven wave propagates perpendicular to the planes I and II. At time t_0, plane I is displaced by a distance u. Because the pressure that drives the wave varies in the x direction, plane II is displaced by a distance of $u + (\partial u/\partial x)dx$. This difference in displacement of planes I and II gives rise to an associated strain, ε:

$$\varepsilon = \frac{\Delta l}{l_0} = \frac{\left[\left(u + \dfrac{\partial u}{\partial x}dx\right) - u\right]}{dx} \tag{3.66}$$

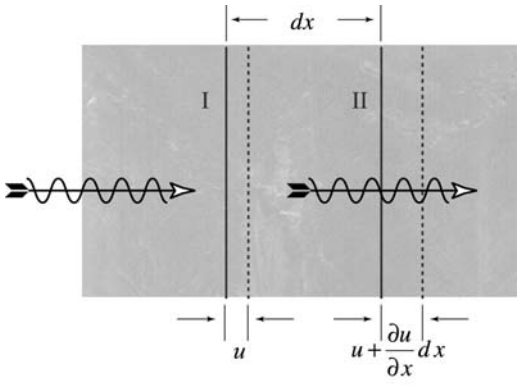


FIGURE 3.11 A longitudinal pressure wave acting on two parallel planes in a homogeneous, isotropic medium.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Therefore,

$$\varepsilon = \frac{\partial u}{\partial x} \tag{3.67}$$

From Hooke's law for linear elastic media (one-dimension), this strain is related to the stress through Young's modulus as:

$$\sigma = E\varepsilon = E\frac{\partial u}{\partial x} \tag{3.68}$$

Given that this stress (force/unit area) is equivalent to the acoustic pressure driving the traveling elastic wave,

$$p = E\frac{\partial u}{\partial x} \tag{3.69}$$

Recalling from Eq. (3.21), the longitudinal *phase velocity* for homogeneous, isotropic medium is $v = \sqrt{E/\rho}$ or $E = \rho v^2$. Substitution yields:

$$p = \rho v^2 \frac{\partial u}{\partial x} \quad \text{(Acoustic pressure)} \tag{3.70}$$

Combining the above equation with Eq. (3.65) yields:

$$Z = \frac{p}{v_{\text{particle}}} = \frac{\rho v^2 \dfrac{\partial u}{\partial x}}{\dfrac{\partial u}{\partial t}} = \rho v^2 \frac{\partial t}{\partial x} = \frac{\rho v^2}{v} = \rho v \tag{3.71}$$

for a longitudinal wave. Therefore, both the material density and elastic properties affect the acoustic impedance. Materials with high impedances are often called *ultrasonically hard*, and materials with low impedances are called *ultrasonically soft*.

For example, consider the difference in the specific acoustic impedance of crown glass ($\rho_{\text{glass}} = 2.5\,\text{g/cm}^3$) and polyetrafluoroethylene (teflon) ($\rho_{\text{teflon}} = 2.2\,\text{g/cm}^3$) whose densities are comparable. From Table 3.2, the longitudinal velocities are $v_{\text{glass}} = 5.6\,\text{km/s}$ and $v_{\text{teflon}} = 1.35\,\text{km/s}$. Therefore,

$$Z_{\text{glass}} = \left(2.5\frac{\text{g}}{\text{cm}^3}\right)\left(5.66\frac{\text{km}}{\text{s}}\right) = 14.15 \times 10^6 \frac{\text{kg}}{\text{m}^2\text{s}}$$

$$Z_{\text{teflon}} = \left(2.2\frac{\text{g}}{\text{cm}^3}\right)\left(1.35\frac{\text{km}}{\text{s}}\right) = 2.97 \times 10^6 \frac{\text{kg}}{\text{m}^2\text{s}}$$

A more generalized form of the specific acoustic impedance must include absorption and anisotropic material effects. Thus, the acoustic impedance would be a complex variable and represented as:

$$Z = R + iX \tag{3.72}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

where R is the real resistive component, X represents the reactive component (related to absorption) and $i = \sqrt{-1}$ [21].

At a boundary, the difference in acoustic impedances between the two materials governs how much of an incident wave will reflect from or transmit through the boundary.

3.2.4 Reflection and Refraction at an Interface

At almost every waking moment of our lives, we encounter the interaction of a wave with a boundary. Seeing your face in a mirror, partial reflection of images off windows or a pool of water, the pencil that seems to bend when partially submerged in water—all are examples of light waves interacting with boundaries. Sound waves interact with boundaries just as often, as we know from hearing an echo return from a distant wall, eavesdropping on muffled voices in the next room, or making music (focused sound) by blowing through a horn.

The behavior or change of a propagating ultrasonic wave as it encounters a boundary is fundamental to ultrasonic material evaluation. Voids, cracks, inclusions, and coatings can be distinguished, provided that they possess different elastic properties from the parent material. We can think of a parent material containing a void, crack, etc. as two materials having different specific acoustic impedances (Z).

Deviation of, or changes in, the wave as it encounters a boundary (interface) can take one or more of the following forms:

A reflection and/or transmission
A propagation of the wave along the interface
A change in direction of travel (*refraction*)
A conversion from one type of wave to another type of wave (*mode conversion*)

The familiar analogy of light waves explains these possible responses and their interdependence.* Most people understand the notion of partially, or totally, reflected or transmitted waves, for example, a light wave striking a mirror or a partially reflecting surface. Even the notion of the bending (refraction) of light waves from an object partially submerged in water is quite familiar. But the concept of mode conversion is less familiar. In mode conversion, an incident wave of one type (mode) is converted into one or more modes of waves. Figure 3.12 shows an example of a light wave responding to a boundary between air and a calcite crystal. The light wave initially propagating in air impinges on the air/calcite boundary and is refracted into two distinct waves, resulting in a double

*Since any wave will respond to a boundary in one or more of the prescribed manners, it doesn't matter which type of energy (electrical or mechanical) is propagating.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

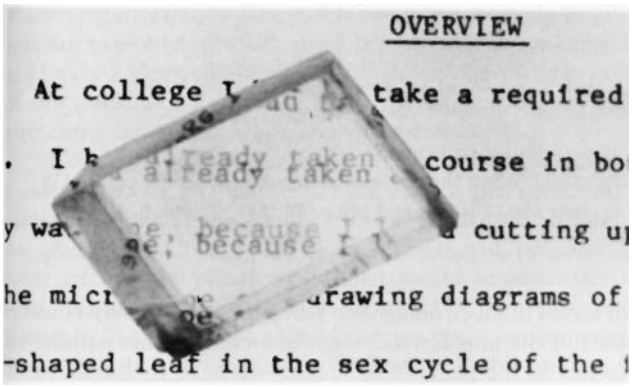


FIGURE 3.12 A light wave is split (refracted) into two modes (birefringence) upon incidence on an air/calcite interface. (H Benson. University Physics. New York: John Wiley, 1991, p 779. Reprinted by permission John Wiley & Sons.)

image. Because the resulting waves are refracted at different angles, they must propagate at different velocities and therefore be different modes.* Refraction of a wave into two distinct waves is called *birefringence*. In acoustics, birefringence and even trirefringence (three resultant modes) are common.

Predicting specifically how a wave will change when it interacts with a boundary is a typical boundary value problem: the boundary conditions (those physical parameters that must be continuous across the boundary) are identified and applied to the equations governing wave propagation. The boundary conditions for ultrasonic wave propagation are

Continuity of particle velocity
Continuity of acoustic pressure
Continuity of phase of the wave (Snell's law)†

The particle velocity on both sides of the boundary must be equal; otherwise, the two materials would not be in contact.‡ Continuity of the acoustic pressure is a direct result of the continuity of tractions (stresses). Note that

*This claim is a requirement to satisfy Fermat's Principle of Least Time, and will be supported later through Snell's law.

†Snell's law is not really a true boundary condition; rather, it follows from the other two boundary conditions. We state it as a separate boundary condition to emphasize the continuity of the phase.

‡In the special case of a solid and a nonviscous liquid (e.g. water), the liquid moves freely parallel to the boundary; therefore, only the perpendicular component of the particle velocity is continuous.

Copyright © Marcel Dekker, Inc. All rights reserved.

boundary conditions 1 and 2 are interdependent—one cannot exist without the other. Snell's law requires that at any point on the boundary, the phase of all resultant waves leaving the boundary must be equal to the phase of the incident wave. Stated another way, the tangential component of the wave vector $\vec{k}$ must be continuous across the boundary.

How do we predict wave behavior at a boundary? To understand this phenomenon more completely, let's first examine the simplest case, normal incidence, and then examine the more complex case, oblique incidence in homogenous isotropic materials.

Normal Incidence

In normal incidence, a propagating plane wave impinges at 90° onto a smooth surface separating materials 1 and 2 (Figure 3.13). Unlike in cases of oblique incidence, the resultant transmitted or reflected waves do not refract (bend) away from the direction of the incident wave. This nonrefraction implies that there will be only one possible transmitted and one reflected wave.

How much of the incident wave is reflected, and how much is transmitted? This problem is depicted in Figure 3.13, where subscripts 1 or 2 indicate properties of material 1 or 2, respectively. The first boundary condition is continuity of particle velocity; the sum of the particle velocities at the boundary in medium 1 equal those in medium 2:

$$\vec{v}_{\text{particle}_i} + \vec{v}_{\text{particle}_r} = \vec{v}_{\text{particle}_t} \tag{3.73}$$

where $\vec{v}_{\text{particle}} = \vec{V} e^{j(\omega t - \vec{k}\cdot\vec{r})} = \partial\vec{u}/\partial t$, $\vec{V}$ is the amplitude coefficient, and subscripts i, r, and t refer to the incident, reflected, and transmitted waves. We note the

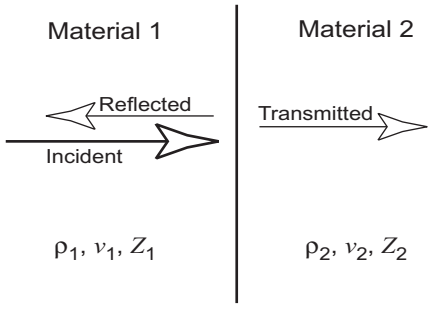


FIGURE 3.13 Reflection and transmission of a normal incident wave at a boundary.

Copyright © Marcel Dekker, Inc. All rights reserved.

particle velocity of each wave is confined to *phase* propagation in the x direction ($\vec{k} \cdot \vec{r} = k_x x$). Thus

$$V_i e^{j(\omega_1 t - k_1 x)} + V_r e^{j(\omega_1 t - k_1 x)} = V_t e^{j(\omega_2 t - k_2 x)} \tag{3.74}$$

Knowing that the acoustic pressure varies as $p = Pe^{j(\omega t - \vec{k}\cdot\vec{r})}$, the second boundary condition, continuity of acoustic pressure (stress), can be written as

$$P_i e^{j(\omega_1 t - k_1 x)} + P_r e^{j(\omega_1 t - k_1 x)} = P_t e^{j(\omega_2 t - k_2 x)} \tag{3.75}$$

Because the boundary is placed at $x = 0$ and the wave is propagating in the x-direction, $\vec{k} \cdot \vec{r} = k_x x = 0$. Assume continuity of frequency across the boundary $\omega_1 = \omega_2$. At the boundary, Eqs. (3.74) and (3.75) can be written as

$$V_i + V_r = V_t \tag{3.76}$$

and

$$P_i + P_r = P_t \tag{3.77}$$

Now to answer "How much of the propagating pressure wave is reflected and how much is transmitted?" we define the *reflection coefficient* as

$$r = \frac{P_r}{P_i} \tag{3.78}$$

and the *transmission coefficient* as

$$t = \frac{P_t}{P_i} \tag{3.79}$$

These equations will indeed yield the pressure amplitudes of the reflected or transmitted wave; but they are awkward, since the variables are not readily measurable. Recall the initial suggestion (Section 3.2.3) that the reflection and transmission of an incident wave is governed by the difference of the acoustic impedances, Z_1 and Z_2. $Z = \rho v$ and is readily calculated from tabulated values of density and phase velocity. Using Eq. (3.65), the acoustic impedance in materials 1 and 2 can be represented as

$$Z_1 = \frac{p_i}{v_{\text{particle}_i}} = \frac{P_i}{V_i}$$

$$Z_1 = \frac{p_r}{v_{\text{particle}_r}} = \frac{P_r}{V_r}$$

and

$$Z_2 = \frac{p_t}{v_{\text{particle}_t}} \quad \text{or} \quad Z_2 = \frac{p_i + p_r}{v_{\text{particle}_i} + v_{\text{particle}_r}} = \frac{P_i + P_r}{V_i + V_r} \tag{3.80}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

After some algebraic manipulation, the reflection and transmission coefficients can be written in terms of acoustic impedances:

$$r = \frac{Z_2 - Z_1}{Z_2 + Z_1} \tag{3.81}$$

and

$$t = \frac{2Z_2}{Z_2 + Z_1} \tag{3.82}$$

These amplitude ratios are called the *acoustic Fresnel equations* for isotropic, homogeneous medium with a wave at normal incidence.

Let us examine in more detail at what the reflected and transmitted acoustic pressures might look like for different boundary materials.

Example: Two common and exemplary cases of reflection and transmission are depicted in Figure 3.14. In both cases the incident wave strikes the boundary of steel and water at an incident angle of 90° from the interface. In Figure 3.14a the initial wave propagates in the water and impinges on the steel (water/steel interface), whereas in Figure 3.14b, the incident wave is initially propagating in the steel and impinges on the water (steel/water interface).

1. For simplicity, assume $P_i|_{\text{steel}} = P_i|_{\text{water}} = 1$. From Table 3.2 the acoustic impedances of water and steel are $Z_{H_2O} = 1.5 \times 10^6\,\text{N s/m}^3$ and $Z_{\text{steel}} = 45 \times 10^6\,\text{N s/m}^3$.

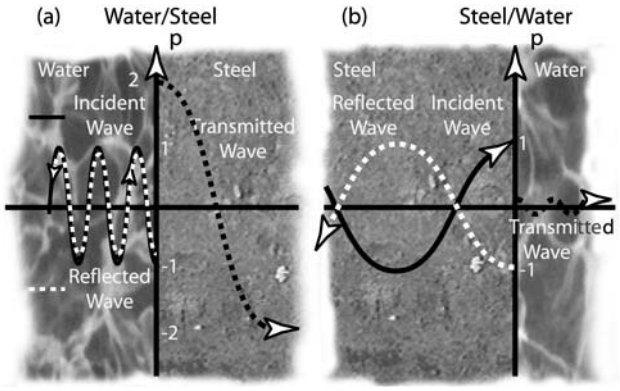


FIGURE 3.14 Pressure amplitudes of the reflected and transmitted waves at (a) a steel/water interface and (b) water/steel interface. Incident wave is a longitudinal mode and normalized to unity. (Adapted from Ref. 51. Reprinted by permission of Springer-Verlag.)

Copyright © Marcel Dekker, Inc. All rights reserved.

2. From Equations (3.81) and (3.82), we calculate the reflection and transmission coefficients:

   "water/steel" interface

$$r_{H_2O/steel} = \frac{Z_2 - Z_1}{Z_2 + Z_1} = \left(\frac{45 \times 10^6 - 1.5 \times 10^6}{45 \times 10^6 + 1.5 \times 10^6}\right) = 0.935$$

$$t_{H_2O/steel} = \frac{2Z_2}{Z_2 + Z_1} = \frac{2(45 \times 10^6)}{45 \times 10^6 + 1.5 \times 10^6} = 1.935$$

   "steel/water" interface

$$r_{steel/H_2O} = \frac{Z_2 - Z_1}{Z_2 + Z_1} = \left(\frac{1.5 \times 10^6 - 45 \times 10^6}{1.5 \times 10^6 + 45 \times 10^6}\right) = -0.935$$

$$t_{steel/H_2O} = \frac{2Z_2}{Z_2 + Z_1} = \frac{2(1.5 \times 10^6)}{45 \times 10^6 + 1.5 \times 10^6} = 0.0645$$

3. Using $P_i|_{steel} = P_i|_{water} = 1$, the reflected and transmitted pressures *at the boundary* are

   "water/steel" interface

$$P_r|_{H_2O} = r_{H_2O/steel} P_i|_{H_2O} = (0.935)(1) = 0.935$$

$$P_t|_{steel} = t_{H_2O/steel} P_i|_{H_2O} = (1.935)(1) = 1.935$$

   "steel/water" interface

$$P_r|_{steel} = r_{steel/H_2O} P_i|_{steel} = (-0.935)(1) = -0.935$$

$$P_t|_{H_2O} = t_{steel/H_2O} P_i|_{steel} = (0.0645)(1) = 0.0645$$

Notice that $P_r|_{steel}$ is negative, and $P_t|_{steel}$ is greater than the pressure of the original incident wave. These results are real, and they do not violate any conservation of energy laws. To understand these responses, let's review the details of the two graphs in Figure 3.14 very carefully, keeping in mind the required boundary condition of $p_i + p_r = p_t$ and using a little common sense.

Example: Assume a wave strikes a boundary; what do we expect to happen to the pressures of the involved waves? To develop a "gut feeling" level of understanding, it is useful to look at the extreme cases.

Case I (Figure 3.14a) Assume a longitudinal plane wave propagates in an ultrasonically soft (small Z) medium and encounters an extremely hard medium (large Z), e.g., water/steel. If the materials at the boundary are displaced equally (i.e., remain in contact), consider the amount of pressure required to displace H_2O a distance Δx compared to that needed to move the steel the same distance. The *excess pressure* required to displace the steel manifests itself as a reflected wave.

Copyright © Marcel Dekker, Inc. All rights reserved.

If this still seems a little unsettling remember that the transmitted, reflected and incident waves all exist simultaneously at the boundary.*

Case II (Figure 3.14b) Assume a longitudinal plane wave propagates in an ultrasonically hard medium and encounters an extremely soft medium, e.g., steel/water. As in Case I, the two materials must stay in contact. Because the pressure to launch a transmitted wave into water is so small compared with the pressure required in steel to produce the equivalent displacements, the reflected wave pressure must be 180° out of phase so as to cancel the *excess pressure* of the incident wave.

Still find it a little unsettling that a transmitted pressure can be greater than the incident pressure? Let's look at the problem from the viewpoint of energy conservation. The energy intensity (energy per unit time per unit area) of a wave is given by

$$I = \frac{1}{2} v \rho v_{\text{particle}}^2 = \frac{1}{2} Z v_{\text{particle}}^2 = \frac{1}{2} \frac{p^2}{Z} \tag{3.83}$$

i.e., $I \propto p^2/Z$. Therefore, for a given acoustic energy, the pressure can be quite large in an acoustically hard medium (higher Z).

Conservation of energy requires that

$$I_i = I_r + I_t \tag{3.84}$$

Note that the sum of the reflected and transmitted wave intensities is equal to the intensity of the incident wave. Contrast this energy intensity equation with the continuity-of-acoustic-pressures equation, $p_i + p_r = p_t$.

In experiment, intensity ratios are more important (and easier to understand) than pressure ratios. The intensity ratios are the *reflectance*

$$R = \frac{I_r}{I_i} \tag{3.85}$$

and the *transmittance*

$$T = \frac{I_t}{I_i} \tag{3.86}$$

Consequently, due to conservation of energy, $R + T = 1$.

*In fact, separate incident, reflected, and transmitted waves exist only after the wave leaves the boundary. At best, one could argue that the events at the boundary are a superposition of the *concepts* of independent waves, and thus a mathematical convenience that allows us to consider separate incident, reflected, and transmitted waves.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Because $I \propto p^2$, then $R \propto r^2$ and $T \propto t^2$. Recognizing that the incident and reflected wave propagate in the same material, i.e., $Z_i = Z_r = Z_1$, then

$$R = r^2 \frac{Z_r}{Z_i} = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1}\right)^2 \tag{3.87}$$

$$T = t^2 \frac{Z_t}{Z_i} = \frac{4Z_1 Z_2}{(Z_1 + Z_2)} \tag{3.88}$$

Recall that these equations for r, t, R, and T have been derived in this section *assuming* a plane wave impinging at an angle of 90° to the interface.

Oblique Incidence

Oblique incidence complicates the problem presented in the case of normal incidence. When a wave impinges on a boundary at an oblique (not normal nor parallel) angle, we must account for

A reflection and/or transmission
A change in direction of travel of the resultant waves (refraction)
Conversions from one type of wave to another type of wave (mode conversion)

As in the previous section, the reflection and transmission amplitude coefficients for the *scattered* (transmitted or reflected) waves are determined by the acoustic Fresnel equations. For oblique incidence, however, the wave can be multiply refringent, i.e., it can result in multiple scattered waves, each with its own Fresnel equation. A convenient trigonometric relation known as *Snell's law* gives the propagation directions of these scattered waves. For clarity, this section focuses on pure shear or pure longitudinal waves in isotropic homogenous materials.

Before discussing Snell's law, let's distinguish between the two special cases of shear or transverse waves that are created by the existence of the boundary. Although the shear wave particle displacement is perpendicular to the propagation direction, the displacement can be either *parallel* or at an *oblique angle* to the plane of the interface (Figure 3.15). If the displacement is parallel (*shear horizontal*, SH) to the interface, the reflected and transmitted waves are also SH waves; no mode conversion takes place. If the displacement is oblique (*shear vertical*, SV) to the interface, then a component of the displacement, perpendicular to the interface, can excite a longitudinal mode. Snell's law does not distinguish between SV and SH waves (they propagate at the same velocity). Consequently, Snell's law can predict a direction for a reflected or transmitted longitudinal wave for an incident SH wave. The Fresnel equations, however, would yield a zero amplitude for these longitudinal waves. In summary, SH waves couple only to scattered SH modes, whereas SV modes couple to scattered SV and longitudinal modes, but not to SH waves.

Copyright © Marcel Dekker, Inc. All rights reserved.

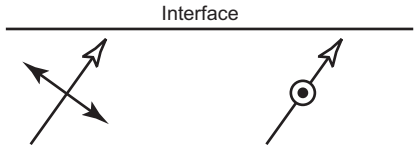


FIGURE 3.15 Particle displacement *polarization* of shear waves relative to a boundary: (a) shear vertical, SV, (b) shear horizontal, SH, ($\odot$ indicates an arrow pointing out of the page, i.e., perpendicular to the plane of the page.)

Snell's Law. Snell's Law (or the continuity-of-phase law) is used to determine the directions of propagation of transmitted and reflected waves. In part, the utility of Snell's law is its simplicity of application. Snell's law (derived in Section 3.6.4) is given as

$$\frac{\sin\theta_i}{v_i} = \frac{\sin\theta_{rs}}{v_{rs}} = \frac{\sin\theta_{rl}}{v_{rl}} = \frac{\sin\theta_{ts}}{v_{ts}} = \frac{\sin\theta_{tl}}{v_{tl}}$$

or

$$\frac{k_i}{\omega}\sin\theta_i = \frac{k_{rs}}{\omega}\sin\theta_{rs} = \frac{k_{rl}}{\omega}\sin\theta_{rl} = \frac{k_{ts}}{\omega}\sin\theta_{ts} = \frac{k_{tl}}{\omega}\sin\theta_{tl} \tag{3.89}$$

where $v = \omega/k$ and s and l indicate shear and longitudinal modes; v_{rs} represents the velocity of a reflected shear wave.* Note that the incident wave, subscript i, could be either in the shear or the longitudinal mode. Eq. (3.89) includes the possibility of multiple reflected and transmitted waves due to mode conversion. *Also note that in ultrasonics the angle of incidence θ is measured from the surface normal. Therefore $\theta = 0^\circ$ for a normal incident wave.*

Example: Consider a longitudinal wave in a plastic, incident at an angle of 30° to a semi-infinite section of Inconel. The velocity in the plastic is 2.7 km/ms (longitudinal) and 1.1 km/ms(shear). The velocity in Inconel is 5.7 km/ms (longitudinal) and 3.0 km/ms(shear). Four types of waves can emerge from this interaction: reflected shear, reflected longitudinal, refracted shear, and refracted longitudinal. According to Snell's law:

$$\theta_{rl} = \sin^{-1}\left(\frac{1.1}{2.7}\sin 30^\circ\right) = 12^\circ$$

$$\theta_{rs} = \sin^{-1}\left(\frac{2.7}{2.7}\sin 30^\circ\right) = 30^\circ$$

$$\theta_{ts} = \sin^{-1}\left(\frac{3.0}{2.7}\sin 30^\circ\right) = 34^\circ$$

$$\theta_{tl} = \sin^{-1}\left(\frac{5.7}{2.7}\sin 30^\circ\right) = \sin^{-1} 1.06$$

*In the literature both s and t are used as subscripts to indicate a transverse wave, e.g., θ_{tt} or θ_{ts} implies a transmitted transverse wave angle.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Note that the reflected longitudinal wave has the same angle as the incident wave. Yes, "angle of incidence equals angle of reflection"—but only for the same wave mode! Note also that the angle of refraction θ_{tl} is too great for the longitudinal wave to enter the Inconel. Therefore, in this case, an incident longitudinal wave in plastic transmits into Inconel a pure mode converted shear only.

Caution! Snell's law can predict angles for wave modes that do not exist. The Fresnel equations, however, will predict a zero amplitude for these modes.

Critical Angles. In the example above, Snell's law yielded a transmitted wave angle, θ_{tl}, greater than 90°. This result simply states that the particular mode (transmitted longitudinally for this case) does not exist for an incident longitudinal angle of 30°, and the energy goes into other modes.

To examine this phenomenon more carefully, we will use the example above and discuss the response of the transmitted waves as the incident longitudinal wave angle ranges from to 0–90°. From our discussion on normal incident waves, at $\theta_{il} = 0°$ (normal incidence) there is no refraction and only one transmitted wave—a longitudinal mode.

Now Snell's law for the transmitted longitudinal wave (from the plastic wedge into the Inconel) is

$$\frac{\sin\theta_{tl}}{\sin\theta_{il}} = \frac{5.7\dfrac{\text{km}}{\text{mm}}}{2.7\dfrac{\text{km}}{\text{mm}}} = 2.1$$

This states that the transmitted longitudinal wave is bent at an angle greater than the incident angle ($\theta_{tl} > \theta_{il}$)—the transmitted angle *leads* the incident angle. Therefore, the transmitted wave will be bent to 90° before the incident wave reaches 90°. When the transmitted wave reaches 90°, it propagates along the interface. Thus, this *first critical angle*, beyond which the transmitted longitudinal mode no longer propagates, is

$$\theta_{\text{critical}} = \theta_{il} = \sin^{-1}\left((\sin 90°)\frac{2.7}{5.7}\right)$$
$$= 28.3°(\text{first critical angle})$$

In the current example, the mode-converted, transmitted shear wave also leads the incident longitudinal. Thus, it too has a critical angle of $\theta_{\text{critical}} = 64.2°$. Beyond this *second* critical angle, no wave is transmitted into the second medium and the incident wave is said to be *totally internally reflected*—all the energy goes into the reflected modes.

Notice that if the wave speed in the transmitted medium is higher than that of the incident wave, then there will be a critical angle. Similarly, if the wave speed of a mode converted *reflected* wave is higher than that of the incident wave,

Copyright © Marcel Dekker, Inc. All rights reserved.

TABLE 3.3 Possible Critical Angles for an Incident Shear or Longitudinal Plane Wave Incident on a Planar Boundary

Incident Shear Wave		Incident Longitudinal Wave	
$v_s^1 > v_l^2 > v_s^2$	1 critical angle	$v_l^1 > v_l^2 > v_s^2$	No critical angles
$v_l^2 > v_s^1 > v_s^2$	2 critical angles	$v_l^2 > v_l^1 > v_s^2$	1 critical angle
$v_l^2 > v_s^2 > v_s^1$	3 critical angles	$v_l^2 > v_s^2 > v_l^1$	2 critical angles

Source: BA Auld. Acoustic Fields and Waves in Solids. Vol. II. New York: John Wiley and Sons, 1973, p 7.

it will also have a critical angle. Table 3.3 lists the number of critical angles for combinations of incident wave modes and relative wave speeds.

Experiment: We can observe a critical angle using a glass (or pan) of water and a wrist watch. (Make sure the watch is waterproof!) Fill the container with water. Holding the watch with the dial visible and horizontal, submerse the watch into the water. Viewing at normal incidence $\theta_i = 0^\circ$ (directly above), the face is clearly visible. Tilt the watch. At some angle the face will disappear and the sides of the container will appear. At this critical angle and beyond the incident light wave is totally internally reflected and the crystal of the watch acts like a mirror. If your watch has a slight crown, at the critical angle part of the watch crystal will be a perfect mirror and part will be transparent.

Fresnel Equations. Snell's law determines the direction, but not the amplitude, of scattered waves. As before, the Fresnel equations yield the amplitudes of these scattered waves. The Fresnel equations for oblique incidence are developed via the solution to the boundary value problem, as in the case of normal incidence. The derivations for oblique incidence are relatively complicated and are not presented here. Instead, in the following three cases we present graphical results of the equations for selected materials. Because the results for different types of incident waves and boundaries are significantly similar, only Case I is presented in full detail. The following is principally a summary of results presented in Ref. 22.

Case I: Solid/Free Boundary Interface An ultrasonic *free boundary* means that the free bounding medium is a vacuum, or at least a rarified medium that does not significantly support acoustic waves. Figures 3.16 and 3.17 show the results of the scattered waves, given an incident longitudinal and transverse wave (SV) onto the solid/air interface, respectively. The reflection coefficients* (pressure ratios) are plotted as a function of the angle. The acoustic pressure of the incident wave (left-hand quadrant) is assigned a value of unity.

*There are no transmitted waves, $t = 0$, into the second medium, since it does not support an acoustic wave (by definition of a free boundary).

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

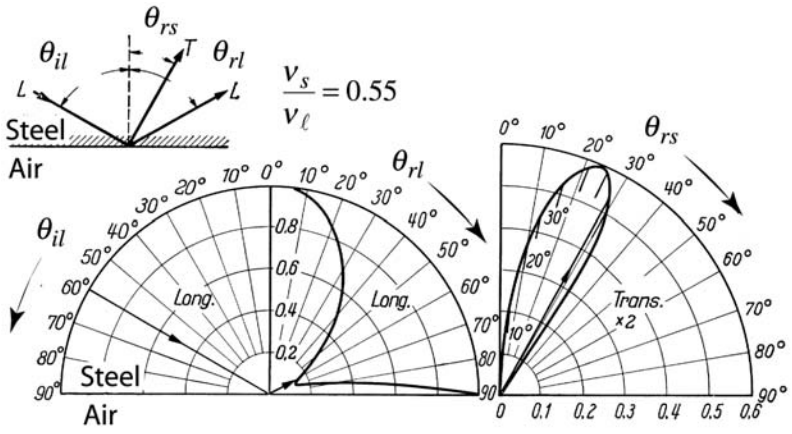


FIGURE 3.16 Plot of the Fresnel coefficients for an incident *longitudinal* wave impinging on a steel/air boundary. (Subscripts *r*, *t*, *s*, and *l* = reflected, transmitted, transverse (shear), and longitudinal respectively. (Adapted from J. Krautkramer, H. Krautkramer, Ultrasonic Testing of Materials. 4th ed Berlin: Springer-Verlag, 1990, p 25. Reprinted by permission of Springer-Verlag.)

The adjacent quadrant shows the pressure amplitude of the reflected wave of the same type or mode as the incident wave, and therefore reflected at the same angle. In the *detached* quadrant, the reflected *mode-converted* acoustic wave pressure is plotted. In each plot, an example is given of a wave at a given incident angle and the associated reflected amplitudes. Snell's law is used to determine the angular relationships between the incident and scattered waves.

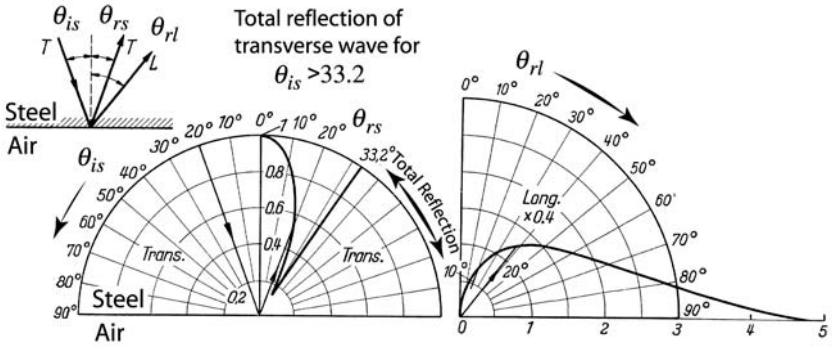


FIGURE 3.17 Plot of the Fresnel coefficients for an incident *transverse* wave impinging on a steel/air boundary. (Adapted from J. Krautkramer, H. Krautkramer. Ultrasonic Testing of Materials. 4th ed. Berlin: Springer-Verlag, 1990, p 25. Reprinted by permission of Springer-Verlag.)

Copyright © Marcel Dekker, Inc. All rights reserved.

For an incident transverse wave (Figure 3.17), a *weak* (low-amplitude) transverse wave is reflected between angles of $\theta_{rs} \cong 25^\circ$ and $\theta_{rs} \cong 33^\circ$ in favor of a *reflected* mode-converted longitudinal wave propagating between angles $\theta_{rl} \cong 50^\circ$ and $\theta_{rl} \cong 90^\circ$. The reflected longitudinal wave encounters a critical angle at $\theta_{is} = 33.2^\circ$ ($\theta_{rl} = 90^\circ$) beyond which the incident transverse wave is totally reflected into the transverse wave and the reflected longitudinal wave no longer propagates.

Case II: Liquid/Solid–Solid/Liquid Interface The case of a wave propagating in water incident onto a solid (liquid/solid) is of great importance in ultrasonic NDE. The specimen being ultrasonically tested (UT) is immersed in a liquid, usually water. The liquid provides a path to couple the UT wave from the transducer to the specimen (excitation) and from the specimen to the transducer (reception). Practical details of immersion UT are presented in Section 3.3. The resultant scattered waves at a liquid/solid interface are represented in Figure 3.18. Recall that the viscosity of water is too low to support transverse (shear) waves. Therefore, the incident and the reflected wave in the liquid are both in *longitudinal* modes with equal angles of incidence and reflection. These waves are represented in the upper two quadrants of Figure 3.18. The attached lower

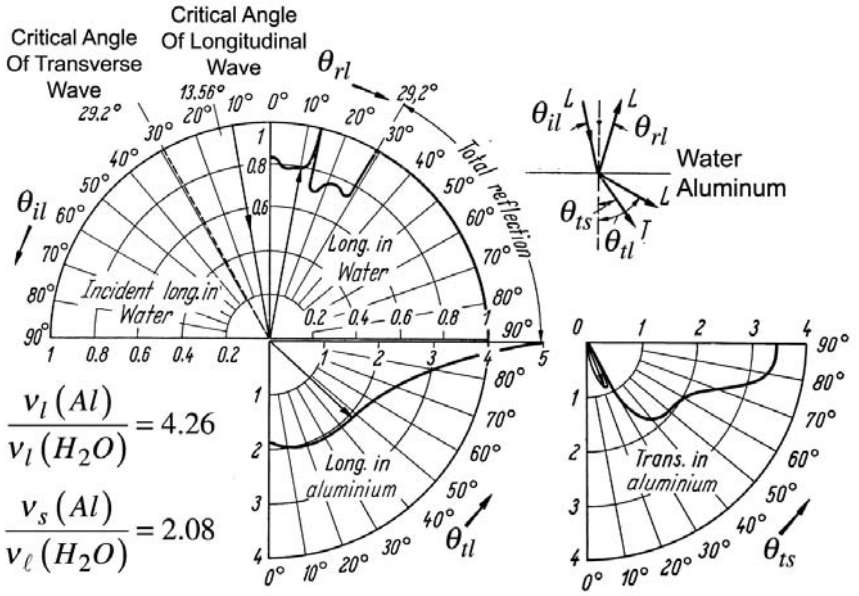


FIGURE 3.18 Plot of the Fresnel coefficients for an incident *longitudinal* wave impinging on a water/aluminum boundary. (Adapted fom J. Krautkramer, H. Krautkramer. Ultrasonic Testing of Materials. 4th ed. Berlin: Springer-Verlag, 1990, p 25. Reprinted by permission of Springer-Verlag).

Copyright © Marcel Dekker, Inc. All rights reserved.

quadrant depicts the results for the transmitted longitudinal wave, and the detached lower quadrant graphs the transmitted, mode-converted shear wave. Because the wave speeds for both the longitudinal and shear wave in aluminum are faster than the longitudinal wave speed in water, there will be two critical angles. Starting from normal incidence, the first critical angle occurs as the *transmitted* longitudinal wave reaches 90°. The corresponding incident longitudinal wave is at $\theta_{il} = 13.56°$. A second critical angle occurs at $\theta_{il} = 29.2°$, when the transmitted transverse wave propagation direction equals 90°. Beyond this second critical angle, total internal reflection occurs, indicated by a pressure amplitude of unity for the reflected longitudinal wave for $\theta_{rl} = 29.2$ to 90°.

In practice, the region between the two critical angles is significant for fluid immersion UT. Only a *single* wave is launched into the second medium when the incident angle is constrained between the two critical angles ($13.56° > \theta_{il} > 29.2°$ for the current example). A single wave interrogating the material reduces the difficulty of interpreting the results.

So far we have analyzed the "liquid to solid" boundary problem, i.e., the modes, associated refracted angles, and pressure amplitude that can be generated in a solid from an incident longitudinal wave propagating in water. If, on the other hand, the ultrasound originates in a solid, what are the reflection and transmission characteristics upon incidence at the solid/liquid interface? The response from an incident wave propagating in aluminum onto a boundary of water (solid/liquid) is graphed in Figure 3.19 for a longitudinal incident wave, and in Figure 3.20 for a transverse incident wave. Note that this condition of propagation from a medium

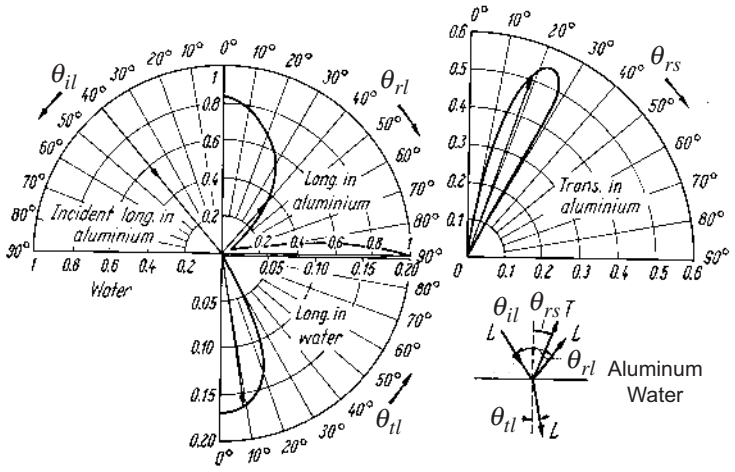


FIGURE 3.19 Plot of the Fresnel coefficients for an incident *longitudinal* wave impinging on a aluminum/water boundary. (Adapted from J. Krautkramer, H. Krautkramer. Ultrasonic Testing of Materials. 4th ed. Berlin: Springer-Verlag, 1990, p 26. Reprinted by permision of Springer-Verlag.)

Copyright © Marcel Dekker, Inc. All rights reserved.

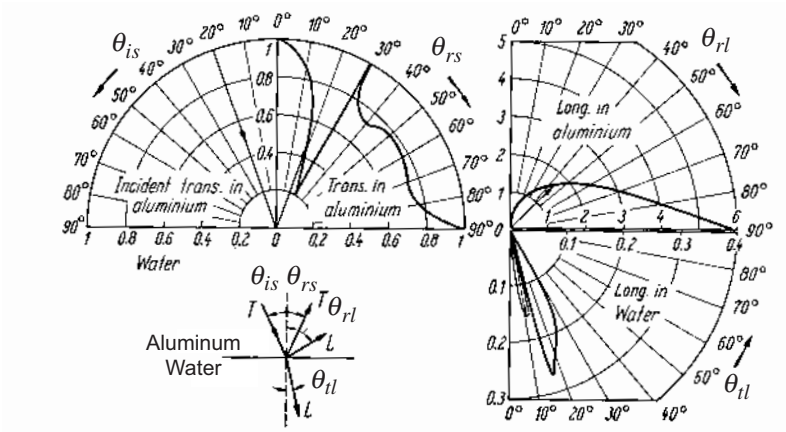


FIGURE 3.20 Plot of the Fresnel coefficients for an incident *transverse* wave impinging on a aluminum/water boundary. (Adapted from J. Krautkramer, H. Krautkramer. Ultrasonic Testing of Materials. 4th ed. Berlin: Springer-Verlag, 1990, p 27. Reprinted by permission of Springer-Verlag.)

with a higher acoustic velocity to a medium of lower acoustic velocity does not produce any critical angles.

Case III: Solid/Thin Liquid Layer/Solid Interfaces A thin layer of liquid sandwiched between two solids is probably the most commonly used configuration to couple ultrasound into and out of an interrogated material. Such a typical arrangement is the generation transducer/liquid couplant/specimen for injection of the ultrasound into the specimen. Figure 3.21 presents details of the reflection and transmission acoustic pressure for the configuration of perspex liquid-coupled to steel. Because shear waves do not propagate in most liquids, the incident wave in the perspex is a longitudinal mode. Only incident angles between the two critical angles are detailed.

Echo Transmission. In a typical application of ultrasonic NDE, the energy from a transducer is transmitted into the specimen, and then the energy is transmitted from the specimen into the transducer—an *echo*. In this scenario, the appropriate Fresnel equations would be applied, first to determine how much of the wave is transmitted into the specimen, then to determine how much of the wave is received at the transducer, as well as other any boundary in the path of the ultrasonic wave.

Lutsch and Kuhn (23) model the echo assuming a unity amplitude incident wave and a planar reflector within the material. Figure 3.22 plots the received energy (echo) for the water/aluminum boundary. A longitudinal wave is generated and received in the water. Figure 3.22a shows the response associated with the longitudinal wave propagating in the aluminum, and Figure 3.22b indicates

Copyright © Marcel Dekker, Inc. All rights reserved.

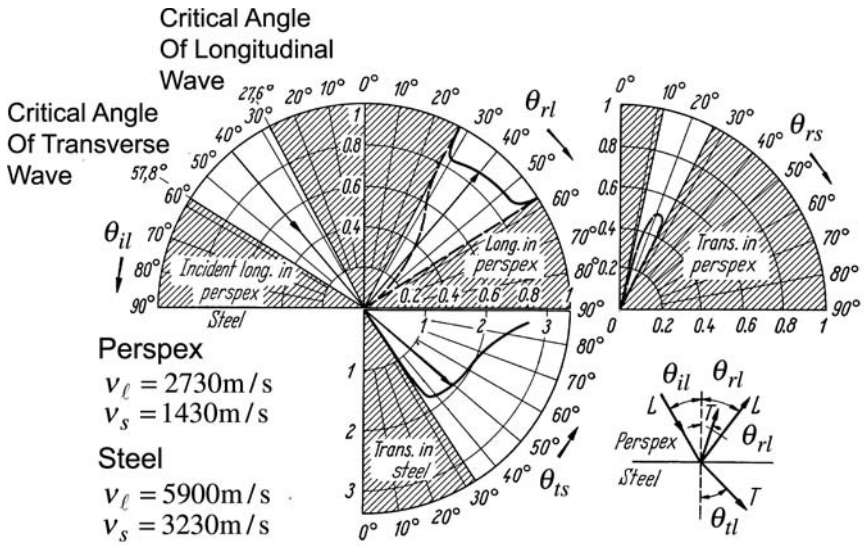


FIGURE 3.21 Plot of the Fresnel coefficients for an incident *longitudinal* wave impinging on a *perspex/liquid/steel boundary*, where the liquid layer thickness is small compared to the wavelength. (J. Krautkramer, H. Krautkramer. Ultrasonic Testing of Materials. 4th ed. Berlin: Springer-Verlag, 1990, p 29. Reprinted by permission from Springer-Verlag.)

the echo response for the mode converted transverse wave propagating in the aluminum. Before the longitudinal critical angle, $0^\circ < \theta_{il} < 13.56^\circ$, the energy is primarily contained in the longitudinal wave up to 30 % of the incident wave (Figure 3.22a). Between the two critical angles $13.56^\circ < \theta_{il} < 29.2^\circ$ the echo reaches nearly 50 % of the incident wave (Figure 3.22b).

3.2.5 Attenuation

So far, we have imagined the ultrasonic wave as a plane wave that does not attenuate in amplitude (decrease in energy) as the wave propagates. In application, however, the amplitude of any propagating wave diminishes with distance traveled, by the following mechanisms:

Absorption: Fundamental dissipation of the energy in the form of heat.
Scattering: Refraction, reflection and diffraction of the wave at discontinuities at the surface of, or within, the medium.
Beam spreading (divergence or geometric attenuation): Geometric expansion of a nonplanar wavefront.
Dispersion: Difference of velocities of different wave modes.

Copyright © Marcel Dekker, Inc. All rights reserved.

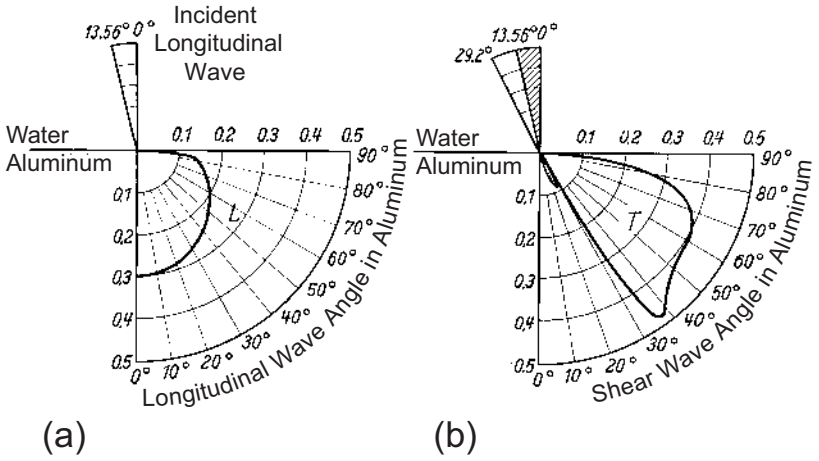


FIGURE 3.22 Echo plots for a water/aluminum boundary: (a) longitudinal and (b) mode converted transverse transmitted waves. (Adapted from J. Krautkramer, H. Krautkramer. Ultrasonic Testing of Materials. 4th ed. Berlin: Springer-Verlag, 1990, p 30. Reprinted by permission from Springer-Verlag.)

Arguably the biggest problem for most ultrasonic testing is low signal-to-noise ratio. Practitioners who understand the mechanisms of attenuation can maximize the signal-to-noise ratio for their application. The knowledgeable practitioner can also determine the limits of attenuation, such as minimum detectable flaw size and maximum usable signal penetration through a material. Table 3.4 compares the competing frequency dependence of the three competing attenuation mechanisms: absorption, scattering, and beam spreading. Choosing the frequency for a particular application requires practitioners to balance among competing loss mechanisms—a common requirement for engineering problems.

TABLE 3.4 Rule of Thumb for Loss Mechanisms as a Function of Frequency

Low frequency (large wavelength)	High frequency (small wavelength)
Large beam spreading losses (large beam divergence)	Large scattering losses ($\lambda \cong$ or $<$ particle size)
Low scattering losses ($\lambda \gg$ particle size)	Low beam spreading losses (Highly collimated beam)

Assume a fixed aperture and scattering particle size.

Copyright © Marcel Dekker, Inc. All rights reserved.

This section discusses the physical and mathematical characteristics of absorption, scattering, beam spreading, and briefly describes dispersion.

Physical Characteristics of Attenuation

Physically, we understand that all traveling waves eventually cease to exist. Where does the vibrational energy go? As with any moving mass, the energy of motion is eventually converted into heat energy, and thus is unavailable for continued motion. Many mechanisms exist by which the energy driving the momentum transfer of an ultrasonic wave, alternating between kinetic and potential like an oscillating Slinky spring, is converted into heat energy. Generically, these processes of energy conversion are called *absorption*.

A traveling wave periodically displaces the material from an equilibrium position. Absorption acts as a damping, or braking, force on these vibrating particles. Higher-frequency waves have a shorter wavelength (recall $v = f\lambda$), resulting in a higher total displacement per length as the wave travels through a material. Therefore, you would expect higher-frequency oscillations to attenuate at a higher rate than lower-frequency oscillations.

Experiment: On a stereo, play music with a little added bass (low frequency vibrations). Go into the next room and listen to the music that travels through the wall. How did the sound of the music change?

Scattering is the response of a propagating wave as it encounters the structure or variations of the material—like ocean waves breaking up on a pier. *Inhomogeneity* is the degree of structure or variations within a material, such as porosity, inclusions, phase changes in metals or ceramics, the grain structure in wood, and the change from one constituent material to another in composite materials. When a wave encounters these material variations, it will reflect, refract, or mode-convert according to the angle of incidence, the change in density, and the change in elastic properties (see Section 3.2.4). In many materials, inhomogeneities are random, whether in their geometry or in their spacing throughout the material. Consequently, the wave scatters—that is, part of the energy from the original (main) wave is sent randomly into different directions within the material.*

In pure scattering, the energy of motion is not converted to heat; instead, it is diverted into waves traveling in directions other than the main wave. These diverted waves appear as *grass* (random noise), potentially obscuring the signal from the main beam (Figure 3.23).

The amount of wave energy that is scattered depends on the size of the scattering particles compared to the wavelength of the ultrasonic wave.

---

*Statistical methods determine the overall scattering effect that these material variations have on the wave.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

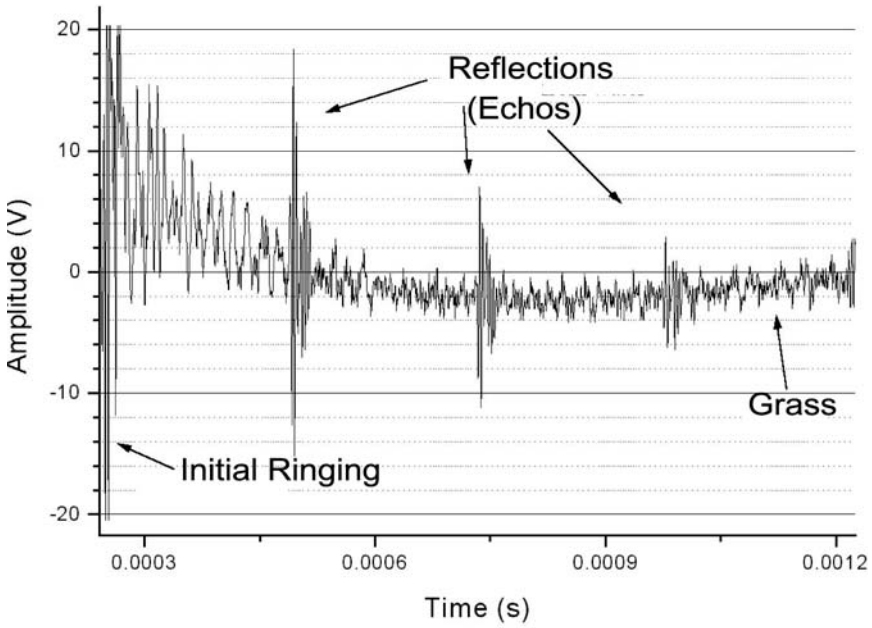


FIGURE 3.23 Response of ultrasound in a coarse grain material. Scattered waves appear as noise (grass) superimposed on the multiple front- and back-face echos.

Consider: How is sound scattered off a canyon wall (large particle) different from sound scattered off a forest of trees (small particle)? The wavelength of audible sound in air from a human voice is on the order of 0.6–1.5 m.

The diagram in Figure 3.24 depict three different regions of wave scattering as a function of particle size and wavelength. The graph shows only the energy that is *backscattered* in the opposite direction of the original wave. The particle diameter is a; the wavelength is λ, and the spatial frequency (wave number) is k.

Region 3 $2\pi a \gg \lambda$ ($ka \gg 1$)*:* In this region, because the particle is much greater than the wavelength, the amount of the wave that is scattered and the directions of the diverted waves are governed by the geometric conditions discussed in Section 3.2.4.

Region 2 $2\pi a \cong \lambda$ ($ka \cong 1$)*:* This is a transitional region between geometric scattering and Rayleigh scattering. At values of a comparable to λ, interface waves can propagate along the surface of the boundary on the particle.

Copyright © Marcel Dekker, Inc. All rights reserved.

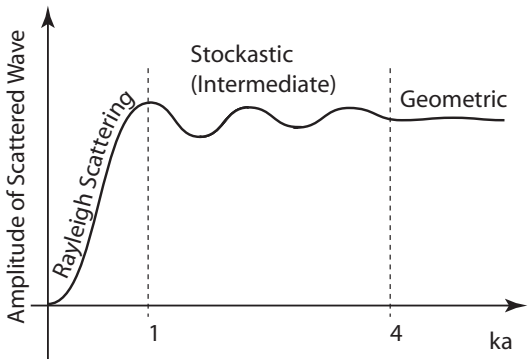


FIGURE 3.24 Scattering of waves as a function of particle size and wavelength. Only the energy scattered directly backwards towards the source is plotted.

This region is characterized by an oscillating magnitude (resonance) of the scattered waves. Note that the maximum scattered energy occurs when $2\pi a = \lambda$.

Region 1 $2\pi a \ll \lambda$ ($ka \ll 1$)*:* In this region the amplitude of scattering has a strong dependence on the size of the particle for a given wavelength of incident ultrasound. Scattering in this region is approximately omnidirectional and is called *Rayleigh scattering*. Ultrasonic measurements to determine particle size are commonly performed in this region.

Both absorption and scattering cause an attenuation of the main beam that imposes a practical limit on ultrasonic testing. But the two mechanisms are fundamentally different (see Table 3.5). Krautkramer et al., presents a clear description indicating the distinction:

> Pure absorption weakens the transmitted energy and the echoes from both the flaw and the back wall. To counteract this effect, the transmitter voltage and the [receiver] amplification can be increased, or the lower absorption at lower frequencies can be exploited. Much more awkward, however, is the scattering. In the [pulse-] echo method, it not only reduces the height of the echo from both the flaw and the back wall, but in addition, produces numerous echoes with different transit times, the so-called *grass*, in which the true echoes may get lost. The scattering [grass] can be compared with the effect of fog in which the driver of an automobile is blinded by his own headlights and is unable to see clearly. Stepping up the transmitter voltage [high beams] or the [receiver] amplification because the grass increases simultaneously cannot counteract this disturbance. The only remedy is to use lower frequencies.(24)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

TABLE 3.5 Mechanisms by Which an Ultrasonic Wave Amplitude is Attenuated: Physical and Mathematical Representations

Amplitude reduction mechanisms	Physical representation	Mathematical representation	Comments
Absorption	Energy conversion (kinetic energy to heat energy)	$e^{-\alpha x}$	Energy of motion is converted to heat and therefore, unavailable to propagate the wave.
Scattering	Energy diversion into other waves (geometric)	$e^{-\alpha x}$	The energy of motion is diverted from the ''main'' wave into waves traveling in different directions.
Beam spreading	Energy is redistributed over a different area but remains part of a single wave.	$\frac{1}{r^2}$ (spherical wave) $\frac{1}{r}$ (cylindrical wave)	A stone is dropped into the water. In each successive ring of rising and falling water, the same energy is distributed over a larger ring than the previous. Consequently, the amplitude of the water displacement must diminish as the wave propagates radially.

Mathematical Characteristics of Attenuation

Absorption and Scattering. Mathematically, attenuation by absorption or scattering of a ultrasonic wave can be represented as a decaying exponential. Assume an acoustic plane wave propagates in a material. Attenuation (scattering or absorption) is simply the drop in acoustic energy ($I - I_0$) or pressure ($P - P_0$) over a distance Δx. The mathematical relationships are, respectively,

$$I = I_0 e^{-\alpha_I \Delta x} \tag{3.90}$$

and

$$P = P_0 e^{-\alpha \Delta x} \tag{3.91}$$

where I_0 = initially measured acoustic intensity, I = measured acoustic intensity at a distance Δx from the initial measurement point, α_I = attenuation coefficient

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

for acoustic intensity, $P_0 =$ initially measured acoustic pressure amplitude, $P =$ measured acoustic pressure amplitude at a distance Δx from the initial measurement point, and $\alpha =$ attenuation coefficient for acoustic pressure. Recall that $I \propto P^2$, therefore, $\alpha_I = 2\alpha$. Whether α_I or α is used depends on the specific reference source.

The attenuation of the acoustic pressure of the wave over a distance Δx is

$$\alpha\Delta x(\text{Np/m})(\text{m}) = -\ln e^{-\alpha\Delta x} = \ln\frac{P_0}{P}(\text{Np}) \tag{3.92}$$

where Np is the abbreviation for nepers—the unit of attenuation of a field variable such as acoustic pressure or electric field. Notice the inversion of ratio of P/P_0 to cancel the negative sign for the attenuation $\alpha\Delta x$.

Another method, one preferred by ultrasonic technicians, is to measure attenuation in decibels (dB) [25–27]. The decibel (10 bels) is based on the base 10, logarithmic scale of the ratio of *energies* or *powers*:

$$\frac{I}{I_0} = 10^{\text{dB}/10} \quad \text{or} \quad \text{dB} = 10\log\frac{P^2}{P_o^2} = 20\log\frac{P}{P_o} \tag{3.93}$$

Therefore, a -20 dB system would be one where the intensity was 10^{-2} (or 0.01) times the original intensity. Engineers often remember that a -3 dB change is equal to $\frac{1}{2}$ the intensity, this is due to the fact that $10^{-0.3} = \frac{1}{2}$ (try this on a calculator).

The attenuation can then be defined in decibels per meter:

$$\alpha_{\text{dB/m}}\Delta x = 10\log_{10}\left(\frac{I_0}{I}\right) = 10\log_{10}\left(\frac{P_0^2}{P_2^0}\right) = 20\log_{10}\left(\frac{P_0}{P}\right) \tag{3.94}$$

We relate the attenuation coefficient α, *derived* from Eq. (92), to the *defined** attenuation constant, $\alpha_{\text{dB/m}}$, in Eq. (94);

$$\begin{aligned}\alpha_{\text{dB/m}}\Delta x &= 20\log\left(\frac{P_0}{P}\right) = 20\log(\text{e}^{\alpha\Delta x}) \\ &= 20\alpha\Delta x\log(\text{e}) \\ &\cong 8.686\alpha\Delta x \\ \alpha_{\text{dB/m}} &= 8.686\alpha\end{aligned} \tag{3.95}$$

Geometric attenuation. Because all ultrasonic sources (or, for that matter, the sources for any wave) are finite in extent, a true plane wave cannot

*A. G. Bell defined the unit of bel from experiments preformed on the range of human auditory response. His experiments showed that for one's perception of sound to increase by a factor of two required a tenfold increase in sound intensity.

Copyright © Marcel Dekker, Inc. All rights reserved.

exist. (The shapes of common wave fronts are discussed in Beam Characterization in Section 3.3.4) Because of this, as a wave propagates, it must either converge or diverge; as it does so, the wave amplitude will *geometrically attenuate*, or amplify. For example, a pebble dropped into a pool of water produces a circularly divergent wave. As the wave propagates along the surface, the displacement of the wave decreases with increasing distance from the source. Assuming low absorption losses (a reasonable assumption) and no scatters on the water surface, the energy in the first ring is equal to the energy in each of the subsequent rings. As the diameter of the waves increases, however, the energy is spread out over a larger area, and the displacement decreases. With a focused source or receiver, the converse would occur. For example, an old-style ear trumpet (hearing-aid horn) increases displacement by focusing the incoming wave.

Attenuation by Dispersion

In many ultrasonic methods, more than one wave mode can propagate at the same time, even when generated from the same source. Because the velocities of the modes are different, transit times between generation and detection are different. Therefore, the total wave energy spreads out over time with increasing transit distance. This phenomenon, known as *dispersion*, reduces the apparent wave amplitude. Dispersion is fundamental to guided modes, and inherent in methods such as laser generation of UT, where the UT wave contains a very broad range of frequencies.

3.2.6 Guided Waves

Earlier in this chapter, we considered bulk waves in and of themselves, disregarding any boundaries that a wave might encounter (Waves in Three Dimensions in Section 3.2.2). We then considered reflection and refraction, which occur when a boundary changes the amplitude and direction of a wave, and introduced critical angles (Critical Angles in Section 3.2.4). In this section, we use this information to consider more carefully what happens to waves as they are critically refracted: in a semi-infinite medium (single boundary) they are bound to the surface, and in a finite medium (two boundaries) they are bounded within the medium and may resonate.

Surface Acoustic Waves (SAWs)

An interface between two objects—such as skipping a stone across the surface of a lake—allows for the creation of a surface acoustic wave (also known as a surface-bound wave). First discovered by Lord Rayleigh in 1885, SAWs have unique properties that are a distinct mix of longitudinal and shear waves. In general, an SAW propagates on the surface of the object, and has an amplitude that decays exponentially with depth into the object. SAWs can propagate for

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

very long distances. There are many different types of SAW, depending on the nature of the materials at the interfaces; the most common are Rayleigh, Scholte, Stoneley, Love, and creeping (28, 29).

Rayleigh Waves. If an SAW travels along a boundary between a semi-infinite solid medium (substrate) and air, the wave is often called a Rayleigh wave. In fact, the term Rayleigh wave is so popular that many scientists (incorrectly) call all SAWs Rayleigh waves. Although a Rayleigh wave starts out as a shear wave, the interface between the substrate and the air causes the wave to transform into a wave that has both transverse and longitudinal wave characteristics (Figure 3.25). Rayleigh waves follow curved surfaces, so they are ideal for finding surface or very near surface defects in curved objects.

The velocity of the Rayleigh wave is given by the root of the Rayleigh equation (30), which can be approximated by

$$\frac{v_r}{v_t} \cong \frac{0.87 + 1.12v}{1 + v} \tag{3.96}$$

As the Poisson's ratio v varies from 0 to 0.5, the Rayleigh wave velocity, v_r varies monotonically from $0.87v_t$ to $0.96v_t$, where v_t is the velocity for transverse (shear) waves.

Rayleigh waves propagating around curved surfaces and corners may encounter velocity changes, as well as some reflection. Where the ratio of radius-of-curvature to wavelength is large, Sezawa (31) has shown that the speed is the same as on a flat plate. As the ratio decreases, however, the velocity increases, such that, for example, if the ratio is 3.5, the Rayleigh wave velocity is approximately equal to that of the transverse wave.

Scholte Waves. Scholte waves propagate on a water–solid interface. The velocity of a Scholte wave is nearly identical to the velocity of a Rayleigh wave. In fact, for typical values of liquid and solid, the velocity changes by less than 0.1 % with the addition of water. The biggest difference between Rayleigh waves

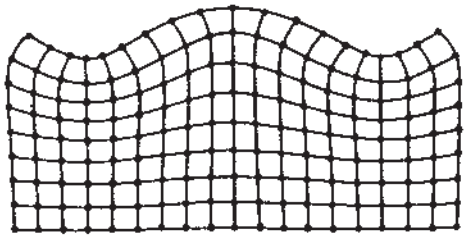


FIGURE 3.25 Surface displacement due to Rayleigh wave: note both longitudinal and shear characteristics.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

and Scholte waves is in their attenuation. In general, Rayleigh waves have very low attenuation, so they are very useful in the field of surface detection. Scholte waves, on the other hand, have very high attenuation. For typical values, the amplitude of the wave decreases to $1/e$ (approximately $\frac{1}{3}$) times its original amplitude when the wave has traveled only 10 wavelengths. Scholte waves are also often called *leaky surface waves*; the energy of the wave leaks away from the surface. This effect has been used to determine if water is present on a material.

If the water layer is thin and has a constant thickness, then the wave that attenuates into the water can be reflected back to the surface and create another surface wave at the interface.

Stoneley Waves. Stoneley waves propagate at the interface between two solid media. For certain specific combinations of material parameters, a wave can become tightly bound to an interface (32). The field distribution is composed of two partial waves decaying away from the surface in each medium.

Numerical computation is required in order to evaluate the wave propagation velocity and the field distributions, but there are some simple guidelines about the existence of Stoneley waves. For example, the Stoneley velocity must lie between that of Rayleigh waves and of shear waves in the denser medium. The solutions are very sensitive to the ratio of longitudinal to shear velocity in the more dense of the two media.

Love Waves. Love waves occur in inhomogeneous surfaces. The original impetus for studies in this area comes from seismology. One of the established facts of seismology is the presence of large transverse (*horizontal plane*) components of displacement in the main tremor of an earthquake. Such displacements, however, are not a feature of Rayleigh waves, which contain displacements only in the vertical plane. It follows that the actual conditions in the earth must differ in some essential aspect from those of a homogeneous isotropic half-space. Love (33) suspected that such waves were a consequence of a layered construction of the earth and that these waves consisted of shear horizontal waves trapped in a surface layer and propagated by multiple reflections within that layer.

Longitudinal Creeping Waves. If a longitudinal wave is critically refracted (Oblique Incidence in Section 3.2.4) onto the interface between two semi-infinite substances, a *longitudinal creeping wave* (or *head wave*) travels along the interface at the longitudinal wave speed. (The term *creeping wave* is a misnomer, since the wave actually travels with velocity of a longitudinal wave.) Creeping waves decay rapidly as their energy is mode converted to shear waves. These waves propagate on the order of centimeters, do not follow the contour of curved surfaces, and are not affected by surface roughness. Nonetheless, (leaky) creeping waves are not uncommon in NDE. Note, however, if a shear wave is created on an interface, then the wave will become a SAW.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Bounded Waves

Ultrasonic waves in objects of finite size can create bounded waves, due to the borders (bounding) of the finite medium. Practitioners refer to bounded waves as *plate waves* if the structure is a multilayer, and as *Lamb waves* if it is a single layer. Consider a longitudinal wave traveling at an angle through a plate of thickness d (Figure 3.26).* When the longitudinal wave in the plate hits a surface, both a longitudinal and a shear wave can be reflected (see Reflection and Refraction at an Interface in Section 3.2.4). These waves then approach the other surface and can split into two additional waves upon reflection. This pattern is then repeated ad nauseam. It is clear that this problem approaches chaos. However, the waves interfere with each other so that certain resonant waves occur. These resonant waves, which we call bounded waves, are either symmetric around the center of the plate or antisymmetric around the center (Figure 3.27). Not surprisingly, they are combinations of shear and longitudinal waves.

There are an infinite number of harmonic symmetric (S_x) and antisymmetric (A_x) modes in the plate. (The subscript x distinguishes the specific harmonic.) The velocity of the Lamb wave depends on the thickness of the plate, the frequency of the wave, the properties of the material and the mode of the wave. The equations

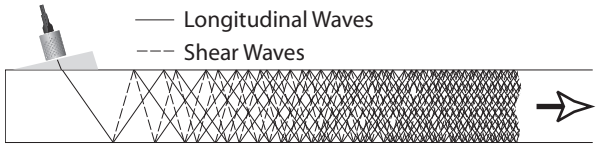


FIGURE 3.26 Creation of Lamb wave by reflecting a longitudinal wave.

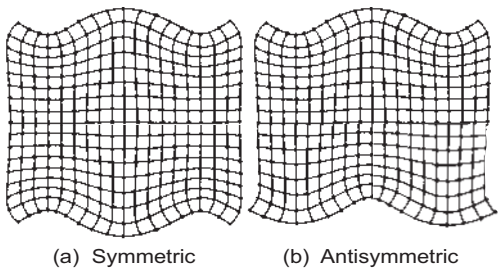


FIGURE 3.27 Lamb wave particle displacement is either (a) symmetric or (b) antisymmetric about the center plane of the plate.

*A shear vertical (SV) wave can also create a plate wave by the same process.

Copyright © Marcel Dekker, Inc. All rights reserved.

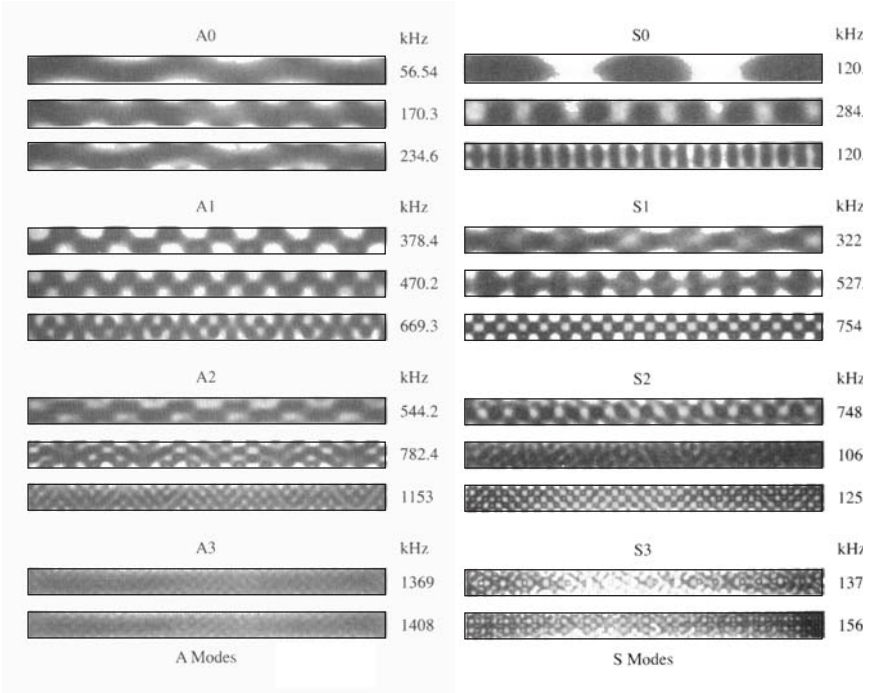


FIGURE 3.28 Visualization of the first few symmetric and antisymmetric Lamb modes traveling in a glass plate. (HU Li, K Negishi. Visualization of the Lamb Mode Patterns in a Glass Plate. Ultrasonics 32:4:243–250. Reprinted by permission.)

relating these facts are complicated to derive and require numerical methods to solve. Photographs of the first few modes for symmetric and antisymmetric modes in plate waves in a glass plate (with photoelastic methods*) are shown in Figure 3.28.

The relationship between the velocity, v and the thickness, d, times the angular spatial frequency, k are called dispersion curves. Dispersion curves are very useful for the creation of Lamb waves. Recall from Snell's law:

$$\frac{\sin\theta_1}{v_1} = \frac{\sin\theta_2}{v_2} \tag{3.97}$$

If one wished to create a specific wave at a specific frequency and thickness, one could determine the velocity of that wave using dispersion curves (Figure 3.29).

*Photoelastic method uses the stress induced birefringent (two speeds of light) nature of certain transparent materials to optically distinguish between different stresses.

Copyright © Marcel Dekker, Inc. All rights reserved.

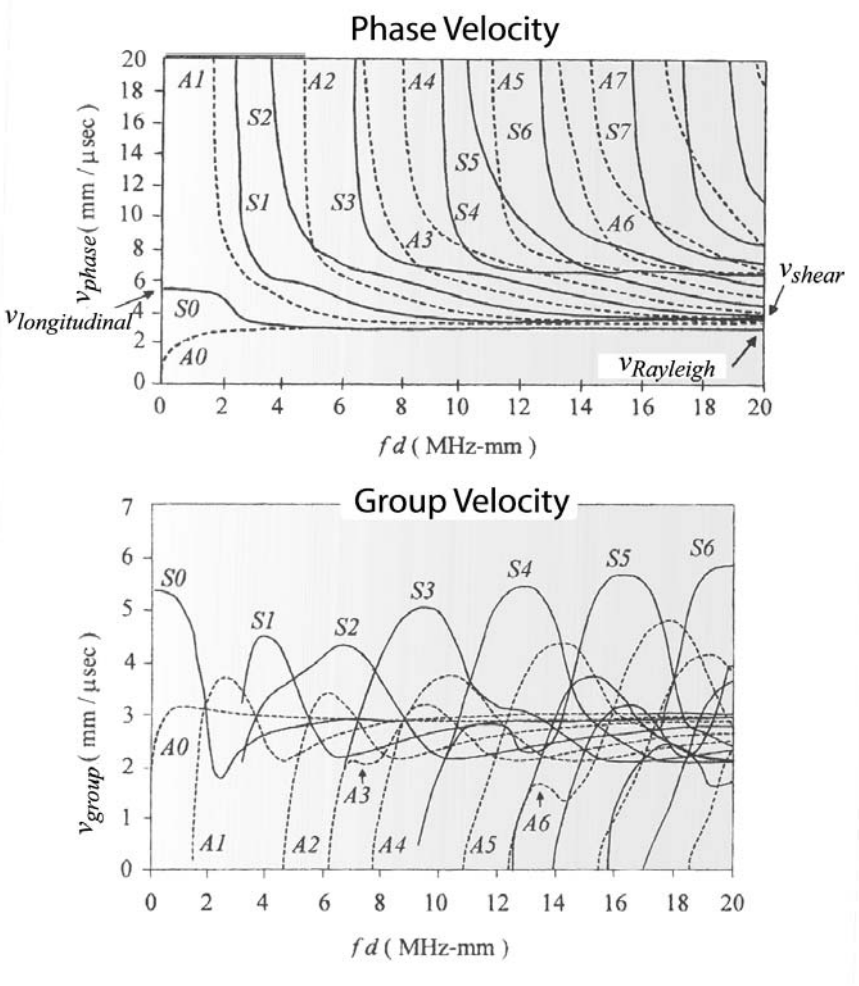


FIGURE 3.29 Dispersion curves for Lamb modes propagating in a plate. (See Section 3.6.2 for details of the group velocity.)

To create such a wave with a transducer on a plastic wedge, one could then determine the angle of the wedge from Snell's law, where the angle in the plate is 90°.

Example: Imagine that one wished to create the S_1 wave in an aluminum plate 4 mm thick with a 1 MHz transducer. If the velocity in plastic is 3 km/ms, what angle of plastic wedge would be required?

Answer: According to Figure 3.29, if the frequency times thickness is 4 MHz-mm, then the S_1 wave has a velocity of approximately 6 mm/μs.

Copyright © Marcel Dekker, Inc. All rights reserved.

From Snell's Law:

$$\sin\theta = \frac{3}{6}\sin 90^\circ = 0.5$$
$$\sin^{-1}(0.5) = 30^\circ$$

3.3 TRANSDUCERS FOR GENERATION AND DETECTION OF ULTRASONIC WAVES

Where do ultrasonic waves come from? In the previous sections, we did not discuss the origins of ultrasonic waves, or how to detect them; we simply assumed an ultrasonic wave existed, and worked from there. But in reality, ultrasonic waves are created (or *transduced*) from electrical or optical signals by *ultrasonic transducers*. These devices also detect ultrasonic waves by transducing the ultrasonic waves back into electrical or optical signals. Because a transducer functions as both the source and the detector of ultrasound in a medium, the limitations of a given transducer dictate our ability to use ultrasound in general. Various transducers offer different methods for creating ultrasonic waves; the most common are piezoelectric, EMAT, and laser (optical) methods.

3.3.1 Piezoelectric Transducers

Piezoelectric transducers are the most common method used today for creating and detecting ultrasonic waves. In the *direct piezoelectric effect*, discovered in 1880 by Jacques and Pierre Curie, a piezoelectric material responds to a mechanical deformation by developing an electrical charge on its surface. The direct piezoelectric effect relates the applied mechanical stress to the output charge and is a measure of the quality of the material to act as a receiver. The reverse phenomenon, the *indirect piezoelectric effect* (discovered by Lippmann in 1881), produces a mechanical deformation when the piezoelectric material is subjected to an electric field. The indirect piezoelectric effect relates the input charge on the material surface to the output strain, and is a measure of the quality of the material to generate ultrasound.

The quality of a particular piezoelectric material to act as a good receiver or as a good generator of ultrasound depends on its specific material properties. Let α and β represent the direct and indirect piezoelectric coefficients such that $V_{\text{output}} = \alpha\Delta d_{\text{applied}}$ and $\Delta d_{\text{output}} = \beta V_{\text{applied}}$, where Δd = thickness and V is a voltage. Ideally, a transducer material would be both a good receiver and a good generator of ultrasound. But α and β are competing effects, so some compromise is required. Typically, materials are chosen with higher α to improve reception, and simply pulsed at a higher voltage to generate the ultrasound. If separate transducers are used for generation and reception, the piezoelectric crystals can be

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

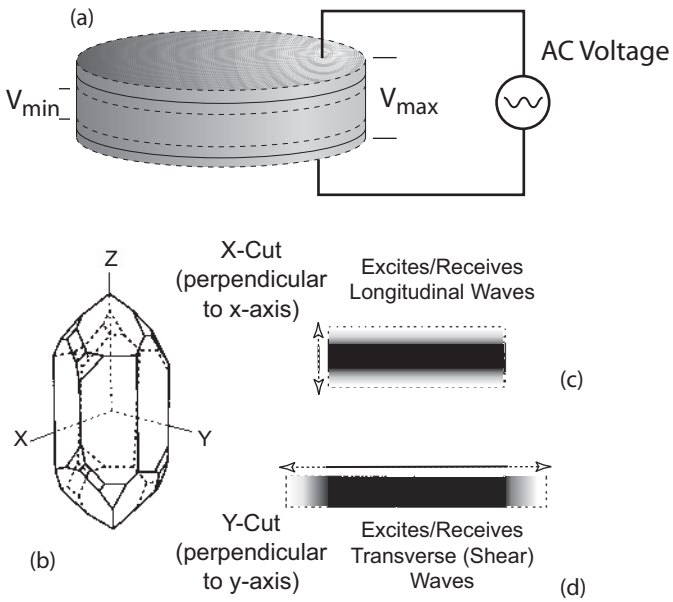


FIGURE 3.30 (a) Response of a piezoelectric disk to an alternating voltage. Piezoelectric crystal: (b) axis definition, (c) *X-cut* disk vibrations are longitudinal, and (d) *Y-cut* disk vibrations are transverse.

chosen to match their function as in a pitch–catch element transducer. For example, PVDF, a polymer film used in the packing industry, has an exceedingly high α value, and acts as excellent receiver material. However, its low mechanical durability limits its use.

To understand piezoelectric generation of ultrasound, consider a thin disk of piezoelectric material with electrodes attached to its surface (Figure 3.30a). If an alternating voltage pulse is applied to the faces, the disk contracts and expands, creating an ultrasonic wave. Conversely, if an ultrasonic wave passes through the disk, the disk contracts and expands, creating an alternating electric field (i.e., a voltage pulse) across the electrodes.

In general, piezoelectric transducers are set up to vibrate at the fundamental frequency, defined as:

$$f = \frac{v_{\text{crystal}}}{2d} \tag{3.98}$$

where d is the thickness of the material and v_{crystal} is the sound velocity in the crystal. Note that this frequency is the same as the wavelength equaling twice the thickness of the material.

Copyright © Marcel Dekker, Inc. All rights reserved.

Piezoelectric materials can occur naturally or artificially.* Quartz and tourmaline are perhaps the best known naturally occurring crystals that exhibit piezoelectric effect; modern (manmade) piezoelectric materials include ceramics such as barium titanate (BaTi), lead zirconate titanate (PZT), and lead metaniobate (PMN). PVDF, a polymeric film, and composite transducers are also gaining acceptance in the production of piezoelectric transducers. Unlike naturally occurring piezoelectric crystals, modern ceramics and films need to be *poled* to align the randomly oriented piezoelectric (ferroelectric) domains, just as we might pole a magnet to align the magnetic (ferromagnetic) domains. Piezoelectric materials are poled by applying strong electric fields just below the *Curie temperature* and holding the materials in the field as the temperature is slowly reduced. Above the Curie temperature, the materials (just like magnets with the magnetic Curie temperature) are not piezoelectric.

Piezoelectric Transducer Characterizations

We can categorize piezoelectric transducers by the type of wave generated, by the beam orientation, by the environment where used, and in other ways. Some of the most common classifications are:

1. Type of wave(s) generated (longitudinal, shear or surface)
2. Contact, immersion, or air-coupled
3. Normal or angled beam
4. Single or multiple elements
5. Transducer face—flat (normal) or shaped (contoured or focused)
6. Broadband or narrowband
7. Ambient or high temperature

Type of Wave(s) Generated (longitudinal, shear, or surface). If the transducer material is made from a quartz crystal, then the way the crystal is cut determines whether it is a longitudinal or shear transducer. In piezoelectric crystals, the x-axis is defined as the axis that expands and contracts (Figure 3.30b). In *X-cut crystals* (longitudinal), the crystal is cut perpendicular to the x-axis (so that the surfaces are y–z surfaces), and the vibrating crystal creates a longitudinal wave (Figure 3.30c). In *Y-cut crystals* (shear), the system is cut perpendicular to the y-axis (so that the surfaces are x–z), and the vibrating crystal creates a shear wave (Figure 3.30d).

Contact, Immersion, or Air-Coupled. Piezoelectric transducers can also be defined by their coupling mechanisms: contact, immersion, or air-coupled. Contact transducers, although they appear to come into direct contact with the test

*Many common materials exhibit a weak piezoelectric effect: bone, wood, and even ice. Also, some wintergreen breath mints have a piezoelectric effect—if you break them in the dark, you can see sparks.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

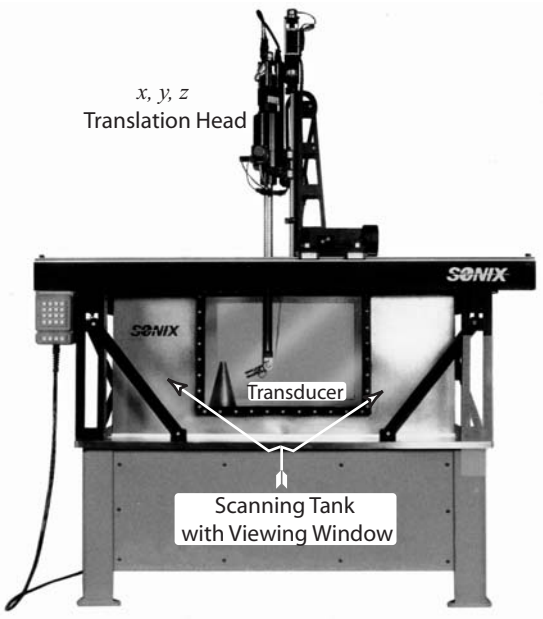


FIGURE 3.31 UT scanning tank. (Courtesy Sonix.)

specimen, in fact require some form of liquid or dry couplant to adequately couple the ultrasound to and from the test specimen.* Immersion transducers operate in a scanning tank where the specimen and the transducer are fully immersed with in the fluid (typically water). In this configuration, scanning can occur with varying stand-off distances (Figure 3.31).

Because of the high impedance mismatch between the piezoelectric crystal and air, air-coupled transducers require a special layer of material to help send a pulse through air to the sample. This layer, called an *impedance matching layer*, should have an impedance that is intermediate between the acoustic impedances of the transducer and the air. In addition, the thickness of the matching layer can be optimized for maximum transmission. Imagine an ultrasonic beam that passes directly through the sample and one that reflects off the surfaces of the matching layer. The beam that reflects off the matching layer/transducer interface has a phase change of 180°. In addition, that beam has a phase difference due to the distance traveled. The total phase difference, ϕ, is thus

$$\phi = 360^\circ \frac{2h_m}{\lambda} + 180^\circ$$

*Contact transducers can be also glued directly to the surface and in very special cases, direct contact without coupling maybe used.

Copyright © Marcel Dekker, Inc. All rights reserved.

where h_m is the thickness of the matching layer. Constructive interference occurs if the phase is 360°, so that:

$$360^\circ \frac{2h_m}{\lambda} = 180^\circ$$
$$h_m = \frac{\lambda}{4} \tag{3.99}$$

Since the matching layer has a thickness that is a quarter of the wavelength, it is often referred to as a *quarter wave-matching layer.*

Normal or Angled Beam. Transducers can also be categorized as normal beam or angled beam. *Normal beam* transducers transmit a wave normal to the face of the transducer, whereas *angled beam* transducers transmit a wave at an angle (not normal) to the face of the transducer. Angled beam or *refracting* transducers are usually just longitudinal normal-beam transducers attached to a wedge (often made of plastic). It is important to remember that, due to Snell's law, the angle in the wedge is not the same as the angle of the beam created in the sample. Angled-beam transducer manufacturers commonly label their transducers according to the angle of injection; when they do, they assume a steel test piece (Figure 3.32). For operation on any material other than steel, the correct angle must be calculated anew. In addition, the incident longitudinal beam within the wedge arrangement can create both shear and longitudinal waves, as well as surface waves, in the sample (Figure 3.32 and Section 3.2.4). This existence of multiple waves can make it rather confusing to detect features. Because the shear and longitudinal velocities are different, the reflections off the same object will arrive at different times. To avoid this confusion, manufacturers produce wedge angles designed to operate between the first and second critical angle, to create only a mode-converted shear wave.

Single or Multiple Elements. So far, we have been discussing single-element transducers. But piezoelectric transducers commonly contain multiple (two or more) elements. *Dual-element transducers* (Figure 3.33) contain separate transmitter and receiver elements housed in a common case. This arrangement allows improvement of the near surface resolution (Section 3.4.3). A set of several normal piezoelectric transducers acting together is called an *array.* If, through either electrical or physical means, the transducers are driven at slightly different times, so that the individual waves interfere constructively and destructively into an angled wave (Figure 3.34), this is called a *phased array.* The individual transducers are usually as small as possible, so that the individual elements appear as point sources.

Flat (normal) or Shaped (shaped or focused). Previously, we have assumed a planar or flat transducer face. (The device is called a *flat* or *normal*

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

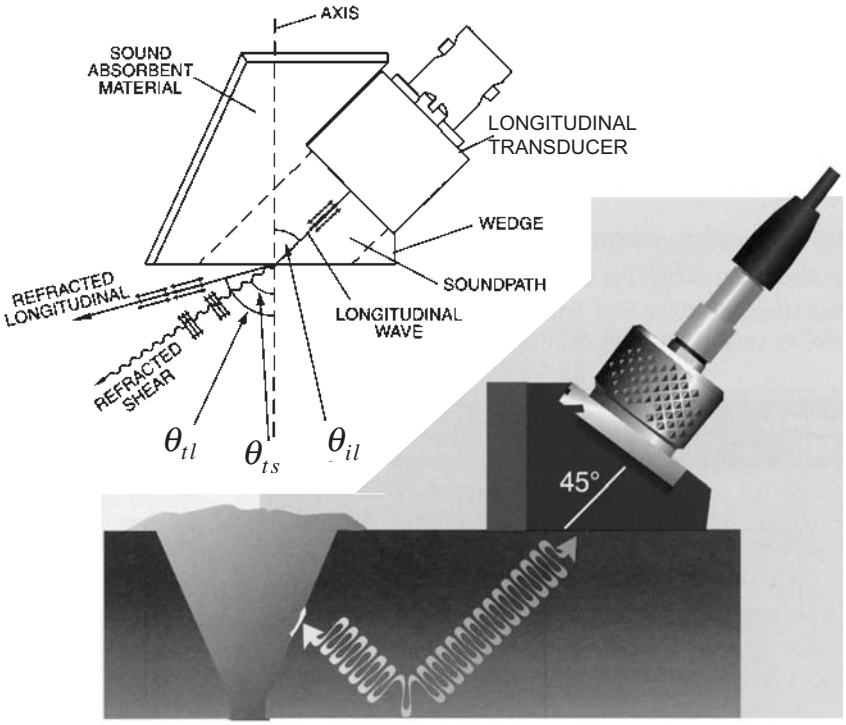


FIGURE 3.32 Angled beam transducers. (Adapted: Courtesy Panametrics.)

transducer.) While most applications call for flat transducers, shaped transducers are very common. Shaped transducers have two basic functions: matching the probe face to the contour of the sample (Figure 3.35a) or focusing the ultrasonic energy. Figure 3.35b shows two common types of focus, line and point. As with any geometric focusing, there will be a finite spot size that will limit resolution and a depth-of-field within which the spot size is focused (See Generating Focused Waves in Section 3.4.2).

Broadband or Narrowband. The choice of the transducer's range of operating frequencies (bandwidth and center frequency) depends greatly on the specific application. Materials with high attenuation, for example, require lower frequencies; UT interferometry works best with a narrow band transducer. Transducers that create or receive a large range of frequencies are called broadband; they are poor transmitters, but excellent receivers. Narrowband transducers have the reverse characteristics. Broadband transducers are commonly used for time-of-flight measurements. For more information on frequency response, see Section 3.3.4.

Copyright © Marcel Dekker, Inc. All rights reserved.

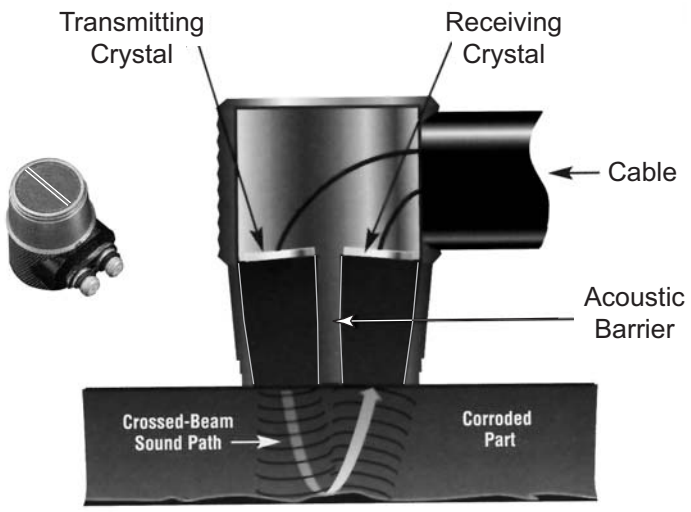


FIGURE 3.33 Duel element transducer: The receiver and transmitter are housed in the same case separated by sound isolating material. This arrangement improves near surface resolution. (Adapted: Courtesy Panametrics.)

Ambient or High Temperature. Because piezoelectric probes must operate well below their Curie temperature, other methods, such as optical UT, are used at very high temperatures. Standard piezoelectric probes typically operate between −20 and 60°C. Specially designed high-temperature delay lines extend this range to about 200°C. Actively cooled probes are also available.

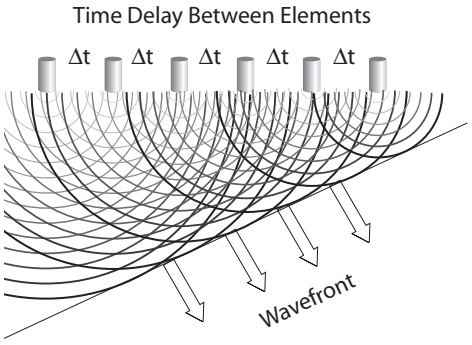


FIGURE 3.34 Phased array used to control the propagation angle.

Copyright © Marcel Dekker, Inc. All rights reserved.

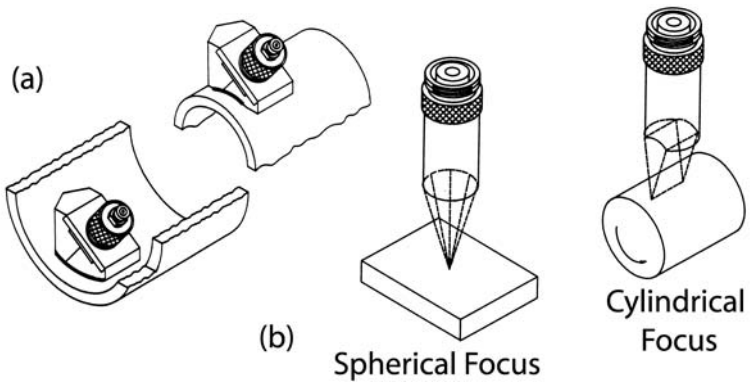


FIGURE 3.35 Shaped transducers: (a) face shaped to match sample contour and (b) face shaped to focus ultrasonic beam. (Courtesy Panametrics.)

Constructing Piezoelectric Transducers

Piezoelectric transducers contain three main components: an active element, a backing material, and a wear plate (Figure 3.36). In general, the piezoelectric materials are made into a thin disk, called the *active element* or piezoelectric crystal, whose thickness determines the natural frequency of the element. The active element is then coated with a metallic film (silver) and connectors are soldered to the faces. The back of the disk is then covered with a thick layer of backing material made of acoustically absorbing material. Thus sound leaving the backside of the active element will be absorbed thereby reducing unwanted signal (noise) caused by the sound reverberating within the transducer housing. Finally, the piezoelectric disk and backing are placed into a housing, and a thin (often plastic) face called a *wear plate* covers the exposed face of the active element.

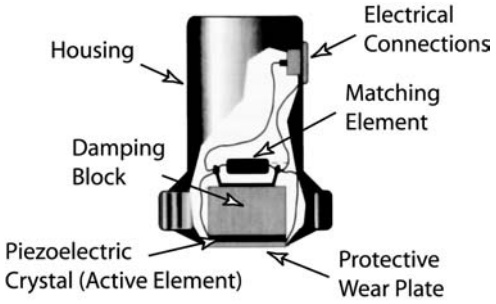


FIGURE 3.36 Basic piezoelectric transducer construction. (Courtesy Krautkramer Branson.)

Copyright © Marcel Dekker, Inc. All rights reserved.

Because transducers are commonly slid across the specimen surface, the housing design often accommodates replaceable wear plates. Occasionally transducers are connected to a curved surface lens to focus the beam or an impedance matching layer.

Transducer Coupling Medium

Piezoelectric transducers need some form of coupling medium to transmit the ultrasound between the transducer and the test specimen. The large acoustic impedance mismatch between the transducer and air prohibits any air gap between the sensor and the specimen.* Since the ultrasonic displacement is on the order of nanometers, it is difficult to make intimate contact between most transducers and the specimen, let alone perform the test in a timely manner. The simplest solution is immersion testing—immersing both the transducer and the specimen in a water bath. When immersion is not practical, contact transducers utilize either a dry or wet coupling material. Thin latex works well as a dry couplant. Grease, gels, or liquids applied manually or automatically work well as wet coupling materials. Special high-viscosity coupling material is designed for normal-incidence shear transducers. The quality of the signal transmission and reception will depend on the thickness of this couplant layer; a thinner layer usually improves efficiency. Because most commercial coupling materials are messy and difficult to remove, many people use honey instead. There are also devices intermediate to full immersion and contact coupling methods; for example, water squirters and bubblers create an acoustic coupling path along a short column of continuously flowing water (Figure 3.37).



FIGURE 3.37 Bubblers use a column of continuously flowing water to couple the ultrasound between the sample and the transducer. (Courtesy Panametrics.)

*Special air-coupled UT transducers exist, but they suffer from very low coupling effieiency.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

3.3.2 Electromagnetic Acoustic Transducers (EMATs)

Electromagnetic acoustic transducers (EMATs) allow for high-speed ultrasonic inspection, as well as for the interrogation of materials where surface contact is prohibited. These are accomplished by using a *noncontacting* coupling mechanism: EMATs generate and receive ultrasound by coupling magnetic fields to the particle movement within a test specimen, which must be conductive. EMATs can also generate shear horizontal (SH) waves.

The ultrasonic transduction mechanism for EMATs generates ultrasonic waves in a two-step process: (a) magnetic induction (a simple transformer), and (b) the force created when current flows through a magnetic field (Lorentz force).

Magnetic induction: Imagine two conductors are placed very close together. Conductor A is connected to a power source and current flows. As expected, we induce an image current that flows in conductor B. This process is the familiar transformer that is commonly used to transform voltages. The induced currents (commonly called *eddy currents*) alternate with the same frequency as the driving current in conductor A.

Lorentz force: Consider the arrangement shown in Figure 3.38, where conductor A is the EMAT (drive) coil and conductor B is the conductive test material in which the ultrasound is to propagate. By placing a permanent magnet in proximity to the material, we superimpose a static magnetic field onto the eddy currents flowing in the specimen. Let $\vec{H}$ and $\vec{J}$ represent the static magnetic field and the eddy currents (both are vector quantities). Subjecting the currents flowing in a material to a superimposed magnetic field creates a force (Lorentz force) that acts on the lattice structure of the material. The magnitude and direction of this force is given by

$$\vec{F} = \vec{J} \times \vec{H} = \vec{a}_n\left(|\vec{J}||\vec{H}| \sin(\theta_{JH})\right), \tag{3.100}$$

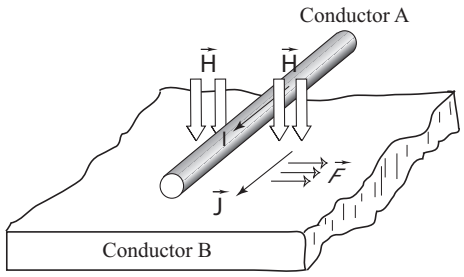


FIGURE 3.38 Schematic of an elementary electromagnetic acoustic transducer (EMAT).

Copyright © Marcel Dekker, Inc. All rights reserved.

where $\vec{a}_n$ points in a direction perpendicular to both $\vec{J}$ and $\vec{H}$, according to the right hand rule. (Holding the right hand open with the fingers pointing in the direction of $\vec{J}$, allow the fingers to curl towards the direction of $\vec{H}$ and the thumb (perpendicular to both) will point in the direction of $\vec{F}$.) θ_{JH} is the angle between $\vec{J}$ and $\vec{H}$. Now $\vec{H}$ is static, and $\vec{J}$ alternates with the frequency of the driving current in conductor A. Therefore, $\vec{F}$ also alternates in a direction perpendicular to and at a frequency identical to the driving current. This alternating force on the lattice structure of the material acts as a source for an ultrasonic wave.

To detect an ultrasonic wave, on the other hand, we simply reverse the process for generating a wave: particle motion in the presence of a superimposed magnetic field generates a current in the conducting material *B*, which in turn excites a current in conductor A. Thus the magnitude and frequency of the ultrasonic wave is transduced into a measurable current flowing in the EMAT (pickup) coil.

Now that we have a mechanism to generate (or receive) an alternating force within a conductive material, let us look at the specifics required to launch or receive a longitudinal or transverse ultrasonic wave. Figure 3.39 shows a typical longitudinal EMAT arrangement of a permanent magnet, an excitation coil, and a specimen. The magnet includes pole pieces that guide the magnetic field parallel to the specimen surface. The resultant orientations of the vectors $\vec{H}$, $\vec{J}$, are $\vec{F}$ shown. The magnetic field and the eddy current are mutually perpendicular, and form a plane parallel to the surface of the specimen. Consequently, the resultant force is perpendicular to the surface and drives a longitudinal ultrasonic wave. The main lobe of the longitudinal directivity pattern is also shown.

To generate a transverse wave (SH), we align the magnetic field of the permanent magnet normal to the surface (Figure 3.40). Two common configurations of magnets are considered: a single magnet and two magnets with alternating pole orientation. The excitation coil is oval-shaped, such that the current flows in a sheet along a constant direction beneath the footprint of the magnet. Because current flows in one direction in half of the coil and in the opposite direction in the other half, the magnets are placed with opposite polarity. In this configuration, $\vec{H}_0$ and $\vec{J}$ form a plane perpendicular to the surface of the specimen. The resultant alternating force that drives the shear horizontal transverse wave is parallel to the surface.

Figure 3.41 shows two additional transverse EMAT geometries: a spiral (pancake) and an oval excitation coil. Because the *entire* circular current pattern is exposed to the magnetic field, the diffraction pattern shows the wave propagating in the direction of the two side lobes, and not into a central lobe.

The pancake coil EMAT exemplifies the potential for beam steering—controlling the direction of the ultrasonic beam (wave). We can manipulate either the static magnetic field or the coil geometry to steer the beam. For example, surface and plate waves are readily generated by an EMAT with a single

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

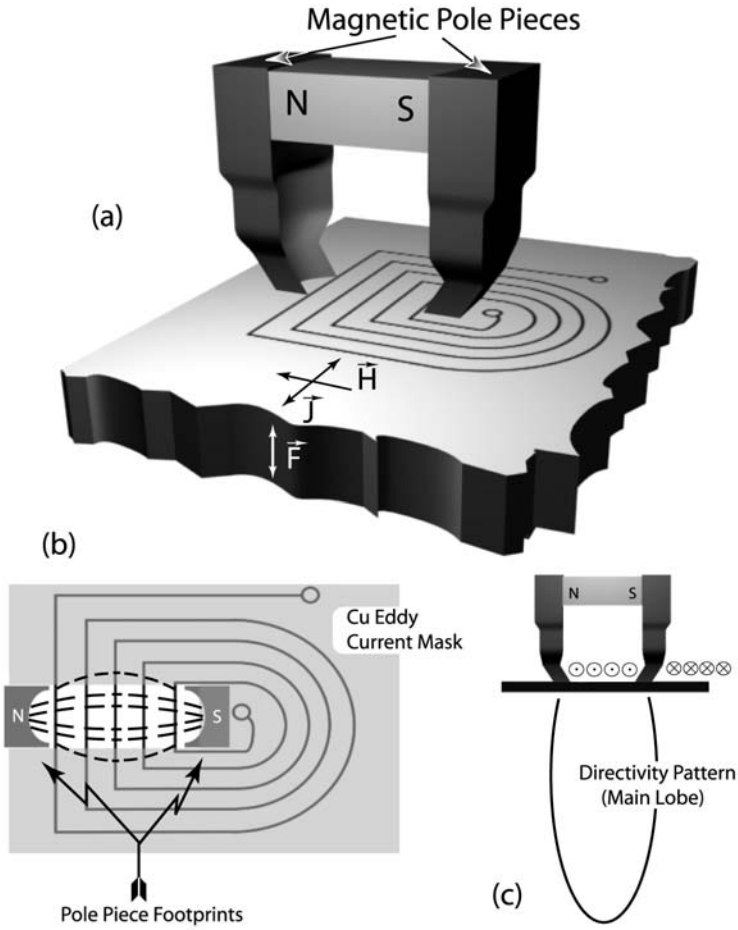


FIGURE 3.39 EMAT for generating longitudinal waves (a) 3-D view, (b) footprint depicting copper shield to remove unwanted eddy currents, and (c) directivity pattern. ($\odot$ and $\otimes$ indicate current flow out and in of the page plane.)

permanent magnet and a meander coil (Figure 3.42). The meander coil reverses direction with each half cycle; thus the EMAT acts as an array of ultrasonic generators. The spacing (periodicity) of the coil is matched to the frequency of the driving current, to steer the beam into a surface or a guided wave.

In practice, it is difficult to spatially orient the magnetic fields and the current to produce *only* the desired mode and direction of an ultrasonic wave. Both the magnetic field and the current must eventually create a loop. In the

Copyright © Marcel Dekker, Inc. All rights reserved.

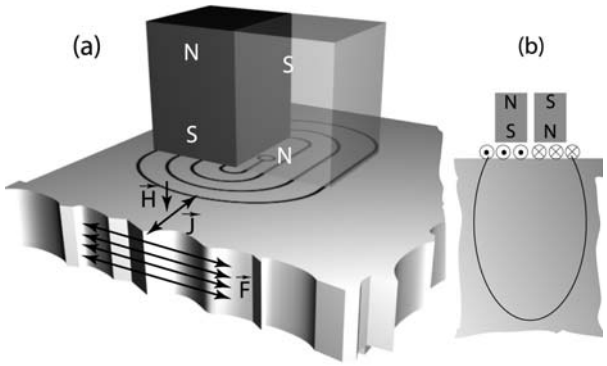


FIGURE 3.40 EMAT for generating transverse waves (a) 3-D view and (b) directivity pattern.

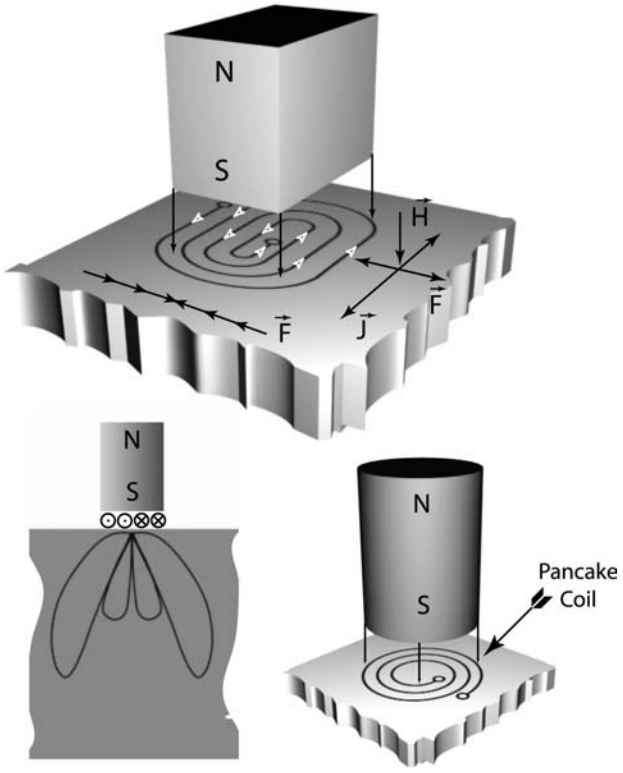


FIGURE 3.41 Alternative transverse EMATs (note the difference in the directivity pattern).

Copyright © Marcel Dekker, Inc. All rights reserved.

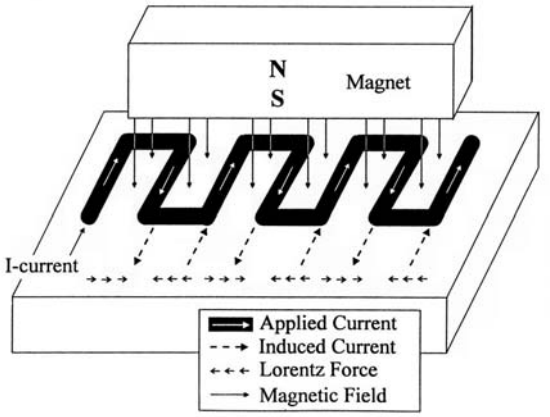


FIGURE 3.42 Meander coil EMAT for generating plate waves.

schematics given in Figures 3.39 through 3.42, only a portion of the arrangement is used to generate the desired wave. To reduce unwanted wave generation, we insert a conductive mask (e.g., copper tape) between the excitation coil and the specimen. The mask is grounded to the EMAT case. The unwanted eddy currents are generated in the mask and shorted to ground. A typical mask is shown in Figure 3.39b.

Low signal-to-noise ratio is an inherent drawback to EMAT operation. To maximize the transduction efficiency, EMAT *liftoff should be less than 1 mm* and the magnetic field strength should be maximized. High-field-strength permanent magnets (such as neodymium iron boron and samarium cobalt) or electromagnets are used. Additionally, impedance matching of the receiver and generation electronics is extremely important (34). Further information on EMATs theory, design, and application can be obtained in Refs. 35–37.

3.3.3 Laser (Optical) Generation and Detection of Ultrasound

Although laser-based ultrasonic testing has significant drawbacks—low sensitivity, susceptibility to vibrations, requirements of substantial operator training, possible damage to the sample surface, and very high cost—nevertheless its ability to operate at large standoff distances (measured in meters) makes it the only UT solution for some applications; in fact, laser-based UT is relatively common. Its two components, laser generation and laser detection, can be used together or separately (in combination with other UT transducers).

Copyright © Marcel Dekker, Inc. All rights reserved.

Laser (Optical) Generation

Laser generation of an ultrasonic wave might be thought of as exciting the wave with an optical hammer. A high energy pulsed optical beam (infrared with $\lambda \approx 1\ \mu$m) is focused onto the specimen surface. The interaction of the optical pulse with the sample occurs in one of two distinct processes, each producing unique acoustic wave patterns within the material—*ablative* and *thermoelastic*.* Whether the excitation falls into one category or the other depends on the optical energy density of the pulse, and on the material's ability to absorb the energy. In higher-energy and absorptive materials, the optical pulse *ablates* the surface as a local explosion (creating an ultrasonic pulse) and material is ejected from the surface (Figure 3.43a). Conversely, at lower optical energy densities and/or in low-absorbing (i.e., reflecting) materials, the interaction between beam and specimen is less violent. Under these *thermoelastic* conditions, the optical energy rapidly heats the material surface; this heating is followed by a rapid cooling as the energy is dispersed. The associated pulse of material expansion and contraction launches the ultrasonic wave.

In addition to the surface damage caused by ablative excitation, thermoelastic and ablative wave excitation differ in how much energy goes into what type of wave and the associated *far field* directivity pattern (see Section 3.3.4). Figure 3.43b shows the directivity patterns for shear and longitudinal waves for both thermoelastic and ablative laser-generated ultrasound. In addition, Figure 3.43c shows the response for each method as measured in through-transmission, with the receiver placed directly opposite the excitation point. This configuration is known as *epicenter detection*. Additionally, both methods produce a large, surface-bound wave.

Laser (Optical) Detection Methods

The laser generation and detection mechanisms operate on different principles. Occasionally, laser-generated waves are detected with conventional EMAT or piezoelectric transducers. More often, however, the ultrasonic waves are measured through optical means. There are a variety of laser detection methods; two methods will be presented—the knife-edge and the Michelson interferometer.

The *knife-edge method* monitors the changes in angular deflection of a laser beam reflecting off the specimen surface as the acoustic wave displaces the surface (Figure 3.44a). A knife-edge is placed in the path of the reflected beam, such that it partially obstructs the light falling onto the photodetector. As the surface moves in the presence of the acoustic disturbance, the angle of reflection

*There are other specialized generation mechanisms in certain materials, such as in silicon where the ultrasound is due to a deformation potential.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

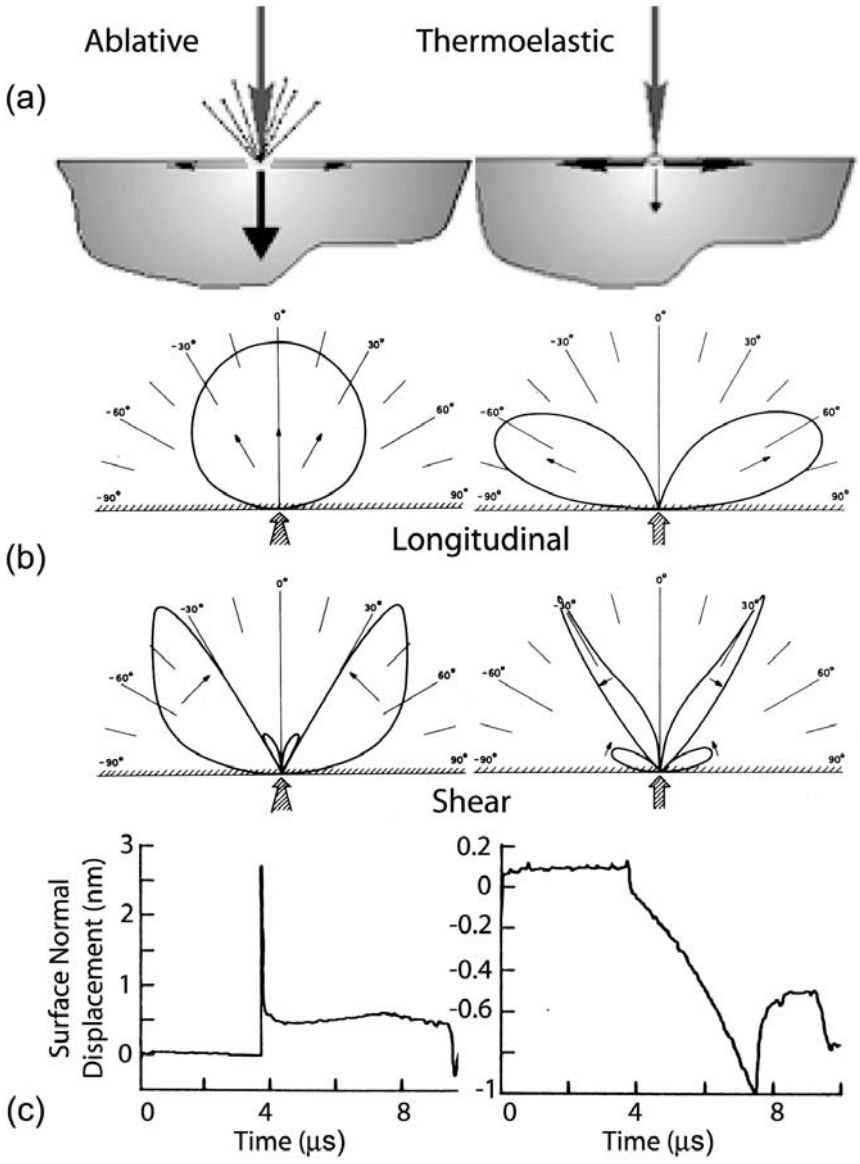


FIGURE 3.43 Laser generated ultrasound in the thermoelastic and ablative regimes. (a) Magnitude of force perpendicular and parallel to the specimen surface, (b) the directivity patterns for the shear and longitudinal waves, and (c) response signal. This signal is measured on the opposite side of specimen directly opposing to the laser strike (epicenter). (Fig. 3.43b CB Scruby, LE Drain. Laser Ultrasonics Techniques and Applications. Bristol: Adam Hilger, 1990, pp 263, 269. Reprinted by permission.) (Fig. 3.43c ADW McKie. Applications of Laser Generated Ultrasound using an Interferometric Sensor. PhD dissertation, University of Hull, England, 1987, p 61. Reprinted by permission.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

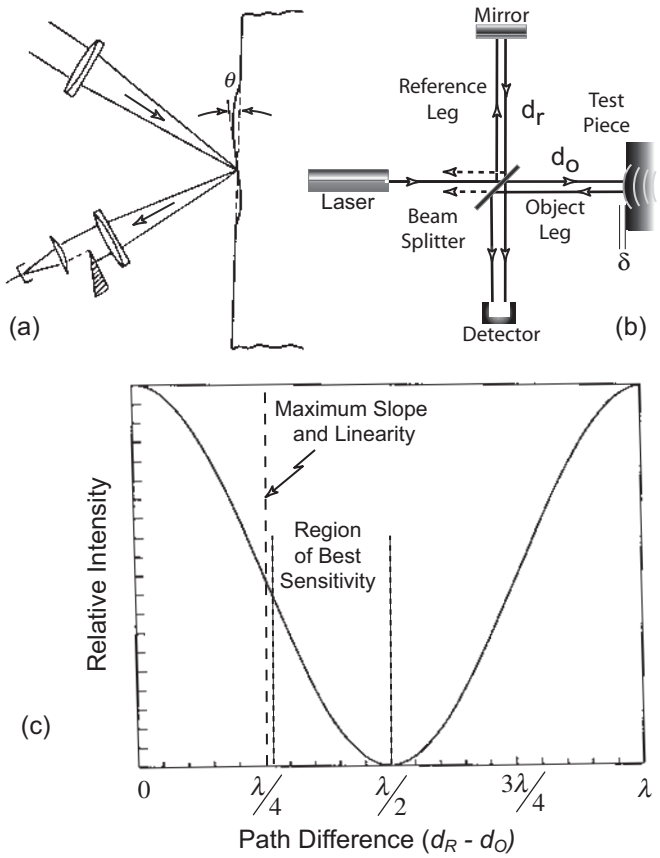


FIGURE 3.44 (a) Schematic of laser beam knife-edge detection of ultrasound, (b) basic Michelson interferometer setup for detection of ultrasound, (c) detector output of an interferometer as a function of displacement.

changes, and more or less of the light beam is obstructed by the knife-edge block. Razor blades are commonly used as the knife edge.

Even more common than the knife-edge method, the *Michelson interferometry method* makes use of the high sensitivity of optical interference (Figure 3.44b). A laser beam strikes a partially silvered mirror (*beam splitter*) and is split into two separate beams, one reflected and one transmitted. One beam reflects off a stationary mirror (the *reference leg*) and the other beam reflects off the test sample (the *object leg*). After they reflect, the two beams are recombined and sensed by a photo detector.

Copyright © Marcel Dekker, Inc. All rights reserved.

Because the two beams originated from the same beam, they are coherent and will constructively and destructively interfere at the detector. The degree of constructive or destructive interference depends on the path-length difference between the reference and object legs of the interferometer. If the lengths are exactly equal (or the difference is an exact multiple of the wavelength), then the beams will add to produce a bright spot. Conversely, if the path length difference is a multiple of a $\frac{1}{2}$ wavelength, then the beams subtract, and produce a dark spot at the detector. As the path-length difference is varied over a wavelength, a sinusoidal pattern of bright-to-dark tracks the movement. Michelson interferometers employing visible helium-neon lasers with a wavelength of 650 nm readily detect the nanometer length surface deflections of an acoustic wave. (Normally, infrared filters are used to block the high-energy light from the laser generation pulse that would otherwise saturate the interferometer's photo detector.)

Both the knife-edge method and the Michelson interferometry method are incredibly sensitive to small table vibrations, the sample roughness, or even temperature gradients in the air. Methods to stabilize path difference caused by environmental influences as well as techniques to improve sensitivity can be found in Refs. 38 and 39. Additionally, these methods require the surface to be partially reflective.

Optical Interference (optional)

Most ultrasonic optical detection methods are based on the interference of two optical waves (beams): a reference beam and an object beam whose path length is modulated by the acoustic wave. Let the reference and object beam be represented by

$$A_R = a_R e^{j(\omega t - k d_R)} \tag{3.101}$$

$$A_O = a_O e^{j(\omega t - k(d_O - 2\delta))} \tag{3.102}$$

where a_O and a_R are the amplitude coefficients, d_O and d_R are the path lengths of the object and reference legs, δ is the acoustic wave–induced path-length variation in the object leg, and k is the wave number (spatial frequency). These two beams combine and are detected by the photo sensor, which senses the *intensity* of the incident light. The intensity of the combined beams is the *squared sum* of the amplitudes. (This is in contrast to two UT waves interfering where the amplitudes simply sum [Section 3.6.1].)

$$\begin{aligned} I_D &= |A_O + A_R|^2 = (A_O + A_R)(A_O^* + A_R^*) \\ &= a_O^2 + a_R^2 + 2a_O a_R \cos(k(d_R - d_O) + 2k\delta) \\ &= \left(a_O^2 + a_R^2\right)\left\{1 + 2\frac{a_R a_O}{a_O^2 + a_R^2}\cos(k(d_R - d_O) + 2k\delta)\right\} \end{aligned} \tag{3.103}$$

where * indicates the complex conjugate.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

In the case of optical detection of ultrasonic waves, the surface displacements are small compared to the wavelength, i.e., $\delta \ll \lambda$; thus the interferometric intensity equation can be simplified:

$$I_D = (a_O^2 + a_R^2)\left\{1 + 2\frac{a_R a_O}{a_O^2 + a_R^2}(\cos k(d_R - d_O)\cos 2k\delta - \sin k(d_R - d_O)\sin 2k\delta)\right\}$$

$$I_D \approx (a_O^2 + a_R^2)\left\{1 + 2\frac{a_R a_O}{a_O^2 + a_R^2}\cos k(d_R - d_O) - 2\frac{a_R a_O}{a_O^2 + a_R^2}\sin k(d_R - d_O)2k\delta\right\} \quad (3.104)$$

The first two terms within the braces represent constant background-light intensities incident on the detector. The third term modulates between bright and dark as the surface is displaced by the acoustic wave. The *modulation depth*, $2(a_R a_O/a_O^2 + a_R^2)$, determines how bright and how dark the light intensity varies. If $a_O = a_R$, then total destructive and constructive interference occurs, i.e., I_D ranges from 0–$4a_O^2$. Figure 3.44c shows the interferometer response as δ, the path difference, varies. This sinusoidal patterns shows that the sensitivity of the interferometer varies depending on where one placed the bias (neutral–no displacement $-\ \delta = 0$) point. Although, it may appear that the maximum sensitivity would be at the point of maximum slope, it turns out that the best sensitivity occurs between $\frac{1}{4}$ and $\frac{1}{2}\lambda$.

3.3.4 Transducer Characteristics

Consider the finite (albeit amazing) abilities of your own sound transducer, your ear:

- It has a nominal frequency response that allows detection between 20 hz–17 kHz—an astounding factor of $\sim 10^3$.
- Its response to all the frequencies within this range is approximately equal.
- It damps out one sound quickly enough to detect the next sound without interference or ringing of the previous sound.
- It can detect amplitudes from a soft whisper in a soundproof room (1 dB*) to the sound of a jet engine (120 dB)—12 orders of magnitude!

*0 dB is defined as the nominal threshold of human hearing.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Ultrasonic transducers have the same characteristics as your ear: finite frequency bandwidth, amplitude response, directivity pattern (diffraction), damping, and lower (threshold) and upper (saturation) intensity limits. And, as with your ear, the finite size of ultrasonic transducers (aperture size) implies that a true plane wave can neither be generated nor detected.

Frequency and Amplitude Response

Consider the transducers for your stereo—the speaker system. To accommodate the broad frequency range of music, a typical speaker system employs three transducers: a woofer for low frequencies, midrange for midrange frequencies, and a tweeter for high frequencies. Each of these transducers has a different frequency range within which the amplitude is relatively constant—the *frequency and/or amplitude response*.

Like stereo speakers, ultrasonic transducers have a frequency-dependent efficiency or amplitude response. Figure 3.45 shows the frequency response for two different piezoelectric ultrasonic sensors. To describe a transducer frequency

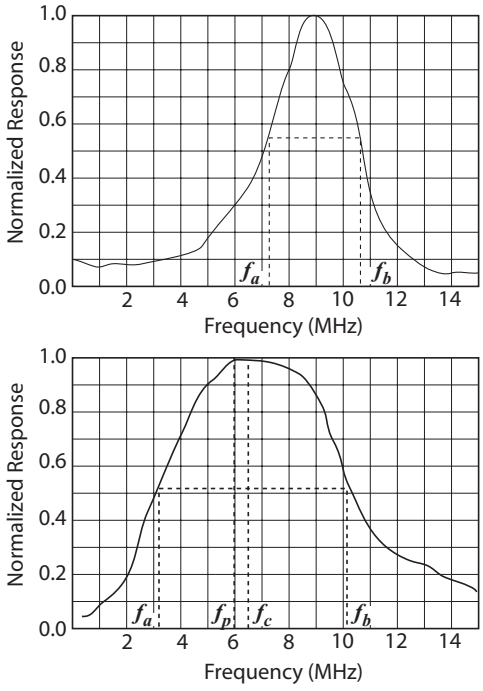


FIGURE 3.45 Frequency response (bandwidth) of different commercial piezoelectric transducers.

Copyright © Marcel Dekker, Inc. All rights reserved.

response, transducer manufacturers specify the bandwidth, the center frequency, and a skew frequency. Note that these are nominal values, not the actual values of a particular transducer.

The transducer *bandwidth* is defined as

$$BW = f_b - f_a \tag{3.105}$$

f_a and f_b represent the points where the transducer power efficiency has dropped to 50 % or −3 dB of its peak value (Figure 3.45). The *center frequency* is defined as

$$f_c = \frac{1}{2}(f_a + f_b) \tag{3.106}$$

Most transducer frequency response spectrums are not symmetric about the center frequency. Therefore, a *skew frequency* is defined as:

$$f_{\text{skew}} = \frac{f_{\text{pk}} - f_a}{f_b - f_{\text{pk}}} \tag{3.107}$$

where f_{pk} is the frequency at peak efficiency. If the system is perfectly symmetric, then

$$f_{\text{skew}} = 1$$

If a transducer responds only to a narrow range of frequencies, then its electronics must be tuned to produce or receive a similarly narrow bandwidth of frequencies. Typically, a *narrowband transducer* is excited by applying a *gated tone burst*—10 or more cycles of a single frequency.

With a *broadband transducer*, on the other hand, its broad spectrum of frequencies allows for excitation or reception of an ultrasonic pulse or spike. (Remember that a spike comprises a large range of frequencies, and thus would excite a broad range of frequencies in the system.)

Overall, pulsed laser generates an extremely broadband ultrasonic wave, and PVDF, a flexible, polymer transducer material, has a very broadband frequency response. In contrast, ceramic transducers have relatively narrow bandwidths. Manufacturers of ceramic transducers still define narrow, medium, and broad bandwidths. For the purposes of ultrasonic NDE, the optimal frequency range and bandwidth depend entirely on the system you're investigating.

In most UT applications, we simply specify the transducer frequency response to match the needs of the application. On occasion, the specific amplitude response as a function of frequency must be known. For example, if we want to know the response of a material to guided waves of different frequencies, ideally we would excite the material with a short-duration constant frequency wave (gated tone bursts) of constant amplitude and then vary the frequency. Because no UT excitation transducer has a constant (*flat*) response with frequency, we can compensate for amplitude variations at different frequencies by normalizing the data to the amplitude response of the specific transducer.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Beam Characterization

How many times have you heard voices from around the outside corner of a building or, if you are a hunter, a rifle shot from the other side of a hill? In fact, there is no straight-line path (or path for reflections) between the source and the sensor (your ear); the sound wave truly bends around the corner or over the hill! So far, we know of two ways in which a traveling wave can change direction: reflection and refraction. But neither of these terms can explain the bending sound waves. Therefore, we define a new term, *diffraction*, as the bending of waves at the edge of an opening or obstruction. Diffraction has two physical characteristics: (a) the spreading (divergence) of the beam as it encounters the edge of an obstacle, and (b) the subsequent intensity variations of the wave in this expanding wave field.

Simple wave *interference* explains both of these diffraction characteristics. (When two or more UT waves of the same frequency cross paths [interfere] their amplitudes will add [*constructive interference*] or subtract [*destructive interference*] depending on the phase difference between them; see Section 3.6.1.) Imagine a *particle* stream obstructed by an aperture. After passing through the aperture, the particle flow retains the aperture dimensions and has a clear shadow region (Figure 3.46a). In contrast, when a wave encounters such an aperture, the waves spread out into the shadow region. The amount of expansion into the shadow region depends on the ratio of the wavelength to the width of the aperture (Figure 3.46b). Figure 3.47 shows this diffraction phenomenon for an actual ultrasonic wave from a circular aperture, where $\lambda/d = 6.7$. The light and dark indicate high and low acoustic pressure, respectively.

Experiment (Diffraction): Equipment: two razor blades, pin, a laser pointer, a sheet of aluminum foil 3 cm by 3 cm, a sheet of white paper, and a darkened room.

Create a wedge with the two razor blades, such that one end of the wedge is approximately 1 mm in width and the other end narrows to zero width (overlap the blades a little). Using the white paper as a screen about 2 m from the aperture, shine the laser pointer through the wedge aperture, starting at the widest point. Slowly move the beam along the aperture towards the aperture apex. Note that the shadows of the razor blade edges are distinct and well defined when the aperture is large. As the aperture narrows, the shadow edges as seen on the paper become undefined in space; the light beam broadens, and intensity variations appear within the light wave field.

Using the pin, prick a hole (aperture) as small as possible in the aluminum foil and shine the laser pointer through the hole toward the white paper screen. Examine the pattern of the light that passes through the hole. You should be able to see a pattern of circles (like a target) on the screen. What happens to the pattern if you vary the distance between the aperture and the screen?

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

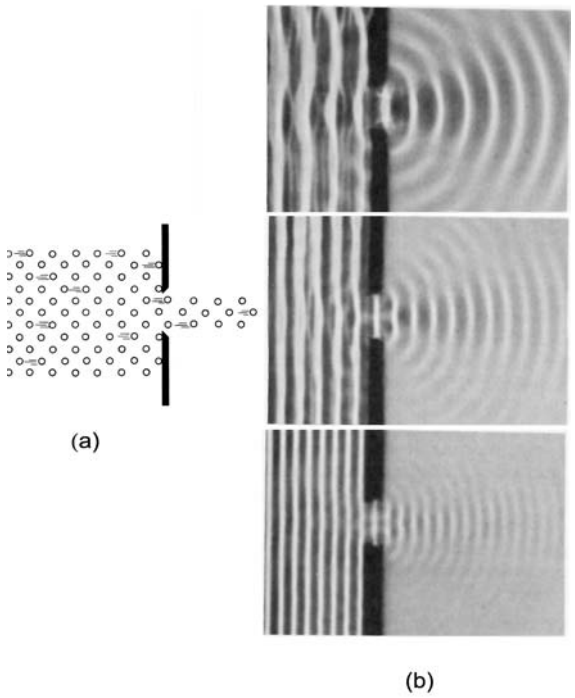


FIGURE 3.46 The contrast between (a) a particle stream versus (b) waves response to encountering an aperture. Ripple tank demonstration of diffraction dependence on the wavelength-to-aperture width ratio. Here the aperture width is fixed. (Adapted from Fundamental University Physics by M. Alonso and E.J. Finn p 901, © 1967 and Optics by E. Hecht p 393, Copyright © 1997, 1974 by Addison-Wesley Publishing Co. Inc. Reprinted by permission of Pearson Education, Inc.)

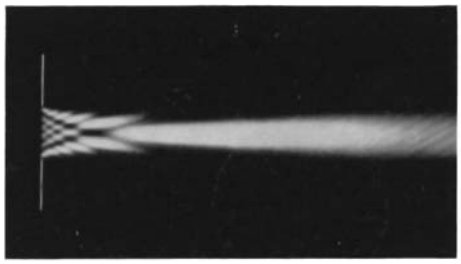


FIGURE 3.47 Sound field diffraction pattern of an actual circular transducer with $D/\lambda = 6.7$. Light indicated high pressure. (J Krautkramer, H Krautkramer. Ultrasonic Testing of Materials. 4th ed. Berlin: Springer-Verlag, 1990, p 59. Reprinted by permission Springer-Verlag.)

Copyright © Marcel Dekker, Inc. All rights reserved.

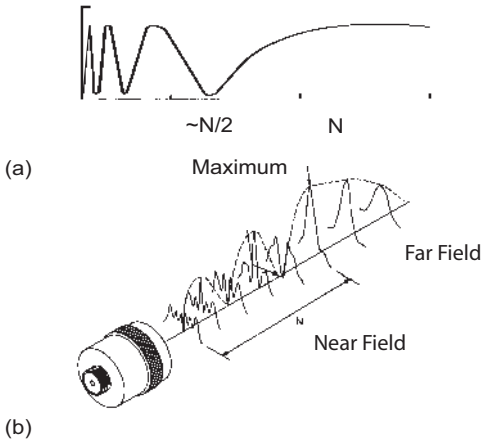


FIGURE 3.48 (a) Plot of sound field variation along the central axis as a function of distance from the transducer. (b) Schematic of the sound field. (Courtesy Panametrics.)

Near Field and Far Field. Consider an ultrasonic beam from a piston-shaped piezoelectric transducer. You might expect that if you graphed the intensity of the beam in the center of the beam as a function of distance, then the intensity would slowly decrease. In fact, however, the real graph of this phenomenon is quite different (Figure 3.48a). At long distances (*far field*), the intensity changes as expected, but at short distances (*near field*) the intensity fluctuates wildly. (Note: This description of oscillation in the near field assumes a single frequency or narrow-band excitation. For broadband excitation, the multiple frequencies smooth the variations in the near field.)

To see why this modulation in intensity from a finite aperture occurs, we can apply the Huygens-Fresnel wavelet principle. Huygens modeled the waves in the aperture by a row of point sources (Figure 3.49). Each source expands forward spherically. Huygens stated that the sum of these waves would approximate the original wave; Fresnel later added the concept of constructive and destructive interference to Huygens's model.* As shown,

*Spherical wavelets are used to present Huygens' wavelet model (1629–1695). We understand, however, that no wave propagates backwards towards the original ultrasonic wave source. Remember, Huygens' principle is a conceptual model, not a rigorous one. Additionally, Huygens' principle did not account for the intensity variations fundamental to diffraction. Fresnel combined Huygens' wavelet principle with the idea of interference, resulting in the Huygens-Fresnel principle. Ironically for our discussion of ultrasonic wave diffraction in elastic medium, Fresnel was describing the diffraction of light as it traveled through the ether—an *elastic* medium.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

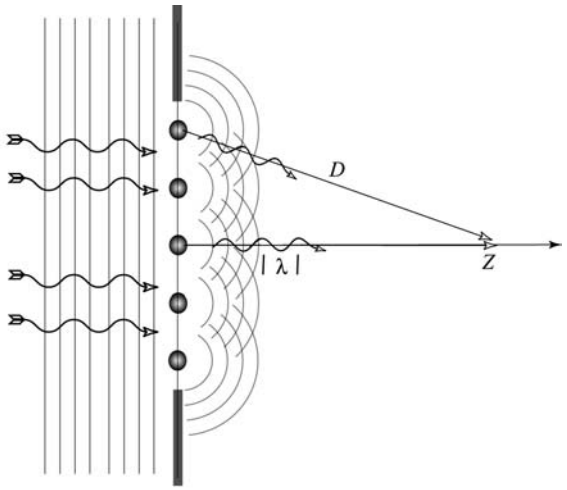


FIGURE 3.49 Wave diffraction from an aperture as explained by Huygens-Fresnel wavelet principle.

the wave at point Z along the central axis is the sum of all the wavelets expanding through this point. Depending on the path lengths, these wavelets can add constructively (path-length difference is a multiple of λ) or destructively (path length difference is a multiple of $\lambda/2$). In other words, intensity fluctuates with distance. But the path-length difference becomes smaller with increasing distance from the aperture. After a certain distance, the "beams" have less than a $\lambda/2$ path-length difference and can no longer combine destructively. The demarcation between these two regions separates the near field (characterized by high intensity fluctuations) from the far field (characterized by a constant decay in acoustic pressure with distance). Figure 3.48b clearly shows these two regions.

It is easy to derive the distance of the boundary between the near and the far fields (*near field length*, N). The near field length should occur where the *maximum* path length between beams is $\lambda/2$, or:

$$\begin{aligned} D - N &= \lambda/2 \\ \sqrt{N^2 + (d/2)^2} - N &= \lambda/2 \end{aligned} \tag{3.108}$$

where d is the diameter of the transducer, D is the maximum distance to the point from the transducer and λ is the wavelength (Figure 3.49). Rearranging

Copyright © Marcel Dekker, Inc. All rights reserved.

the Eq. (3.108) leads to:

$$\begin{aligned} N^2 + \frac{d^2}{4} &= \frac{\lambda^2}{4} + \lambda N + N^2 \\ \lambda N &= \frac{d^2}{4} - \frac{\lambda^2}{4} \\ N &= \frac{d^2 - \lambda^2}{4\lambda} \end{aligned} \tag{3.109}$$

Since the wavelength is usually much smaller then the diameter of the transducer, sometimes the near field length is taken to be approximately

$$N \approx d^2/4\lambda \tag{3.110}$$

The distance $N/2$ approximates the location of the last minimum. In general, most experiments are conducted in the far field regime as the intensity fluctuations in the near field can cause an inspector to miss, or misread, flaws in the sample.

Beam Divergence (circular aperture transducer). Huygens-Fresnel wavelets also predict the expansion of the acoustic wave into the shadow regions, e.g., the wave bending around a corner. Obviously, each point source wavelet expands laterally, as well as normal to the aperture face (Figure 3.49). These laterally expanding "beams" interfere and create regions of fluctuating intensity. If the aperture were infinitely large, all of the laterally traveling waves would cancel each other and only a forward traveling wave would exist—obviously, since an infinitely large aperture is no aperture at all! At the other extreme, if the aperture is much smaller than the wavelength, then only one wavelet would represent the wave in the aperture, and would expand hemispherically into the downstream side of the aperture.

Because most UT measurements are made in the far field, let us examine a cross section of the far-field acoustic intensity pattern of a circular aperture transducer at some distance Z perpendicular to the aperture face. This intensity is expressed in two basic formats called *directivity patterns* (Figure 3.50). As expected, the highest energy occurs along the central axis—the *main beam* or *central lobe*. In addition, a series of *side lobes* decrease in intensity as Φ increases (Φ is measured from the central axis). In a circular aperture these side lobes create a circular "target." These beam patterns are often displayed in polar diagrams, where the intensity is graphed versus the angle for a set radius. Note that the lines in the diagram do not represent the position of the wave, but the *intensity* of the wave at that angle and radius.

Note also that in the far field, the main beam has a much higher intensity then the outer rings (side lobes). Therefore, in the far field, the ultrasonic wave

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

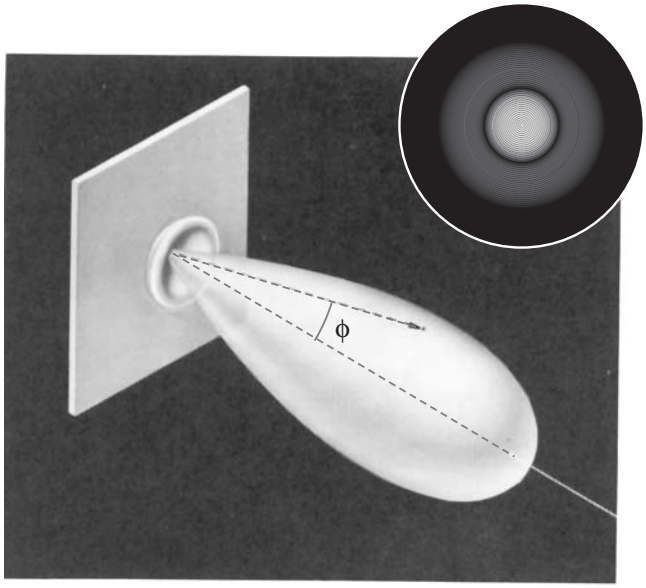


FIGURE 3.50 Directivity pattern of the far field acoustic intensity at a distance z from the aperture: polar plot and intensity ''bull's eye'' plot (inset). Both plots represent the same acoustic field. The polar plot also depicts the angle of divergence of the main beam. (Adapted from FW Sears, MW Zemansky. University Physics. 3rd ed. Reading, MA: Addison-Wesley, 1964, p 518. Reprinted by permission of Pearson Education, Inc.)

acts as a single beam which spreads out slowly with distance, much like a flashlight. The angle of divergence of the main beam is given by:

$$\sin \Phi_0 = \frac{1.22\lambda}{d} \tag{3.111}$$

where d is the diameter of the transducer, λ is the wavelength, and Φ_0 is the half angle of the beam when the intensity is -20 dB of the maximum intensity (Figure 3.50).

In summary, if the wavelength increases or the aperture diameter decreases, the divergence of the beam will increase. Additionally, because the wavelength equals the velocity divided by the frequency, if the frequency is increased, then the beam divergence will decrease. And if the material has a higher velocity, it should also have higher beam divergence.

Example: Assume a 1 MHz longitudinal wave is generated by a 20 mm–diameter transducer. Compare the beam divergence for the wave propagating in water to its propagation in aluminium. Given:

$$f = 1 \times 10^6 \text{ Hz} \qquad d = 20 \text{ mm} \qquad v_{\text{al}} = 6320 \text{ m/s} \qquad v_{\text{H}_2\text{O}} = 1483 \text{ m/s}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Solution:

$$\lambda_{al} = \frac{6320 \text{ m/s}}{1 \times 10^6 \text{ cycles/s}} = 6.320 \text{ mm}$$

$$\lambda_{H_2O} = \frac{1483 \text{ m/s}}{1 \times 10^6 \text{ cycles/s}} = 1.483 \text{ mm}$$

$$\phi_{al} = \sin^{-1}\left(1.22\frac{\lambda}{d}\right) = \sin^{-1}\left(1.22\frac{6.32 \text{ mm}}{20 \text{ mm}}\right) = 22.8^\circ$$

$$\phi_{H_2O} = \sin^{-1}\left(1.22\frac{1.483 \text{ mm}}{20 \text{ mm}}\right) = \sin^{-1} 0.091 = 0.091 \text{ radians} = 5.2^\circ$$

3.4 INSPECTION PRINCIPLES

To perform an ultrasonic inspection, you need to select:

The variable to be measured
The appropriate transducer for generating a specific wave type (longitudinal, shear, surface, or focused)
The transducer configuration: pitch–catch or pulse–echo
The type of excitation pulse: spike, tone burst, or continuous wave (CW)
The type of signal processing and display method, e.g., time vs. amplitude or position on the part vs. attenuation, as well as whether a point measurement or an area scan will fit the application better

3.4.1 Measurements

Ultrasonic inspection directly measures four variables: time of flight, amplitude, phase, and frequency. All other information (such as attenuation, elastic constants, homogeneity, and structure) is derived from this set of variables.

Time of Flight

When an ultrasonic pulse is sent through an object and then reflected (or transmitted), the wave transit time can be measured, and is called the *time of flight*. Assuming that the sample thickness is known, the time of flight measurement can be used to determine the velocity:

$$v = \frac{\text{Total distance}}{\text{Time of flight}}$$

From this velocity data the material density and/or the elastic constants can be calculated. Conversely, if the velocity is known, then time of flight can be used to

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

determine the position of a flaw. It is important to note that if the signal is reflected, then the measured time of flight represents the *total* distance traveled by the wave—that is, twice the distance from the source/receiver to the object.

Amplitude

Attenuation (the change in amplitude due to absorption, beam spreading, scattering, or dispersion; see Section 3.2.5) is measured by observing the amplitude decay of a sequence of reflections. Recall from Eqs. (3.92) and (3.95):

$$\begin{aligned} \alpha l &= \ln\left(\frac{A_{\text{final}}}{A_{\text{initial}}}\right) \\ \alpha_{\text{db}} l &= 8.686\ \ln\left(\frac{A_{\text{final}}}{A_{\text{initial}}}\right) \end{aligned} \tag{3.112}$$

where α is the attenuation constant in Nepers, α_{db} is the attenuation constant in decibels, l is the distance traveled and A is the amplitude in volts. Although you could also measure the distance traveled by measuring the change in amplitude, time of flight is a much easier and more accurate measure. Instead, attenuation is used to measure the homogeneity, volume fraction of dispersed particles, or other absorptive or scattering qualities of a material? The data may also be used to interpret the density, or the elastic constant, or variations in either.

Phase

Occasionally, experimenters measure the phase of the wave. If a wave reflects off an object with higher impedance, then the phase will change by 180° (Reflection and Refraction at an Interface in Section 3.2.4). Phase measurement is generally done in conjunction with amplitude and time-of-flight measurements. In some cases, such as ultrasonic interferometry, the response signal phase can be measured against the phase of a reference signal.

Frequency

Measuring frequency can determine resonance. First, a continuous wave is placed in the system; the frequency is then adjusted until the system creates a standing wave. At the highest frequency the system is harmonic when the wavelength λ is equal to twice the length d, or:

$$\begin{aligned} \lambda_1 &= 2d = \frac{v}{f} \\ f_1 &= \frac{v}{2d} \end{aligned} \tag{3.113}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

For each corresponding frequency, the wavelength is:

$$\lambda_n = \frac{2d}{n} = \frac{v}{f}$$
$$f_n = \frac{v}{2d} n \tag{3.114}$$

Therefore, for two neighboring frequencies:

$$f_{n+1} - f_n = \frac{v}{2d} \tag{3.115}$$

this equation can measure velocity or thickness. If there is a defect in the sample, the resonant frequencies adjust accordingly.

3.4.2 Wave Generation

As you know, different transducers create different types of waves—longitudinal waves, shear waves, surface acoustic waves, and focused waves. In practice, there are many different types of transducers that will produce a specific type of wave. Wedges and other additions also create flexibility for generating specific wave types.

Generating Longitudinal Waves

There are four major methods for generating longitudinal waves in samples:

1. A piezoelectric transducer
2. A phased array of piezoelectric transducers
3. An EMAT, and
4. Laser generation

Of these methods, by far the most common is the first, which is performed either by coupling a piezoelectric transducer to a sample with a thin layer of coupling gel, or by placing both the transducer and the sample in coupling water bath. Either way, this method creates a normal longitudinal wave in the sample. To create an angled wave, the transducer can be attached to a wedge (commonly made of plastic), which is then positioned on the sample. Effective wave generation requires a thin layer of couplant at all interfaces—transducer/wedge and wedge/sample. The angle of the wedge dictates the incident angle, and Snell's law determines the refracted wave angle. Snell's law also shows, however, that a mode-converted shear wave as well as a longitudinal wave is generated in the sample (Figure 3.51). Because of the shear and longitudinal velocities, the reflections off a given boundary will have different arrival times, making signal interpretation difficult.

If only the angled longitudinal mode is desired, a phased array can be used instead. A phased array is a set of several normal-incidence piezoelectric

Copyright © Marcel Dekker, Inc. All rights reserved.

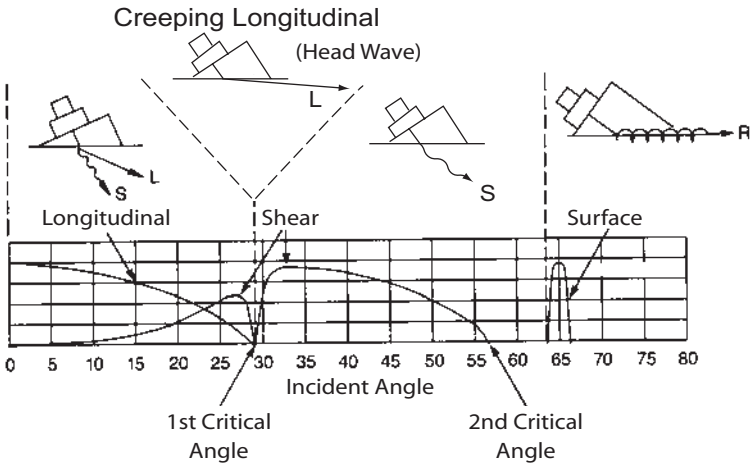


FIGURE 3.51 Using a longitudinal wedge-transducer, one can generate a longitudinal, transverse, and/or surface wave. As indicated, the wedge angle dictates the refracted wave types. (Courtesy Panametrics.)

transducers. Through either electrical or physical means, the transducers are driven at slightly different times, so that the resultant waves interfere constructively and destructively into an angled longitudinal wave (Figure 3.34).

EMAT and laser-generation methods (Sections 3.3.2 and 3.3.3) have the advantage of not requiring contact with the sample. EMAT standoff distance is on the order of millimeters, and laser generation commonly has standoff distances of meters. These methods, however, are complicated to run and can result in waves of much smaller amplitude.

Generating Shear Waves

Just as longitudinal waves can be created with piezoelectric longitudinal transducers, shear waves can be created with normal incidence piezoelectric shear transducers—and angled shear waves can be created with wedges. Unlike longitudinal transducers, however, shear transducers require special high-viscosity couplant and cannot be used in liquid scanning tanks. (Most liquids do not effectively support shear waves; all immersion transducers are longitudinal type.)

Normal-incidence piezoelectric shear transducers tend to be more expensive and create lower-intensity waves than longitudinal transducers. For this reason, shear waves are often created by using, instead, a longitudinal piezoelectric transducer on a wedge. If the angle of the wedge is chosen correctly

Copyright © Marcel Dekker, Inc. All rights reserved.

(between the first and second critical angles), then the incident longitudinal wave will have complete mode conversion into a shear wave (Figure 3.51). Wedges are specifically designed for this single-mode generation. EMAT and laser systems also generate shear waves.

Generating Surface Acoustic Waves

There are four methods of creating a surface acoustic wave (SAW). The simplest and most commonly used is the *wedge method*, in which a longitudinal wave is transmitted through a wedge coupled to a solid; this is the same method as that for generating refracted shear waves (Figure 3.52(a)). Due to Snell's law, the wave will bend depending on the characteristics of the wedge and the material. Therefore, according to Eq. (3.89) the angle should be:

$$\begin{aligned} \frac{1}{v_R}\sin 90^\circ &= \frac{1}{v_w}\sin(\theta_{\text{critical}}) \\ \theta_{\text{critical}} &= \sin^{-1}\left(\frac{v_w}{v_R}\right) \end{aligned} \tag{3.116}$$

where v_w is the speed in the wedge and v_R is the Rayleigh speed in the solid.

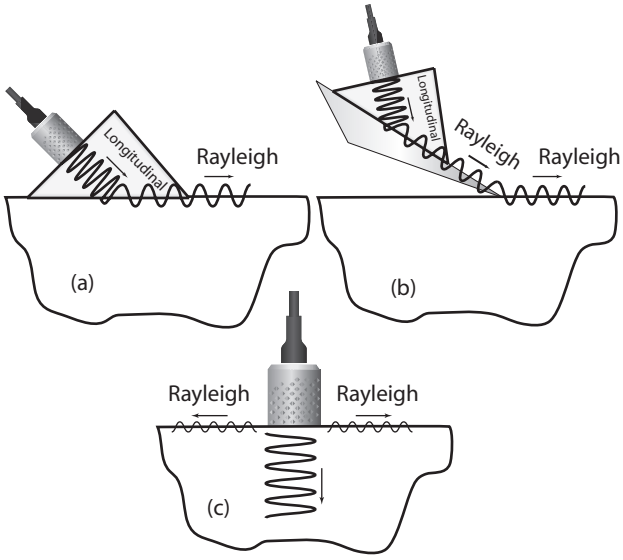


FIGURE 3.52 Various methods to generate SAW: (a) wedge: refracted mode converted, (b) mediator, and (c) normal incidence longitudinal wave.

Copyright © Marcel Dekker, Inc. All rights reserved.

Just as the velocity for an SAW is very close to the velocity for a shear wave, the critical angle for creating SAWs is very close to the second critical angle, or the angle where the shear wave is on the surface. Since the velocity of the SAW is less than that of both the shear wave and the longitudinal wave, both of these waves are totally internally reflected—i.e., not refracted into the solid. But if the Rayleigh velocity in the test medium is less than the longitudinal velocity in the wedge, then the refracted wave cannot be bent to the surface (90°) and no critical angle exists (see Oblique Incidence in Section 3.2.4). In this case, a *mediator* can be used. A mediator is a wedge of metal with a protruding tip and a velocity greater than the velocity in the wedge (Figure 3.52b). The wave in the wedge creates an SAW on the surface of the mediator. The tip of the mediator is then placed on the surface of the sample, and the wave is transmitted onto the new medium.

A second common method of creating SAWs is to create a longitudinal wave straight into a solid (Figure 3.52c). This method works like dropping a stone in a container of water, creating both bulk waves and surface waves.

Third, an SAW wave can be made with a phased array of longitudinal piezoelectric transducers. The advantage of this method is that only the surface wave propagates. By placing the comb's tang (the phased array elements) spacing at the Rayleigh wavelength, λ_R, the Raleigh waves add constructively while the bulk waves can add destructively.

Fourth, a shear wave transducer can be placed against the corner of a plate to create a SAW on the plate.

Generating Focused Waves

Just as with visible light (electromagnetic waves), acoustic waves can be focused or reflected with a lens and mirrors. There are many different approaches to focusing the acoustic wave. Shaping the transducer face, Fresnel zone plates, and liquid-filled lens-shaped containers all rely on refraction (Snell's law) of the wave in the same manner as optical lenses.

Also as in optics, focused beams increase sensitivity and resolution, however, at a cost of dynamic range and flexibility. For example, just as with a camera, the beam is focused only within its depth of field (Figure 3.53a,b). The smaller the spot size or beam waist (i.e., the sharper the focus), the shorter the depth of field; the spot size dictates the resolution. Also, according to Snell's law, as the focused beam passes from medium 1 to medium 2, the focal length shortens ($v_1 < v_2$) or lengthens ($v_1 > v_2$) according to the ratio of velocities of the two medium (Figure 3.53c).

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

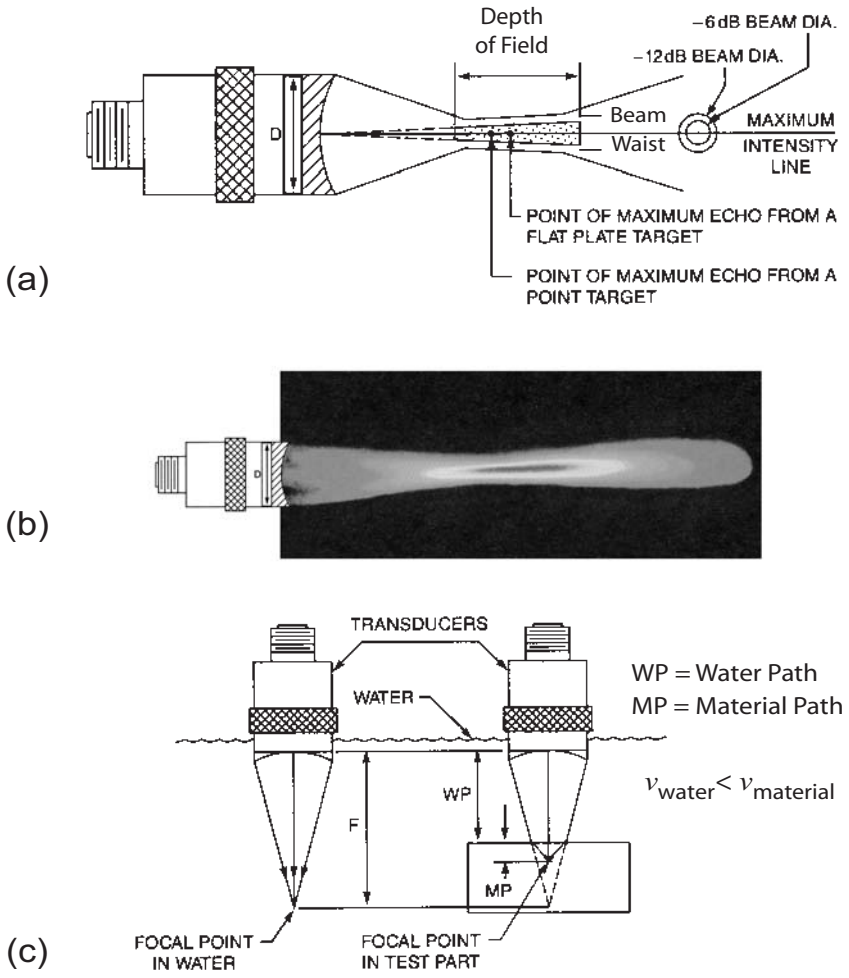


FIGURE 3.53 Focused beams: (a) characteristics of the beam (b) actual sound field intensity of a focused transducer, and (c) focal length reduction as the beam passes to a *higher* velocity medium.

3.4.3 Transducer Configurations

After choosing a transducer and the coupling medium, and deciding whether to use intermediate systems (wedges, lenses, or arrays), one next has to decide between the two basic transducer configurations for generating and receiving the ultrasonic wave. *Pulse–echo* incorporates a single transducer that acts as both the

Copyright © Marcel Dekker, Inc. All rights reserved.

receive and send transducer; *pitch–catch* employs separate generating and receiving transducers.

The dual function of the transducer for the pulse–echo method simplifies the setup and the electronics. Pulse–echo can also be very advantageous where access is limited to a single side (*single-sided access*), e.g., the skin of an aircraft. In Figure 3.54, a single transducer generates the ultrasound and then receives an echo from a flaw or the backwall. Additional echos would be observed as the ultrasonic wave continues to reflect off the front and back surfaces (as well as any flaw). With each round trip, the wave is attenuated, as shown by decreased signal amplitude.

The pulse–echo method has a couple of disadvantages. One occurs when the transducer is excited: the large electrical pulse necessary to generate the ultrasound is also seen by the receiver. As a consequence, the very sensitive receiver electronics are briefly saturated. This signal is called the *main bang*, and appears as the large (off-scale) initial "echo" in Figure 3.55. If the flaw is close to the surface, the main-bang signal will mask the flaw signal creating a *dead zone* in the received signal. Also, in a pulse–echo system with a near-surface flaw or for thin materials, the multiple returning echoes can overlap making accurate signal detection difficult (Figure 3.56)

The pitch–catch method uses separate generation and receiver transducers, effectively solving the problem of the main bang. Pitch–catch therefore has more potential to detect near-surface flaws and to operate on thin materials (recall the dual element transducer [Figure 3.33]). The send and receive transducers in

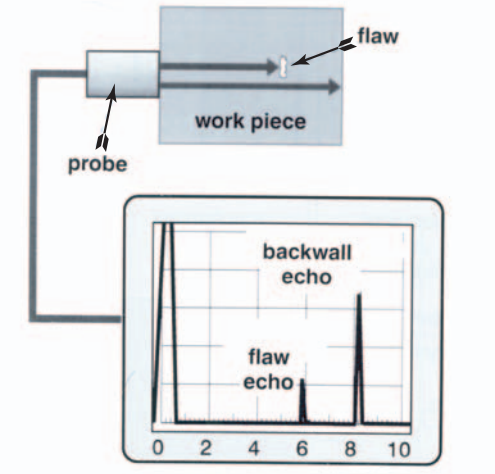


FIGURE 3.54 Pulse–echo flaw detection with echoes displayed. (Courtesy Krautkramer Branson.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

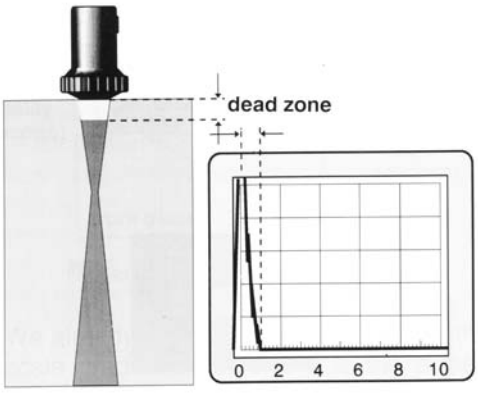


FIGURE 3.55 *Dead zone* in a pulse–echo system. (Courtesy Krautkramer Branson.)

pitch–catch can be placed on the same side of the sample (single-sided operation), or on opposite sides (through transmission). In same-side operation, a distance separates the transducers, and the signal is either an angled reflection off a flaw or the back wall, or the reception of a guided plate or surface wave. Through transmission works well on highly attenuating materials such as composites (including wood and concrete) and polymers, because the wave need traverse only one thickness of the material.

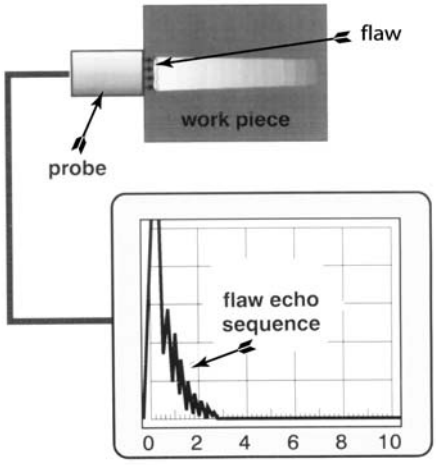


FIGURE 3.56 Overlapping echoes from multiple reflections off a near-surface (near-probe) flaw. (Courtesy Krautkramer Branson.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

But this method has disadvantages as well. If you are operating with normally incident waves and need single-sided access, pitch–catch will not work. Pitch–catch transducers also need to be aligned unlike pulse–echo, which is by definition self-aligning. First you need to know the directivity pattern of the excitation transducer (and to some degree that of the receiver) to position the receiver to capture the maximum signal—this may or may not be normal to the transmitter (Beam Characteristics in Section 3.3.4). Additionally, with shear waves the polarization (vibration direction) of two transducers needs to be coincident.

There are two basic methods for utilizing either pulse–echo or pitch–catch methods: *normal incidence* and *angled beam*. In normal incidence, ultrasound is injected into the specimen perpendicular to the surface (as described above). Angled beam refers to ultrasonic injection at off-normal angles. Figure 3.57 shows the angled-beam method for both pulse–echo and pitch–catch. Ultrasonic weld inspection commonly uses the angled-beam method.

3.4.4 Excitation Pulsers

Three common types of *drivers*—typically called *pulsers*, a term derived from the high-voltage pulsed output wave—provide an electrical pulse to the transducer to

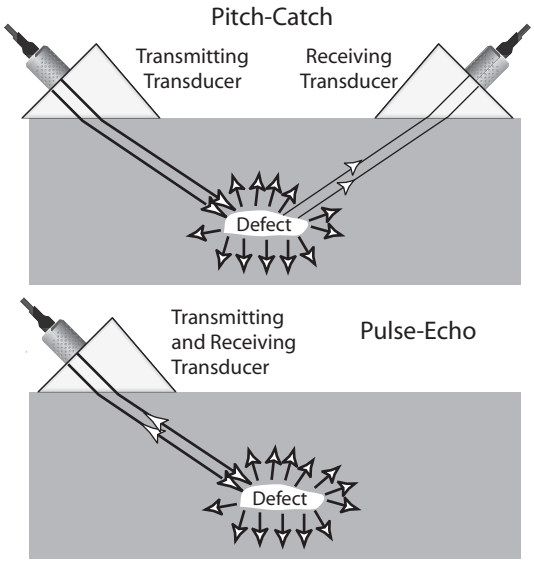


FIGURE 3.57 Angled beam transducers arranged in pitch–catch and pulse–echo.

Copyright © Marcel Dekker, Inc. All rights reserved.

generate an ultrasonic wave. These pulsers are described by the output-waveform type: spike, tone burst, or continuous wave (CW).

The *spike pulser* is the most commonly used general purpose pulser. Excitation with a spike pulser is akin to a hammer strike. The high voltage spike can be generated in a monopolar or bipolar single pulse wave. Typical voltage outputs are in the range of 200–400 V and can be repeated at frequencies (*pulse repetition frequency*) up to ~1 kHz.

Tone burst pulsers (*generators*) provide a constant amplitude sinusoidal pulse to the transducer for a set length of time (usually on the order of microseconds). The frequency of the pulse is set near or at the natural vibrational frequency of the transducer. One method of generating a tone burst is to gate a continuous sinusoidal wave, thus allowing only 3–10 cycles to excite the transducer. This type of pulser is useful if you want a narrow-frequency-band ultrasonic wave. Tone burst pulse repetition is typically less than 100 Hz.

Occasionally, a *continuous wave pulser* is used. A continuous wave pulser, as its name implies, creates a continuous wave, and is only used for resonant-frequency measurements. Sometimes a low-frequency tone burst pulser can be made into a continuous wave pulser by making the pulse width match the pulse repetition rate.

Most pulsers contain amplifiers and filters to adjust the output signal. Typical options with their limits are:

Pulse height (volts)

- Upper limit—can damage transducer or receiver electronics
- Lower limit—Low signal-to-noise

Pulse width (μs)

- Upper limit—ringing interferes with received signal
- Lower limit—difficult to see signal

Pulse repetition rate (ms)

- Upper limit—long time to get data plus many oscilloscopes cannot read
- Lower limit—second pulse overlaps echoes from first pulse

Frequency of pulse (Hz) (tone burst drivers only)

- High pass filter (Hz)—filters out signal with frequency lower then limit
- Low pass filter (Hz)—filters out signal with frequency higher then limit

Amplification (dB)

Copyright © Marcel Dekker, Inc. All rights reserved.

Many devices combine pulsers and receiver electronics in a common box, and offer a choice between pulse–echo or pitch–catch modes.

3.4.5 Receiving the Signal

Once the ultrasonic wave has been detected and converted to an electrical signal by the transducer, the *receiver* amplifies, filters, and performs other signal conditioning before displaying the results. Typical received signals are in the microvolt to millivolt range, with special application as high as several volts. These low-amplitude signals require fast, low-noise amplification. Additionally, the noise (grass) may require substantial filtering.

In general, the received signal is displayed on an oscilloscope screen, interpreted with a computer, and presented as an intensity measure; for dedicated equipment, a specific physical parameter, such as coating thickness, may be displayed. These displays are commonly integrated into a pulser-receiver system (Figure 3.58).



FIGURE 3.58 Typically oscilloscope display on an ultrasonic pulser-receiver system configured in pulse–echo mode. Notice the main bang. (Courtesy Krautkramer Branson.)

Copyright © Marcel Dekker, Inc. All rights reserved.

Oscilloscope-Displayed Signals

It is very common for the signal to be displayed on an oscilloscope as received signal amplitude (voltage) versus time. Generally, the oscilloscope is triggered by the pulser excitation signal. Many integrated systems display time of flight, attenuation, frequency, phase differences, and signal amplitude threshold levels directly on the screen, along with other signal processing features, such as filtering. Figure 3.58 is a typical oscilloscope reading from a transducer configured in pulse/echo mode: note both the initial ringing (main bang) and the reflection from the end of the sample.

Intensity Transformations (imaging)

Imaging means systematically scanning a specimen along a linear distance or an area and displaying the results. In an area scan, the probe is typically moved in a raster pattern over the interrogated area with data acquired at specified intervals. Although each data point in the scan initially is amplitude versus time, the data must be transformed so that it can be presented in a 2- or 3-D plot. The data from each point is analyzed and a value for some predetermined feature is extracted, such as maximum amplitude, decay due to absorption, or the time-of-flight between two specific echoes. These features are then plotted versus position to yield an ultrasonic image of the area scanned.

A-, B-, C-, and M-scans are the most common imaging methods used in ultrasonics. Actually, the A-scan is not really a scan. It simply is the result of a point measure. It is sometimes called a scan because it interrogates the material over some distance, e.g., through the thickness.

In a B-scan, the interrogated specimen is scanned along one axis. The resultant data reduction produces a cross-sectional image of the object. Position is plotted against an A-scan parameter, such as amplitude or time of flight. Most neonatal images are B-scans (Figure 3.59a). To avoid distortion of the image, care must be taken to align the transducer(s) to the specimen. The medical industry uses the term M-scan for a B-scan using time-of-flight data.

C-scans acquire ultrasonic data over an area, usually in a raster pattern. At each data point a parameter (or more than one) is extracted from the time-versus-amplitude signal; typical parameters are time-of-flight, amplitude, phase, or attenuation. Then the parameter is plotted versus position to build a bird's-eye view of the flaw or feature within the specimen, i.e., a slice of the object normal to the UT beam is imaged. The resultant plots present the scanned area in the x, y plane and the measured parameter along the z-axis (*waterfall plot*) or as intensity in grayscale or color, e.g., a bright spot representing a large reflection and dark a small reflection. (Color plots can often be deceptive, if not chosen as a continuous gradation along the color spectrum. People often refer to such data as presented in *false color*.) Figure 3.59b,c shows a C-scan of a Japanese coin. The C-scan system was set (gated) to detect the maximum amplitude between a preset time

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

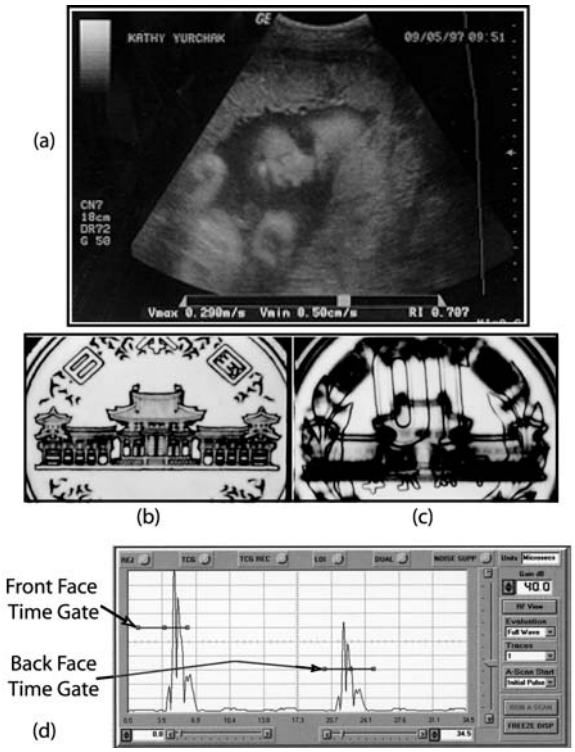


FIGURE 3.59 (a) Neonatal M-scan on Daniel Yurchak Shull. Ultrasonic C-scan of a Japanese coin: (b) gated on the front surface and (c) gated on the back surface. (d) Ultrasonic signal showing the response at the front and the back face and the associated time gates positioned to capture the signals within the gates.

interval. The time gates are shown in Figure 3.59d. For Figure 3.59b, we choose the time interval (time gate) to include any distance variations that might occur on the front surface of the coin. (This allows for the differences in transit time.) Note the high resolution. For the second scan (Figure 3.59c), the system was time gated to detect reflections from the back surface. Since the UT wave must first pass through the entire coin to reach the back surface, the front surface and any interior information will be superimposed onto the back surface information as shown in the figure.

Most NDE scanning couples the transducer via a fluid, typically water. Either the entire specimen and the probe are immersed in a scanning tank, or a special bubbler or water squirter is attached to a probe to facilitate high speed

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

scanning. The bubbler and squirter create a water-column path for the ultrasound from the transducer face to the specimen surface. The scanning-tank method produces less noise, provided the scan speed does not produce turbulence. The medical industry commonly uses a gel as a scan couplant.

3.5 APPLICATIONS

This section elaborates two case studies that illustrate the real-world use of UT principles (40–48).

3.5.1 Roll-by Inspection of Railroad Wheels

The railroad industry is a major user of NDE methods for both new and in-service inspection. Although in the United States rail and vehicle integrity are not as critical as in countries with high-speed passenger trains, a simple accident such as a rail or wheel failure can be very costly (Figure 3.60). Besides safety-critical factors, costs can include lost freight, inevitable disruption of commerce (tracks are often down for several days after an accident) and potential environmental hazards. Although both rail and wheel integrity are important and commonly inspected in the industry, this section focuses on wheels alone.

As massive as railroad wheels are ($\sim \frac{1}{3}$ton) and even with considerable design improvements, a significant number of accidents still occur from wheel failure. In a four-year period, failed wheels caused 134 accidents, resulting in



FIGURE 3.60 Possible result of a failed railroad wheel.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

losses totaling $27.5M. Railroad wheels are subject to very harsh operating conditions: large dynamic and static mechanical loads, extreme temperature fluctuations, and exposure to abrasive and corrosive environments. For example, in areas such as the Rocky Mountains, wheels are commonly drag-braked for an entire prolonged descent, reaching near-glowing temperatures; they are then rapidly quenched when the brakes are released.

This application details a roll-by inspection method for railroad wheels using EMAT-generated Rayleigh waves.

Design Considerations

For automated roll-by wheel-tread inspection, specially designed Rayleigh wave ultrasonic transducers are well suited. In deciding to use Rayleigh waves to inspect wheel treads, there are numerous considerations:

The fact that many tread flaws are nonsurface breaking flaws
Wheel radius
Wheel profile
Sensor liftoff
Wheel-sensor load constraints
Wheel-sensor contact time for roll-by inspection

Rayleigh waves are surface-bound waves. But they do penetrate into the surface, to a depth that depends on wavelength. For this application with $\lambda = 6\,\text{mm}$ and $f = 500\,\text{kHz}$, the penetration depth of 10 mm will detect any cracks that have been closed by the wheel-rail contact. Wheel radius, wheel profile, sensor liftoff, and wheel-sensor load constraints create signal-attenuation problems. As Rayleigh waves propagate along a surface, they may be attenuated or reflected by the surface curvature (Guided Waves in Section 3.2.6). The degree of reflection depends on the rate of curvature compared to the wave length. The large radius of the railroad-wheel tread presents little signal attenuation due to surface curvature. On the other hand, the relatively tight radius of the rounded corners on the wheel edges adequately confines the wave to the tread surface.

The wheel-sensor contact can also present significant problems. One system currently in use incorporates a pair of piezoelectric transducers arranged for pitch–catch. To couple the sensors to the wheel, a liquid-filled polyurethane boot covers the relatively fragile piezoelectric transducers. The boot is also sprayed with water, to aid the ultrasonic coupling. Although the system is well-engineered, the coupling unit is susceptible to lacerations that release the pressurized liquid under the load of the train. Small, extremely sharp slivers of metal are commonly abraded from the rail. To overcome this and other issues, a noncontacting EMAT system was developed.

The wheel-sensor contact time simply requires the wheel to remain in contact long enough for the ultrasound to traverse a minimum of one circumfer-

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

ence of the wheel tread. For a system operating at 500 kHz, a Rayleigh wavelength of 6 mm, and a wheel diameter of about 1 m, a round trip occurs in a little more than a millisecond. To improve sensitivity, the system should have at least two round trips. This limits the train velocity to about 20 mph.

Additional consideration has to be made to accommodate the sensor mounting. For effective and efficient inspection, the EMAT has to be mounted directly into the rail. (The mounting involves cutting away a large portion of the rail. Although you might think that the type of train and its load would be factors in the rail integrity, in fact the locomotive engine is always the heaviest load by far on the rail.) A partial cutout of a single rail was used to house the EMAT sensor. To align the wheel to the sensor a second rail, parallel to the load bearing rail, forced the train wheel against the load bearing rail. Wheels–track profiles allow the load to be carried by the center to the inside of the rail (Figure 3.61a).

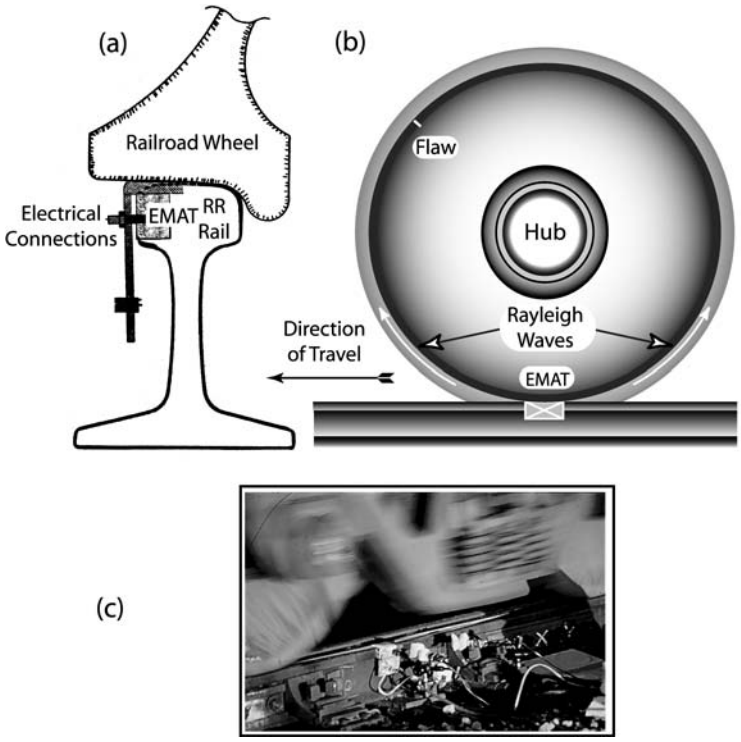


FIGURE 3.61 (a) Rail and wheel mating profile and location of EMAT assembly, (b) schematic of an EMAT inspection system installed in the rail for a roll-by inspection, (c) actual EMAT assembly.

Copyright © Marcel Dekker, Inc. All rights reserved.

Therefore, at slow speeds (under 20 mph), partial reduction of the exterior rail surface poses no great risk.

Inspection Principles

Figure 3.61b,c present the roll-by wheel-inspection setup. The transmitter and receiver Rayleigh-wave transducers are rail-mounted. As a wheel rolls over the sensor, a proximity system (a pressure switch) detects its presence and triggers the transmitter EMAT. A Rayleigh wave is launched into the wheel. The geometry and generation mechanism of the EMAT causes waves to be excited in both the clockwise and counterclockwise directions around the wheel. Thus a flaw can be detected by a reflection of both of the waves. As waves complete round trips or reflected waves pass the receiver transducer, the amplitude versus time is recorded.

The response signal is then processed to determine if the wheel is to be flagged for further inspection. We analyze signals in two ways: by the amplitude, and by the area under the curve of the reflections. Depending on the method chosen, either the reflection amplitude is normalized to the previous round-trip amplitude and then compared to a predetermined threshold value; or, similarly, the area is normalized and compared.

Results

To benchmark system sensitivity, we often use a model system with well-characterized flaws. In this case, we selected two new wheels and one wheel retired from service for thermal checking cracks. We machined a slot 2 mm wide and 10 mm long (the length is perpendicular to the wheel circumference) into the tread one of the new wheels. Figure 3.62 shows the response of these three wheels in a static test. In the graphs, two complete round trips appear as large spikes about 1 ms apart. The initial spike is the response from energizing the transmitter EMAT. Note the low attenuation of the wave between the first and second round trip. Although not shown, we measured 10 round trips (~31 m) before the signal-to-noise ratio reached unity.

Figure 3.62b clearly shows the echos from the slot. We notice, however, that one spike appears within the first round trip and two reflections appear within the next round trip. Why does this happen? As the wave traveling clockwise first reflects off the flaw, it produces the spike within the first 1 ms. But the reflection from the wave traveling counterclockwise must return over a distance greater than the total round-trip distance. Consider the distances traveled by the reflections of the two waves if the flaw were located very near the transducers. As the flaw location nears the transducers, the response signal will become masked by the main bang or subsequent round trip signals. It is left as a problem to estimate the length of these *dead zones* as well as their location. Additionally, the problem asks the student to present a solution to this problem.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

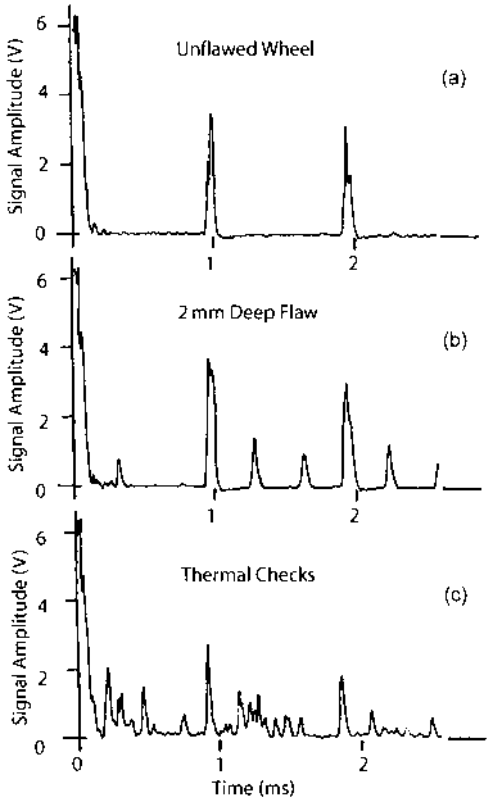


FIGURE 3.62 Ultrasonic response signals under static testing conditions on model wheels/flaws: (a) new unflawed wheel, (b) new wheel with a 10 mm long saw cut, and (c) a retired from service wheel with thermal checking cracks.

Figure 3.62c shows the response to a wheel that deteriorated when subjected to large temperature variations. Under such conditions, cracks with patterns similar to crazing in ceramic coatings appear on the wheel tread. Therefore we do not expect to see a "clean" echo pattern. Instead, the response resembles large amplitude noise. From careful examination of the signal, however, we note that the signal-to-noise ratio for Figure 3.62c is approximately equal to that of the other two response signals.

Whereas Figure 3.62 represents *model* results, Figure 3.63 presents representative responses from three wheels *actually* inspected in a roll-by system. Figure 3.63a depicts the signature from an in-service wheel with no flaws. The wheel represented in Figure 3.63b was flagged for a crack-like flaw.

Copyright © Marcel Dekker, Inc. All rights reserved.

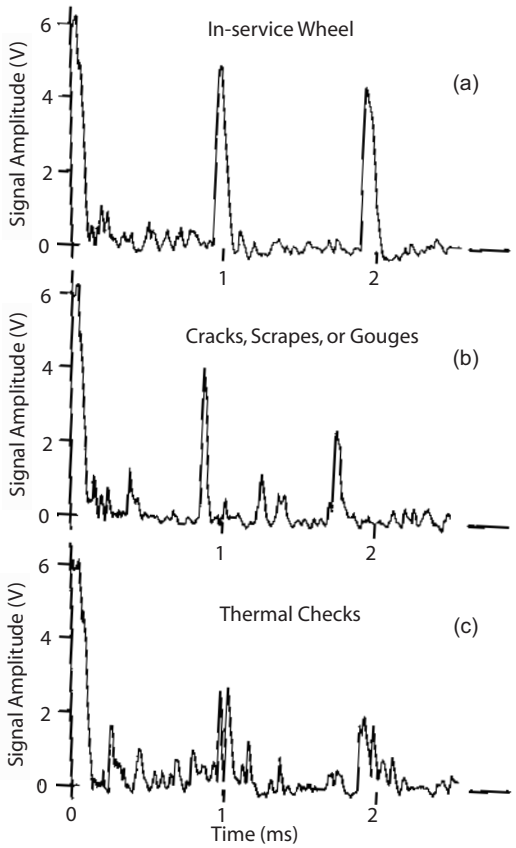


FIGURE 3.63 Ultrasonic response signals under actual roll-by test conditions with a train velocity of 12 mph: (a) in-service wheel with no flaws, (b) wheel flagged for a crack-like flaw, and (c) wheel flagged for thermal checking cracks.

Likewise the wheel represented by Figure 3.63c was flagged for significant thermal checking. Note the increase in baseline noise for an actual roll-by inspection as compared to the static tests. This increase in noise clearly makes flaw detection more difficult and the signal processing method more complex than in the routines suggested above.

System Details

The general details of EMAT transduction process were presented in Section 3.3.2. Recall that an alternating current–carrying wire subjected to a static magnetic field will generate a wave when brought near a conductive material.

Copyright © Marcel Dekker, Inc. All rights reserved.

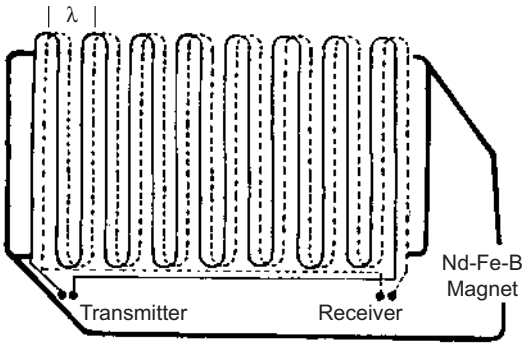


FIGURE 3.64 Geometry of EMAT coils for excitation of Rayleigh waves at 500 kHz with $\lambda = 6$ mm. The EMAT receiver and transmitter coils are stacked and staggered to reduce the system footprint.

To generate a surface-bound wave, the EMAT wires are run in a serpentine fashion (meander coil) with a periodicity equal to the wavelength of the surface wave. In this application $\lambda = 6$ mm.

The large footprint of the EMAT does not allow end-to-end placement of the receiver and transmitter, as in the piezoelectric case (Design Considerations in Section 3.5.1). By placing the receiver on top of the transmitter with the wire spacing staggered such that the receiver did not shield the transmitter, we reduced the footprint to the size of a single EMAT (Figure 3.64). (The generation mechanism requires eddy currents to be induced in the test material. If the receiver and transmitter wires were aligned, the eddy currents would appear in the transducer wires and not the wheel.)

The EMAT system was mounted into the rail on a spring-loaded pivot. As the wheel approached, the system would tilt towards the wheel thereby making contact a few moments earlier than if the system where fixed within the rail. As the wheel traveled the EMAT system would pivot and continue to maintain contact with the wheel. This pivot mechanism more than doubled the contact time over a fixed position system.

The transmit and receive electronics and the computer are housed in a trackside shed.

3.5.2 Ultrasonic Testing of Elevator Shafts and Axles (49)

Elevators are the safest means of transport. But during the long life of an elevator—some 20 or more years, with millions of rides—and despite regular maintenance, invisible symptoms of fatigue may occur, e.g., breaking wire inside

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

the elevator ropes, or cracking in the welds of cabin-suspension elements. In such events, however, the "safety catch" arrests uncontrolled cabin movement.

Material Fractures Involving Safety Risk

A fatigue fracture of the drive shaft of an elevator is critical, especially when the fracture separates the drive pulley from the elevator gear, allowing the pulley to rotate freely (Figure 3.65). The mechanical brake on the drive can brake only the motor and elevator gear, leaving the possibility of a free-wheeling drive pulley. Depending on the location of the fracture, the drive pulley is supported by only two bearings—or even only one. The cabin then runs, uncontrolled, downward or upward, depending on the load condition between counterweight and cabin.

If the drive shaft fractures during a regular cabin operation, a heavy-loaded cabin runs downward (even if it was previously going up), increasing in speed, until the safety catch brakes the cabin due to overspeed. This case is harmless. But a low-loaded cabin runs upward, increasing in speed, with the possibility of a colliding with the shaft ceiling and severely injuring the passengers. The usual safety catch cannot act during upward travel (nor, according to the present regulations, is it even allowed to do so). This collision could be prevented only by a second safety catch on the counterweight. Passengers could also be injured if the drive shaft fails while passengers are entering or leaving the elevator.

Ultrasonic Testing as a Preventive Measure

The most typical shaft flaws are transverse fatigue cracks running parallel to the shaft circumference. These occur most commonly at the fillet or near connections

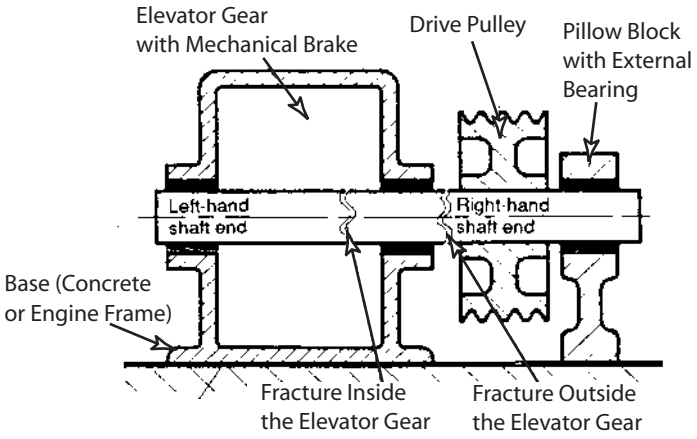


FIGURE 3.65 Elevator drive shaft housing and assembly. (Courtesy Krautkramer Branson.)

Copyright © Marcel Dekker, Inc. All rights reserved.



FIGURE 3.66 Photograph of UT inspection of drive shaft. (Courtesy Krautkramer Branson.)

to the bearings or pulley. If the inspector injects ultrasound axially, these flaws can be readily detected (Figure 3.66).

Figure 3.67 shows a drive-shaft geometry with two changes in shaft diameter. The two fillets oppose one another. Ideally, if the shaft diameter were constant, normal-incidence ultrasonic testing (using a longitudinal wave) would adequately test the shaft. To test the entire shaft, one simply inspects from both ends. But the actual drive-shaft geometry has two high-stress areas where the diameter changes. Normal-incidence ultrasound cannot adequately inspect these

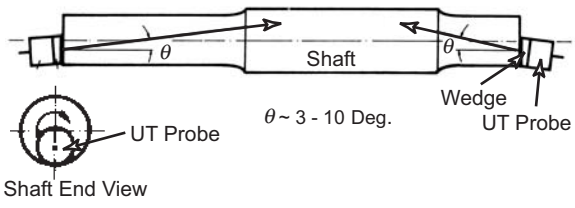


FIGURE 3.67 Typical drive shaft geometry showing two changes in shaft diameter. (Courtesy Krautkramer Branson.)

Copyright © Marcel Dekker, Inc. All rights reserved.

fillet regions. Although the ultrasound diverges, it still propagates in a straight-line path, thus creating shadow regions in the fillet area. It is much easier to inspect the shaft using angled-beam inspection. The specific shaft geometry determines the appropriate wedge angle. In this case, wedge angles between 3–10° produce good results.

Inspection Procedure

For angled-beam inspection, as with normal-incidence inspection, the operator inspects the shaft from both ends if access permits. As indicated in Figure 3.67, the inspector places the angled-beam transducer off-axis on the shaft end. (Many shafts have a recess in the center of the shaft end.) The probe is aligned such that the wedge angle decreases only in the radial direction. This probe-shaft alignment is maintained as the inspector moves the probe circumferentially around the shaft end. This procedure is repeated on the other end of the shaft.

False Indicators

Geometric features on the shaft can produce reflections not associated with actual flaws. The fillets themselves reflect the ultrasound. As the ultrasound propagates, the opposing fillet will reflect the wave and potentially mask any flaw within the fillet region. By inspecting from the opposite end, an actual flaw can be discerned. If access to the opposite end is unavailable, the operator must carefully choose the wedge angle and probe position to minimize the effect of the opposing fillet.

Additionally, keyways and press-fit bearings act as potential false indicators. If the probe is properly aligned, an axial keyway poses a small cross-sectional area to act as a reflector. Press-fit bearings require careful inspection of the response signal to accurately discern flaws. A qualified inspector can test the shaft in about 30 minutes.

3.6 ADVANCED TOPICS (OPTIONAL)

3.6.1 Interference

To understand many ultrasonic phenomena, it is important to first understand *interference*, or the interaction of waves when added together. Waves can add constructively (new amplitude 2A) or destructively (new amplitude 0), although these are the extreme cases; in general, waves can add to create waves with any amplitude between 0 and 2A.* And interference occurs in the same way with all types of waves, from mechanical to optical waves. This section explains constructive and destructive intereference of two waves of the same frequency.

*For a discussion of adding waves with slightly different frequencies, see Advanced Topics/Group and Phase Velocities.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

How does constructive interference work? First, consider two waves of the form

$$u_1(x,\ t) = A\sin(kx - \omega t)$$

and

$$u_2(x, t) = A\sin(kx - \omega t)$$

If these waves are combined, the total wave is:

$$u_{\text{total}} = u_1 + u_2 = 2A\sin(kx - \omega t) \tag{3.117}$$

Thus the waves constructively interfere, i.e., add constructively. Now imagine that one of the waves travels a distance Δx before impinging on the other wave:

$$\begin{aligned} u_1(x + \Delta x, t) &= A\sin(k(x + \Delta x) - \omega t) = A\sin(kx - \omega t + k\Delta x) \\ &= A\sin(kx - \omega t + \phi) \end{aligned} \tag{3.118}$$

where ϕ is called the *phase factor*, expressed in degrees or radians. Thus, the equation for the phase *factor* due to distance traveled is:

$$\phi = k\Delta x = 360^\circ \Delta x/\lambda(\text{degrees}) = 2\pi\Delta x/\lambda(\text{radians}) \tag{3.119}$$

If the first wave and the second wave are combined, the total wave is:

$$\begin{aligned} u_{\text{total}} &= u_1 + u_2 = A\sin(kx - \omega t + \phi) + A\sin(kx - \omega t) \\ &= 2A\cos(\phi/2)\sin(kx - \omega t + \phi/2) \end{aligned} \tag{3.120}$$

Therefore, if the phase factor is zero, then the total wave adds constructively so that the new amplitude is $2A$, as described in Eq. (3.117). The total wave also adds constructively if the phase factor is a multiple of 360° (3.117) (or 2π radians). On the other hand, if the phase factor is 180° and $\cos(180^\circ/2) = 0$, then the total wave will have an amplitude of zero. In that case, the waves destructively interfere, i.e., add destructively, in effect canceling each other. The equations for the phase factors are as follows:

$$\text{Constructive:}\ \ \phi = n360^\circ \quad n = 0,\ 1,\ 2,$$

$$\text{Destructive:}\ \ \phi = \left(n + \frac{1}{2}\right)360^\circ \quad n = 0,\ 1,\ 2,$$

3.6.2 Group Velocity and Phase Velocity

Imagine a wave packet where a wave with a high frequency has a modulating amplitude at a lower frequency (Figure 3.68). The speed of the wave packet is called the group velocity and the speed of the individual waves is called the phase velocity. To understand the difference, imagine that the phase velocity was zero while the group velocity was positive. In this case, the position of the waves

Copyright © Marcel Dekker, Inc. All rights reserved.

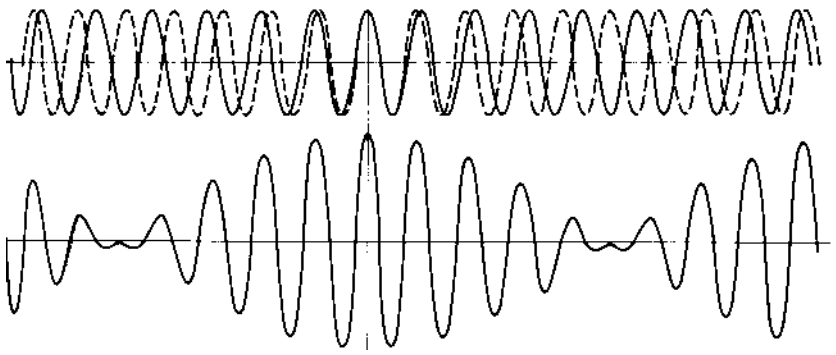


FIGURE 3.68 Frequency modulated wave.

would remain the same while only the amplitude would modulate and the envelope would move. If, on the other hand, the group velocity was zero while the phase velocity was positive, then the waves would move but the envelope would remain constant. If the group velocity equaled the phase velocity, then the shape of the wave would remain constant as the wave traveled.

The group velocity and the phase velocity differ only when different frequencies have different velocities. Consider a simple example of two waves with slightly different velocities and frequencies:

$$u_1 = A\cos(k_1 x - \omega_1 t)$$

and

$$u_2 = A\cos(k_2 x - \omega_2 t) \tag{3.121}$$

with phase velocities of $v_1 = k_1/\omega_1$ and $v_2 = k_2/\omega_2$. The total displacement is then:

$$\begin{aligned} u_{\text{total}} &= u_1 + u_2 = A\cos(k_1 x - \omega_1 t) + A\cos(k_2 x - \omega_2 t) \\ &= 2A\cos(k_{\text{diff}} x - \omega_{\text{diff}} t)\cos(k_{\text{ave}} x - \omega_{\text{ave}} t) \end{aligned} \tag{3.122}$$

where:

$$\begin{aligned} k_{\text{diff}} &= \frac{k_1 - k_2}{2} \qquad \omega_{\text{diff}} = \frac{\omega_1 - \omega_2}{2} \\ k_{\text{ave}} &= \frac{k_1 + k_2}{2} \qquad \omega_{\text{ave}} = \frac{\omega_1 + \omega_2}{2} \end{aligned} \tag{3.123}$$

Therefore, the wave is composed of two waves, one with a velocity of $k_{\text{ave}}/\omega_{\text{ave}}$ (phase velocity) and one with a velocity of $k_{\text{diff}}/\omega_{\text{diff}}$ (group velocity). In other words, the group velocity can be defined as:

$$v_{\text{group}} = \frac{\Delta\omega}{\Delta k} \tag{3.124}$$

this is the common definition of group velocity.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

In which types of waves does the group velocity differ from the phase velocity? For longitudinal, shear, and interface waves the velocity is independent of frequency, so the difference between phase and group velocity is nonexistent. But bound waves (such as plate, tube waves etc.) do have frequency-dependent velocities, so they have different phase and group velocities. The variation of velocity with frequency is called *dispersion* and is graphed in *dispersion curves*. Figure 3.69 is a diagram of one such dispersion curve for an aluminum plate. Dispersion curves can be used to determine the optimum incident angle or comb separation for creating a specific mode at a set frequency and thickness (Guilded Waves in Section 3.2.6).

3.6.3 Method of Potentials

Acoustic waves can be described by a function $\vec{u}(x, y, z, t)$, where $\vec{u}$ describes the particle displacement and is a vector. The function $\vec{u}$ is a function of a longitudinal wave and a shear wave:

$$\vec{u} = \vec{u}_L + \vec{u}_S \tag{3.125}$$

Particle displacement differs for different types of waves.

Longitudinal Potential

Longitudinal waves create particle displacement that is parallel to the direction of propagation. Therefore, if the direction of propagation is $\hat{n}$, then $\vec{u}_L$ can be written as $u_L\hat{n}$, where u_L is a scalar (not a vector). The direction of propagation is proportional to the change in particle displacement over space. In mathematics, the way to describe how something changes over space is with the gradient ∇, defined as:

$$\nabla = \frac{\partial}{\partial x}\hat{i} + \frac{\partial}{\partial y}\hat{j} + \frac{\partial}{\partial x}\hat{k} \tag{3.126}$$

where $\hat{i}$, $\hat{j}$, and $\hat{k}$ are unit direction vectors along the x, y, and z directions. One can thus define the particle displacement due to longitudinal waves as:

$$\vec{u}_L = \nabla\phi_L \tag{3.127}$$

where ϕ_L is called the *potential* of the longitudinal wave and is a scalar.

Shear Potential

Unlike longitudinal waves, shear waves vibrate in a direction perpendicular to the direction of motion. After a little mathematics, one can show that $\nabla \cdot \vec{u}_S = 0$, so that the shear wave displacement can be written as:

$$\vec{u}_S = \nabla \times \vec{\phi}_S \tag{3.128}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

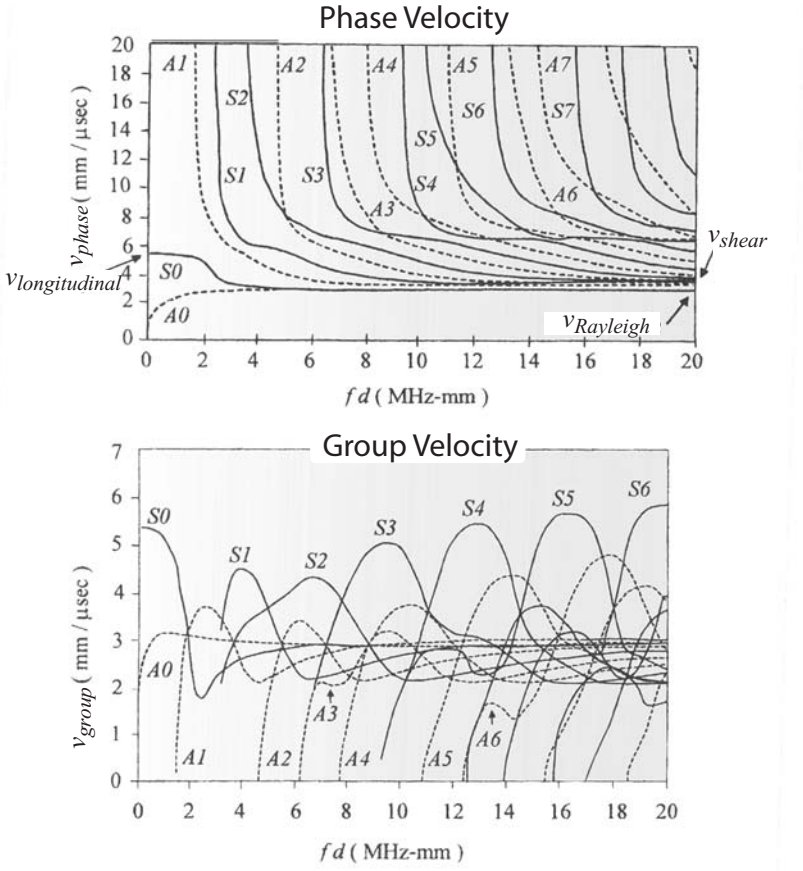


FIGURE 3.69 Dispersion curves (phase and group velocity) for plate waves in an aluminum plate.

where $\vec{\phi}_S$ is called the potential of the shear wave and is a *vector*. Combining Eqs. (3.125, 3.127, 3.128), results in*

$$\vec{u} = \nabla\phi_L + \nabla \times \vec{\phi}_S \tag{3.129}$$

At first, the above equation seems more complicated than helpful. However, due to the nature of the motion of waves, both of the potentials follow the time-

*This is similar to the relationship in electrodynamics where the electric field is the gradient of the scalar electrostatic potential and the magnetic field is the cross of a magnetic potential which is a vector.

Copyright © Marcel Dekker, Inc. All rights reserved.

dependent Helmholtz equation, ignoring diffusion:

$$\nabla^2\phi = \frac{1}{v^2}\frac{\partial^2\phi}{\partial t^2} = -k^2\phi \tag{3.130}$$

where v is the speed of the wave and k is the spatial frequency (or the wavenumber).

To understand how to use the above equations, consider the simple case of a longitudinal wave propagating with speed v in the x direction. Eq. (3.130) can now be written as:

$$\frac{\partial^2\phi(x,\ t)}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2\phi(x,\ t)}{\partial t^2} = -k^2\phi(x,\ t) \tag{3.131}$$

which is solvable for ϕ:

$$\phi(x,\ t) = Ae^{i(kx-\omega t)} \tag{3.132}$$

where A is an amplitude constant. Therefore the displacement vector becomes:

$$\begin{aligned}\vec{\boldsymbol{u}} &= \nabla\phi_L + \nabla\times\vec{\boldsymbol{\phi}}_S \\ &= \frac{\partial}{\partial x}\left[Ae^{i(kx-\omega t)}\right] + 0 \\ &= A'e^{i(kx-\omega t)}\end{aligned}$$

where $A' = A(ik)$. Therefore, the particles vibrate like a harmonic wave in the x direction as expected.

Potential in Surface Acoustic Waves (SAWs)

For simplicity, first consider a wave on the surface of an elastic solid. Assume that the surface is in the x–y plane and the medium is in the z direction, so that both potentials have no y component. An SAW is a combination of both a shear wave and a longitudinal wave. You may wonder how a shear wave and a longitudinal wave can interfere, since they are perpendicular. The answer is interesting: both waves propagate in both the x and the z direction! But since the wave moving in the z direction has an exponentially decaying term, the wave acts like a combined shear-longitudinal wave on the surface. For this reason, both potentials have x and z components.

Consider the Helmholtz equation for the shear and longitudinal potential, from Eq. (3.130):

$$\begin{aligned}\frac{\partial^2\phi_L(x,\ z,\ t)}{\partial x^2} + \frac{\partial^2\phi_L(x,\ z,\ t)}{\partial z^2} &= \frac{1}{v_L^2}\frac{\partial^2\phi_L(x,\ z,\ t)}{\partial t^2} = -k_L^2\phi_L(x,\ z,\ t) \\ \frac{\partial^2\vec{\phi}_S(x,\ z,\ t)}{\partial x^2} + \frac{\partial^2\vec{\phi}_S(x,\ z,\ t)}{\partial z^2} &= \frac{1}{v_S^2}\frac{\partial^2\vec{\phi}_S(x,\ z,\ t)}{\partial t^2} = -k_S^2\vec{\phi}_S(x,\ z,\ t)\end{aligned} \tag{3.133}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

where

$$\begin{aligned}\vec{u} &= \nabla\phi_L + \nabla \times \vec{\phi}_S \\ &= \frac{\partial\phi_L}{\partial x}\hat{i} + \frac{\partial\phi_L}{\partial z}\hat{k} - \frac{\partial\phi_{Sy}}{\partial z}\hat{i} + \frac{\partial\phi_{Sy}}{\partial x}\hat{k}\end{aligned} \tag{3.134}$$

Since the displacement should be zero in the y direction, then $\phi_{Sx} = \phi_{Sz} = 0$ or $\vec{\phi}_S(x, z, t) = \phi_S(x, z, t)\hat{j}$. For this wave to be a coherent wave on the surface of the material, the velocity in the x direction must be the same for the shear and longitudinal waves. Therefore

$$\begin{aligned}\frac{\partial^2\phi_L(x, z, t)}{\partial x^2} &= \frac{1}{v^2}\frac{\partial^2\phi_L(x, z, t)}{\partial t^2} = -k^2\phi_L(x, z, t) \\ \frac{\partial^2\vec{\phi}_S(x, z, t)}{\partial x^2} &= \frac{1}{v^2}\frac{\partial^2\vec{\phi}_S(x, z, t)}{\partial t^2} = -k^2\vec{\phi}_S(x, z, t)\end{aligned} \tag{3.135}$$

where v is the phase velocity in the x direction. This problem has been solved in Eq. (3.132), to be

$$\begin{aligned}\phi_L(x, z, t) &= A(z)\exp(ikx - i\omega t) \\ \phi_S(x, z, t) &= B(z)\exp(ikx - i\omega t)\end{aligned} \tag{3.136}$$

Note that ϕ_L is not equal to ϕ_S, although they have the same x dependence. Placing the above equations back into the differential equations, we obtain

$$\begin{aligned}\phi_L(x, z, t) &= [A_1e^{-kpz} + A_2e^{kpz}]\exp(ikx - i\omega t) \\ \phi_S(x, z, t) &= [B_1e^{-kqz} + B_2e^{kqz}]\exp(ikx - i\omega t)\end{aligned} \tag{3.137}$$

The e^{kpz} and e^{kqz} terms are physically impossible, since the wave would be infinite as $z \to \infty$; therefore, $A_2 = B_2 = 0$.

Finally, the displacement vector on the surface becomes

$$\vec{u}(x, z, t) = [A'\hat{i} + iB'\hat{k}]\exp(ikx - i\omega t) \tag{3.138}$$

where A' and B' are real numbers. The real part of the displacement vector is

$$\vec{u}(x, z, t) = A\cos(kx - \omega t)\hat{i} + B\sin(kx - \omega t)\hat{k} \tag{3.139}$$

The above equation describes an ellipse. The value of A and B can be determined by satisfying the boundary conditions of continuity of the stress. After some mathematics, the amplitude of the constants becomes a function of the Poisson ratio of the materials.

Potential in Plate Waves

The mathematics of plate waves, also called Lamb waves, are amazingly similar to the mathematics of SAWs. As with SAWs, if z is the normal to the surface and

Copyright © Marcel Dekker, Inc. All rights reserved.

x is the direction of total propagation, then the potentials do not depend on the y value. Also, if the shear and longitudinal waves combine to create a new wave, this new wave must have the same velocity in the x direction. Consider the solution for the longitudinal and shear wave potential for a Raleigh wave from Eq. (3.137):

$$\begin{aligned}\phi_L(x,\ z,\ t) &= [A_1 e^{-kpz} + A_2 e^{kpz}]\exp(ikx - i\omega t)\\ \phi_S(x,\ z,\ t) &= [B_1 e^{-kqz} + B_2 e^{kqz}]\exp(ikx - i\omega t)\end{aligned} \tag{3.140}$$

As Lamb waves are finite systems in the z direction, the variables A_2 and B_2 do not equal zero:

$$\begin{aligned}\phi_L(x,\ z,\ t) &= [A_1 e^{-kpz} + A_2 e^{kpz}]e^{(ikx-i\omega t)}\\ &= [A_i \sinh(kpz) + A_{ii}\cosh(kpz)]e^{(ikx-i\omega t)}\\ \phi_S(x,\ z,\ t) &= [B_1 e^{-kqz} + B_2 e^{kqz}]e^{(ikx-i\omega t)}\\ &= [B_i \sinh(kpz) + B_{ii}\cosh(kpz)]e^{(ikx-i\omega t)}\end{aligned} \tag{3.141}$$

The Lamb waves are found to split into two modes groups, one symmetric (S_n) around the center of the plate, and one antisymmetric (A_n) around the plate. In each mode group there are an infinite number of modes. (The subscript n indicates the specific mode.) For this reason, we can divide the potentials into symmetric and antisymmetric cases. In the symmetric case:

$$\begin{aligned}\phi_L(x,\ z,\ t) &= A\cosh(kpz)\exp(ikx - i\omega t)\\ \phi_S(x,\ z,\ t) &= B\sinh(kqz)\exp(ikx - i\omega t)\end{aligned}$$

The displacement vector then becomes:

$$\begin{aligned}\vec{u}(sy) = \{&[ikA\cosh(kpz) - kqB\cosh(kqz)]\hat{i}\\ &+ [kqA\sinh(kpz) + ikB\sinh(kqz)]\hat{k}\}e^{(ikx-i\omega t)}\end{aligned} \tag{3.142}$$

In the antisymmetric case:

$$\begin{aligned}\phi_L(x,\ z,\ t) &= A\sinh(kpz)\exp(ikx - i\omega t)\\ \phi_S(x,\ z,\ t) &= B\cosh(kqz)\exp(ikx - i\omega t)\end{aligned}$$

The displacement vector then becomes:

$$\begin{aligned}\vec{u}(sy) = \{&[ikA\sinh(kpz) - kqB\sinh(kqz)]\hat{i} + [kqA\cosh(kpz)\\ &+ ikB\cosh(kqz)]\hat{k}\}e^{(ikx-i\omega t)}\end{aligned} \tag{3.143}$$

Note that this separation into symmetric and antisymmetric modes was possible only because of the symmetry of the system; it is not possible in tubes. Finding the surface stress gives a relationship between q, p and k that can be

Copyright © Marcel Dekker, Inc. All rights reserved.

calculated numerically. The lowest value is the zeroth order of the wave, the next the first order, and so on.

3.6.4 Derivation of Snell's Law: Slowness Curves

For the moment, let us revisit the boundary conditions that we applied in the case of a *normal* incidence wave. The boundary conditions were:

1. Continuity of particle velocity
2. Continuity of acoustic pressure

From these the two conditions, we obtained the normal incidence reflection and transmission coefficients Eqs. (3.81 and 3.82). But we avoided an in-depth discussion of the continuity of the phase of the waves at the boundary by setting the interface at $x = 0$. We will now explain the origins of this required boundary condition.

Again, assume plane waves propagating in homogeneous, isotropic material. As discussed in Section 3.2.1, a propagating plane wave consists of

1. $P(t, r)$, an amplitude coefficient that may vary in both time and space.
2. $e^{j(\omega t - \vec{k}\cdot\vec{r})}$, a periodic propagation multiplier that specifies how the wave modulates in both space, $e^{-j\vec{k}\cdot\vec{r}}$, and time, $e^{j\omega t}$.* Recall that $\omega t - \vec{k}\cdot\vec{r}$ is the phase of the wave.†

To meet the continuity criteria at the boundary, both components of the wave must be continuous across the interface. The continuity of the phase is what we call Snell's law (Snell's Law in Section 3.2.4).

Continuity of acoustic pressure is

$$P_i e^{j(\omega_i t - \vec{k}_i\cdot\vec{r})} + P_r e^{j(\omega_r t - \vec{k}_r\cdot\vec{r})} = P_t e^{j(\omega_t t - \vec{k}_t\cdot\vec{r})} \tag{3.144}$$

This must be true for any real physical wave problem. Assume P_i, P_r, and P_t are constants. Pick a fixed position on the interface; $\vec{r}$ = constant. Therefore, $e^{-j\vec{k}\cdot\vec{r}}$ terms are constants. Let P_i', P_r', and P_t' represent the combined constant amplitude coefficients and constant $e^{-j\vec{k}\cdot\vec{r}}$ terms:

$$P_i' e^{j(\omega_i t)} + P_r' e^{j(\omega_r t)} = P_t' e^{j(\omega_t t)} \tag{3.145}$$

For this continuity condition to be true for all time, $\omega_i = \omega_r = \omega_t = \omega$. Because $e^{j\omega t}$ is common to all terms, it can be dropped from the *equation*, as is commonly

*No matter how complex the waveform, Fourier analysis will show that the wave can be represented as a series of sinusoidal waves.

†The temporal and spatial frequencies, ω and k, are not functions of position or time.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

done in the literature. This leaves the phase components of the form $e^{-j\vec{k}\cdot\vec{r}}$. Recall that $\vec{k}\cdot\vec{r} = k_x x + k_y y + k_z z$. By a rotation of the coordinate system to $y = 0$, the boundary plane is defined by $z =$ constant. Therefore, $e^{-jk_z^i z}$, $e^{-jk_z^r z}$, and $e^{-jk_z^t z}$ terms are constants albeit different (The *superscripts* on k indicate the incident, reflected, or transmitted wave.) Because P_i, P_r, P_t, and $e^{-jk_z z}$ terms are constants, we can combined the constants as P_i'', P_r'', and P_t'' to yield

$$P_i'' e^{j(k_x^i x)} + P_r'' e^{j(k_x^r x)} = P_t'' e^{j(k_x^t x)} \tag{3.146}$$

For this boundary condition to be valid for an arbitrary value of x, $k_x^i = k_x^r = k_x^t = k_x$. k_x is the magnitude of the component of the spatial frequency vector (wave vector) parallel (tangent) to the interface. Thus, $k_x^i = k_x^r = k_x^t$ is a statement of continuity of the tangential component the spatial frequency vector. Recall $k = \omega/v$ and $k_x = k\sin(\theta)$* are the tangential component of $\vec{k}$. Combining these, we have Snell's law:

$$\frac{\omega}{v_i}\sin(\theta_i) = \frac{\omega}{v_r}\sin(\theta_r) = \frac{\omega}{v_t}\sin(\theta_t) \tag{3.147}$$

Even from this cursory derivation, Snell's law is clearly an integral part of the each of the two boundary conditions on particle velocity and normal tractions (acoustic pressure).

Snell's law, continuity of tangential spatial frequency vector (wave vector), can be implemented in a useful graphical format. This graphic is shown in Figure 3.70, where $\vec{k}$, the spatial frequency vector, is plotted in each of the materials as a function of direction. The magnitude of $\vec{k}$, the spatial frequency, is represented as the distance from the curve to the origin on the boundary. The direction of $\vec{k}_i$ is given by θ_{il}, which is measured from the boundary normal. Because $k = \omega/v$ (frequency/velocity), these curves are called *slowness curves*.† Figure 3.70a shows the slowness curves for acoustically isotropic materials, i.e., the wave velocity is the same in all directions. The multiple slowness curves in each material represent *multirefringence*, i.e., the different wave modes that can occur in reflection and refraction (transmission). Figure 3.70b shows slowness curves for one *anisotropic* material and one isotropic material. In anisotropic material the wave velocity is *not* the same in all directions. Simply, the slowness curves trace the length (magnitude) of the spatial frequency vector $\vec{k}$, for all possible angles of incident or scattered waves. The phase of the wave travels in the direction of $\vec{k}$ with phase velocity v, but the energy (momentum transfer) of the wave propagates in a direction normal to the curve at the group velocity. For isotropic material, these directions coincide.

---

*$\vec{k}\cdot\vec{r} = |\vec{k}||\vec{r}|\cos(\phi) = |\vec{k}||\vec{r}|\sin(\theta)$, where ϕ is measured from the interface plane to the direction of the wave, i.e., $\phi = 90^\circ - \theta$.

†In three dimensions the slowness curve is a surface.

Copyright © Marcel Dekker, Inc. All rights reserved.

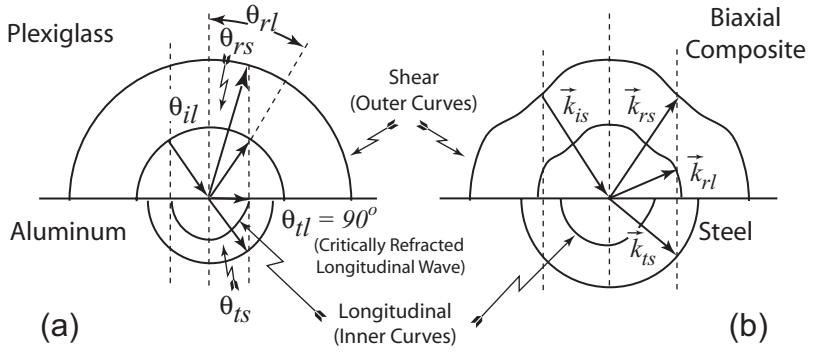


FIGURE 3.70 Slowness surfaces for (a) two isotropic materials and (b) one anisotropic and one isotropic.

To determine the number and direction of possible scattered modes, we match the magnitude of the tangential component of the *scattered* modes to the tangential components of the *incident* wave. Assume a longitudinal wave is incident on the interface as shown in Figure 3.70a. The *length* of the arrow represents k_{il} as given by the definition of the slowness curve. The incident angle is θ_{il} and is measured from the interface normal as usual. (For clarity one curve indicates the angle and the other indicates the spatial frequency.) For continuity of the tangential component of the spatial frequency vector (wave vectors), the directions of propagation of $\vec{k}_{rs}$, $\vec{k}_{rl}$, $\vec{k}_{ts}$, and $\vec{k}_{tl}$ must be θ_{rs}, θ_{rl}, θ_{ts}, and, θ_{tl}, respectively, as shown in Figure 3.70a. Note $\theta_{tl} = 90^\circ$, therefore, θ_{il} is the longitudinal critical angle. Figure 3.70b shows the potential scattered waves for an incident shear (SV) wave.

PROBLEMS

Introduction

1. Can you hear the sound of a ships horn faster if you are swimming underwater or if you are floating on top? Why?
2. Describe the basic theory of Ultrasonics.
3. Name three important historical advances in ultrasonic science.
4. What are the two biggest limitations of ultrasonic NDE? The two biggest advantages?

Theory

A. Introduction to Wave Propagation

5. Suppose someone shouted at you from across the room, would the vibrating air particles near the person reach you? Why or why not?

Copyright © Marcel Dekker, Inc. All rights reserved.

6. What is the difference between longitudinal and transverse waves? Which has a higher velocity?
7. Is the wave in a guitar string longitudinal or transverse? How about the surface wave in water? The sound from the beating of a drum?
8. What is the difference between the phase velocity and the particle velocity?
9. If the peak of a wave at time equal to 0 s is at 2 mm, where is the peak at 5μs if the wave moves at a speed of 3 mm/μs?
10. If the period is 5μs, and the wavelength is 2 cm, what is the linear spatial frequency, linear temporal frequency, angular spatial frequency, angular temporal frequency, and velocity?
11. If the velocity of a wave is 5 mm/μs and the linear *temporal* frequency, f, is 5 MHz, what is the wavelength, the linear *spatial* frequency, the angular spatial frequency, the angular temporal frequency and the period?
12. Show that $\omega/k = \lambda/\tau = f/\chi =$ distance traveled/time = velocity.

B. Wave Motion and the Wave Equation

13. Show that a wave of the form $u(x, t) = A\cos(kx + \omega t)$ satisfies the wave equation (Eq. [15])
14. Show that a wave of the form $u(x, t) = A(kx - \omega t)^2$ does not satisfy the wave equation (Eq. [15]).
15. If the density of a material doubles, what happens to the longitudinal phase velocity? The shear phase velocity?
16. If two materials have the same density but material 1 has 4 times the longitudinal phase velocity as material 2, what is the ratio of their Young's modulus?
17. If the Poisson's ratio $v = 0.3$, by what factor does that change the longitudinal phase velocity from the one dimensional value?
18. Show that $e_{ji} = 0$ for simple shear shown in Figure 3.9, use the definition of rigid angular rotation and strain.

C. Interference

19. What is the phase needed so that the total amplitude of the sum of two waves of amplitude A is A?
20. What is the minimum difference in path length to create destructive interference between waves with a wavelength of 3 mm?
21. What is the minimum difference in path length to create constructive interference between waves with a wavelength of 3 mm?
22. If one knows the difference in path length to create constructive interference, is this enough information to determine the wavelength of the wave? Why or why not?

D. Specific Acoustic Impedance and Pressure

23. What is the change in density needed to double the acoustic impedance?

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

24. What is the acoustic impedance of aluminum (longitudinal velocity of 6.42 km/s, and density of 2.7×10^6 kg/m^3)?
25. Show that if $u(x, t) = A\sin(kx - 2\omega t)$, then $\dfrac{\partial u/\partial x}{\partial u/\partial t} = \dfrac{1}{v}$.
26. Is the acoustic impedance greater, lesser or the same for shear waves as for longitudinal waves? Why?

E. Reflection and Refraction at an Interface

27. What are the three boundary conditions for ultrasonic wave propagation?
28. If the incident pressure is 1, what is the reflected and transmitted pressures if a wave impinges normal to a steel/water interface? The impedances are $Z_{H_2O} = 1.5 \times 10^6$ N s/m^3 and $Z_{steel} = 45 \times 10^6$ N s/m^3. What are the pressures if the wave travels from water to steel?
29. If the incident intensity is 1, what is the reflected and transmitted intensities if a wave travels from steel to water (normal to the interface)? The impedances are $Z_{H_2O} = 1.5 \times 10^6$ N s/m^3 and $Z_{steel} = 45 \times 10^6$ N s/m^3. What are the intensities if the wave travels from water to steel?
30. Show that the transmitted intensity equals the initial energy if $Z_1 = Z_2$.
31. Show that the reflected intensity equals the initial energy if $Z_1 \gg Z_2$.
32. If the reflected energy is equal to half the initial energy, what is the ratio of Z_1 to Z_2
33. What is the difference between a shear horizontal wave and a shear vertical wave?
34. Why does one need an oblique angle to an interface to describe shear horizontal and shear vertical waves?
35. What are the refracted and reflected angles of the longitudinal and shear waves if a shear wave in plastic impinges on inconel at an angle of 45 degrees? The velocity in plastic is 2.7 km/ms (longitudinal) and 1.1 km/ms (shear). The velocity in inconel is 5.7 km/ms (longitudinal) and 3.0 km/ms (shear).
36. What is the first critical angle for a longitudinal wave propagating in plastic incident onto inconel? (Velocities given in Problem 35.)
37. What is the second critical angle for a longitudinal wave in plastic with inconel? (Velocities given in Problem 35.)
38. Why is the first critical angle *always* less then the second critical angle?

F. Attenuation

39. Describe the four basic methods of ultrasonic attenuation.
40. Which would attenuate faster, a 2.5 MHz wave or a 1 MHz wave? Why?
41. Would a wave attenuate faster in a high density material or a low density material? Why?
42. What is the difference between wave scattering and wave absorption?

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

43. If an experimentalist wished to determine the *size* of a defect, what frequency would he or she use: high, medium, or low and why? (Hint: recall that $v = f\lambda$)
44. If an experimentalist wished to determine the *shape* of a defect, what frequency would he or she use: high, medium or low and why? (Hint: recall that $v = f\lambda$)
45. If the intensity of an acoustic wave decreases to $1/e$ of the initial intensity in 20 mm, what is the attenuation constant?
46. If the intensity is 0.002 times the original intensity, what is the ratio of intensities in dB?
47. A -40 dB intensity change implies that the intensity is equal to the original intensity divided by what number?
48. For the values in Problem 45, what is the dB attenuation constant?

Transducers for Generation and Detection of Ultrasonic Waves

A. Piezoelectric Transducers

49. Describe the piezoelectric effect and the reverse-piezoelectric effect.
50. What thickness piezoelectric crystal would be used to create a frequency of 2 MHz if the velocity in the crystal is 11.1 km/s?

B. Electromagnetic Acoustic Transducers (EMAT)

51. Describe the basic theory of EMAT transducers.
52. Give a verbal explanation of why the EMAT geometry shown in Fig. 3.41 produces a diffraction pattern that does not show any energy propagating perpendicular to the aperture.
53. Could a DC current be used in the EMAT wires? Why or why not?

C. Optical Transducers

54. Why does the method to create ultrasonic waves with lasers have to be different then the method to measure the ultrasonic wave?
55. What are some difficulties in using optical measurement methods?
56. What is the purpose of the reference beam? The object beam?

D. Transducer Characteristics

57. Show that Eq. (110) gives basically the same answer as Eq. (109) if the wavelength is 2 mm and the aperture width is 40 mm.
58. Show that Eq. (110) is not viable if the wavelength equals the aperture width.
59. If the aperture width is much greater then the wavelength, by what factor does the near field length change by if the aperture is doubled?

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

60. What is the angle of divergence for a beam if the wavelength is 2 mm and the aperture width is 40 mm?
61. Would a shear wave or a longitudinal wave of the same frequency have a greater angle of divergence? Why?

Inspection Principles

A. Measurements

62. Name the four variables that an ultrasonic inspector can measure.
63. What should be measured to determine the size of an object?
64. What should be measured to determine the position of a flaw?
65. What should be measured to determine the existence of a flaw?
66. What should be measured to determine the density of a material?
67. If a sample has a resonant frequency at 1.10 MHz and 1.22 MHz, in a material 20 mm thick, what is the velocity of the system?

B. Creation of Waves

68. Describe the four basic methods to create longitudinal waves.
69. Describe the four basic methods to create shear waves.
70. Describe the four basic methods to create interface waves.
71. What angle of plastic would be used to create the Rayleigh wave in steel if a longitudinal transducer is used? The velocity in plastic is 2.7 km/ms (longitudinal) and 1.1 km/ms (shear). The velocity in steel is 5.9 km/ms (longitudinal) and 3.2 km/ms (shear).

C. Transducer Setup

72. Describe the difference between the pitch–catch and the through transmission method. Which method is better for very thin samples? Which method is better for highly attenuating methods? Which method is easier to align?

D. Driver for Transducers

73. Describe the three types of pulse generators (drivers).
74. What are the advantages and the disadvantages of a large pulse width?

E. Receiving the Signal

75. How is an oscilloscope used to interpret the ultrasonic signal?
76. Describe the B-Scan, M-Scan, and C-Scan.
77. What is the difference between a C-Scan and a SAM.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Advanced Topics

A. Group Velocity and Phase Velocity

78. Describe the difference between the group velocity and the phase velocity.
79. Show that if the phase velocity is independent of the frequency then the group velocity is the same as the phase velocity. Hint: define ω as a function of k and v.
80. Using Figure 3.69, what angled plastic wedge would you use to create the A_1 wave in a plate of inconel 1.5 mm thick with a 2 MHz longitudinal transducer. ($v_{\text{plastic}} = 2.7\,\text{km/s}$). Hint: use Snell's law where the angle in the plate is 90°.

B. Method of Potentials

81. If the longitudinal potential is a function of x, what direction is the longitudinal wave in?
82. If the shear potential vector is a function of x, what direction is the shear wave in?
83. How can a longitudinal wave and a shear wave have the same velocity in the x direction?

C. Slowness Curves and the Derivation of Snell's Law

84. What is the slowness curve a graph of and why is it useful?

GLOSSARY

Acoustic Fresnel equations: Pressure ratios for reflection and transmission.
Acoustic pressure: The driving force that produces the resulting ultrasonic wave.
Anisotropic: Having different elastic constant for different directions.
A-scan: A measure of the amplitude of a longitudinal wave that reflects off an object.
Attenuation: Loss of energy in wave due to interactions with a material or due to beam spreading.
Bandwidth: The range of frequencies of a transducer.
Beam spreading: Geometric expansion of a nonplanar wavefront.
Birefringence: Property of a material with two velocities of waves.
B-mode: Single ultrasonic pulse reading versus time transformed into an intensity signal.
B-scan: A B-mode that is scanned across the length of an object and converted into an intensity measurement to create a 2-D slice of the object.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Continuous wave generator: Driver that produces a continuous wave; used often in resonant frequency measurements.

Critical angle: Angle where the refracted wave is at 90 degrees from the surface normal.

First critical angle: Angle where the refracted longitudinal wave is at 90 degrees from the surface normal.

C-scan: Intensity from a specific point scanned across the surface of the sample and transformed into a 2-D slice parallel to the pulse direction.

Decibel: 10 bels, based on base-10, logarithmic scale of the ratio of *energies* or *powers*.

Delaminations: Separations of layers of a composite material.

Dilatational wave: A bulk longitudinal wave that changes the volume of the volume elements.

Dispersion: Variation of velocity with respect to frequency.

Dispersion curve: Plot of velocity versus frequency.

Doppler Effect: Change in frequency of a wave due to the interaction of a wave with a moving target and/or source.

Drivers: Electronic device that provide the electromagnetic pulse to the transducer.

Electromagnetic Acoustic Transducer (EMAT): Transducer that uses a combination of currents in wires and external magnets to create or receive ultrasonic waves (shear horizontal, shear vertical, and longitudinal).

Fabry–Perot method: Optical method that uses interference, resonance, beats and the Doppler effect to detect ultrasonic waves.

Far field: Area where the intensity from the center of a piezoelectric transducer does not fluctuate with distance; area usually used for measurements.

Fresnel equations: Equations to determine the intensity of reflected and refracted waves

Frequency

- *Linear temporal frequency:* Cycles per time.
- *Linear spatial frequency:* Cycles per length.
- *Angular temporal frequency:* Radians per time.
- *Angular spatial frequency:* Radians per length.

Fundamental frequency: Lowest-mode resonant frequency for a bar or disk, when the wavelength equals twice the thickness.

Geometrical attenuation: Attenuation due to the geometric spreading of the wave, e.g., a spherical wave where the waves far from the source have lower amplitude then those close to the source.

Grass: Random noise due to diverted waves from scattering.

Hertz: Vibrations per second.

Homogeneous: Having constant density.

Impedance: Resistance that restricts movement to a finite velocity.

Inhomogeneity: Measure of density variation within a material.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Interference: Interaction of waves when added together.
Constructive: Addition of two waves where the new amplitude is sum of individual amplitudes.
Destructive: Addition of two waves where the new amplitude is difference of individual amplitudes.
Isotropic: Having equal elastic constants along every axis.
Knife-edge method: Optical method to detect ultrasonic waves by reflecting a laser off a sample onto edge of a barrier.
Lamb wave: Resonant vibration in flat plates with free surfaces, also called plate wave
Lamé constant: Constant used to relate stress and strain.
Laser interferometry: Method using lasers and interference to determine motion.
Longitudinal wave: Wave where the particle motion is parallel to the wave propagation direction; also known as *pressure wave* or *p-wave*.
Mediator: Wedge of metal with a protruding tip used with the wedge method as an intermediate medium for a surface wave.
Michelson interferometry: Optical method to detect ultrasonic waves by the interference of optical waves.
Mode conversion: The act of a wave reflecting or refracting and creating a different type of wave (e.g., longitudinal to shear or shear to longtudinal).
M-scan: Single ultrasonic pulse reading versus time transformed into an intensity signal that is examined as a function of time; used often in medical imaging.
Near field: Area from the center of a piezoelectric transducer where the intensity fluctuates with distance.
Near field length, N: Distance separating the near field from the far field.
Optical transducer: System that can transform optical energy into ultrasonic waves.
Oscilloscope: Machine that graphs voltage difference versus time.
Piezoelectric
  Effect, direct: The property of creating a potential difference due to mechanical deformation.
  Effect, reverse: The property of creating mechanical deformations due to a potential difference.
  Transducers: Objects that transform electrical energy to ultrasonic energy and visa-versa.
Period: Amount of time required for one complete cycle to move past a stationary observer.
Phased array: Set of several normal piezoelectric transducers driven at slightly different times, so that they interfere constructively and destructively into an angled longitudinal wave.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Pitch–catch: Method where one transducer transmits and receives signal.
Plane wave: A wave whose direction of propagation is always linear.
Polar diagrams: Diagrams of intensity from a transducer versus angle.
Pressure waves: See *Longitudinal waves*.
Pulse generator: Driver that produces a voltage spike.
Pulse–echo method: Ultrasonic method whereby the same transducer receives a reflected signal as created the signal.
P-waves: See *Longitudinal waves*.
Rayleigh scattering: Scattering off a particle that is much smaller then the wavelength.
Reflectance: Fraction of reflected energy to incident energy.
Reflection coefficient: Fraction of reflected pressure amplitude to initial amplitude scattering; refraction, reflection and diffraction of the wave at discontinuities at the surface of or within a medium.
Second critical angle: Angle where the refracted shear wave is at 90°.
Shear wave: Wave where the particle motion is perpendicular to the direction of wave propagation; also known as *transverse wave*.
Shear horizontal wave (SH): Shear wave vibrating parallel to an interface.
Shear vertical wave (SV): Shear wave vibrating at an oblique angle to an interface.
Side lobes: Off normal directions of local high intensity ultrasound caused by diffraction of the wave generated in a finite size transducer.
Simple shear: Shear wave whose direction of motion remains in one plane
Skew frequency: A measure of the uniformity of a transducer, related to difference in frequency from average frequency.
Slowness curves: Diagrams of the wavenumber, k, versus the angle from the boundary normal
Through transmission: Ultrasonic method whereby one transducer produces an ultrasonic wave and one transducer receives the wave.
Time of flight: Time that a wave traveled in a medium.
Tone burst generator: Driver that provides a sinusoidal pulse to the transducer for a set length of time.
Total internal reflection: The action when a wave approaches a boundary at an angle greater than the critical angle, so that all of that mode is reflected and none is refracted.
Transducer: Object that changes one type of energy to another, e.g., electrical energy to ultrasonic energy.
Transmission coefficient: Fraction of transmitted pressure amplitude to incident amplitude.
Transmittance: Fraction of transmitted energy to incident energy.
Transverse wave mode: See *Shear wave*.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Ultrasonic free boundary: Boundary with a vacuum or a rarified medium that does not support acoustic waves.
Ultrasonically hard: Having high acoustic impedances.
Ultrasonically soft: Having low acoustic impedances.
Velocity
group: Velocity of pulse envelope.
particle: Velocity of moving mass.
phase: Velocity of ultrasonic pulse.
Wave vector: See *Angular spatial frequency.*
Wavefront: Crests and valleys in wave.
Wavelength: The period of space or distance required for one complete cycle to occur, assuming fixed time.
Wavenumber: See *Angular spatial frequency.*
Wedge method: Method of creating an angled beam by a passing wave through a wedge on the surface of an object.
X-cut crystals: Piezoelectric crystals cut to create longitudinal waves.
Y-cut crystals: Piezoelectric crystals cut to create shear waves.

VARIABLES

v	Phase velocity
v_g	Group velocity
t	Time
τ	Period
λ	Wavelength and Lamé constant
μ	Lamé constant
v	Possion's Ratio
f	Temporal frequency (linear)
χ	$1/\lambda$
ω	Temporal frequency (angular)
k	Angular spatial frequency, wave number $k = 2\pi\chi$
u	Particle displacement ($u_i = U_i e^{j(k_1 x - \omega t)}$)
$\dot{u}$	$= \partial u/\partial t$
p	Acoustic pressure $p = Pe^{j(\bar{k}\cdot\bar{r} - \omega t)}$
r	$\frac{p_r}{p_i}$ reflection coefficient
t	$\frac{p_t}{p_i}$ transmission coefficient
Z	Acoustic impedance
I	Acoustic intensity (W/m^2)
R	$\frac{I_r}{I_i}$ reflectance

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

ϕ $\frac{1}{2}$ angle of divergence

T $\frac{I_t}{I_i}$ transmittance $(R + T = 1)$

θ_i Incident angle measure from normal

ε Strain

σ Stress

E Young's modulus

ρ Density

γ Engineering strain $(\gamma = 2\varepsilon)$

REFERENCES

1. K Graff. A History of Ultrasound. In: RE Sharpe, ed. Physical Acoustics. Vol. XV. New York: Academic Press, 1981, pp 1–97.
2. RB Lindsay. The Story of Acoustics. J Acoust Soc Am 39:630–638, 1966.
3. AD Pierce. Acoustics: An Introduction to Its Physical Principles and Applications. New York: McGraw-Hill, 1981, pp. 1–6.
4. AS Birks, RE Green Jr. Ultrasonic Testing, Vol. 7. In: P McIntire, ed. Nondestructive Testing Handbook. 2nd ed. U.S. ASNT, 1991.
5. JL Rose. Ultrasonic Waves in Solid Media. New York: Cambridge University Press, 1999.
6. B Banks, GE Oldfield, H Rawding. Ultrasonic Flaw Detection in Metals. Engelwood Cliffs, NJ: Prentice-Hall, 1962, pp 1–8.
7. Lord Rayleigh. On the Theory of Sound. Macmillan, 1877.
8. Lord Rayleigh. On waves propagated along the plane surfaces of an elastic solid. Proc London Math Society 17:4–11, 1855.
9. H Lamb. On Waves in an Elastic Plate. Proc of the Royal Soc (London) Ser. A 93:114, 1917.
10. S Sokolov. Means for indicating flaws in materials. U.S. Patent 2,164,125, 1937.
11. F Firestone. The supersonic reflectoscope, an instrument for inspection of the interior of solid parts by means of sound waves. J Acoust Soc Am 17(3):287–299, 1945.
12. F Firestone, I Fredrick. Refinements in supersonic reflectoscopy. J Acoust Soc Am 18(1):200–211, 1946.
13. F Firestone. Flaw detecting Device and Measuring Instrument. U.S. Patent 2,280,266, 1942.
14. I Asimov. Understanding Physics Vol 1. New York: Walker and Company, 1966, p 148.
15. H Benson, University Physics. New York: John Wiley & Sons, 1991, p 320.
16. RP Feynman, TB Leighton, M Sands. The Feynman Lectures on Physics. Vol. I Reading, MA: Addison-Wesley, 1963, pp 47–53.
17. RP Feynman, TB Leighton, M Sands. The Feynman Lectures on Physics. Vol. I Reading, MA: Addison-Wesley, 1963, pp 47–54.
18. E Hecht. Optics. 2nd ed. Reading, MA: Addison-Wesley, 1990, pp 19–21.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

19. B Auld. Acoustic Fields and Waves in Solids. Vol I. New York: John Wiley & Sons, 1973, pp 1–50.
20. BB Muvdi, JW McNabb. Engineering Mechanics of Materials. New York: Macmillan, 1980, p 165
21. K Uno Ingard. Theoretical Acoustics. New York: McGraw-Hill, 1968, pp 257–270.
22. J Krautkramer, H Krautkramer. Ultrasonic Testing of Materials. 4th ed. Berlin: Springer-Verlag, 1990, pp 24–30.
23. A Lutsch, G Kuhn. Generation and reception of elastic waves via a plan boundary. J Acoust Soc Am 36:428–436, 1964.
24. J Krautkramer, H Krautkramer. Ultrasonic Testing of Materials. 4th ed. Berlin: Springer-Verlag, 1990, p 109.
25. WH Buchsbaum. Buchsbaum's Complete Handbook of Practical Electronic Reference Data. 2nd ed. New Jersey: Prentice-Hall, 1973, p 5.
26. B Carlin. Sound, Noise, and Ultrasonics. In: T Baumeister, ed. Mark's Standard Handbook for Mechanical Engineers. 8th ed. New York: McGraw Hill, 1980, pp 12–137.
27. F Tse, IE Morse. Measurement and Instrumentation in Engineering. New York: Marcel Dekker, Inc. 1989, pp 207–208.
28. JL Rose. Ultrasonic Waves in Solid Media. New York: Cambridge University Press, 1999.
29. I. A. Viktorov. Rayleigh and Lamb Waves. New York: Plenum Press, 1967.
30. B Auld. Acoustic Fields and Waves in Solids. Vol II. New York: John Wiley & Sons, 1973, pp 88–93.
31. IA Viktorov. Rayleigh and Lamb Waves. New York: Plenum Press, 1967
32. B Auld. Acoustic Fields and Waves in Solids. Vol II. New York: John Wiley & Sons, 1973, pp 103–104.
33. AEH Love. A Treatise on the Mathemathical Theory of Elastisity. 4th ed. New York: Dover Publications, 1944.
34. CM Fortunko. Ultrasonic Testing Equipment. In: P McIntire, ed. Nondestructive Testing Handbook. 3rd ed. U.S., ASNT, 1991, pp 119–120.
35. HM Frost. Electromatic-Ultrasonic Transducers: Principles, Practice, and Applications. In: WP Mason and RN Thurston eds.. Physical Acoustics, Vol. XIV. New York: Academic Press, 1979, pp 179–270.
36. ER Dodds. Electromagnetic Generation of Ultrasound. In: RS Sharpe, ed. Research Techniques in Nondestructive Testing. Vol. II. London: Academic Press, 1973, pp 419–441.
37. J Monchalin, G Alers, G Huebschen, G Khuri-Yakub, B Maxfiield, W Repplinger, H Salzburger, R Smith, R Thompson, J Wagner, A Wilbrand. Other Ultrasonic Techniques; Electromagnetic Acoustic Techniques. In: P McIntire, ed. Nondestructive Testing Handbook. 3rd ed. United States, ASNT, 1991, pp. 326–340.
38. JW Wagner. Optical Detection of Ultrasound. In W.P Mason and R. N. Thurston, eds. Physical Acoustics Vol. XIX. New York: Academic Press,1990, pp 201–266.
39. CB Scruby and LE Drain. Laser Ultrasonics: Techniques and Applications. Bristol: Adam Hilger,1990.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

40. RE Schramm, AV Clark Jr., DV Mitrakovic, PJ Shull. Flaw Detection in Railroad Wheels Using Rayleigh-Wave EMATs. In: DO Thompson, DE Chimenti, eds. Review of Progress in Quantitative Nondestructive Evaluation, Vol. 7B. New York: Plenum Press, 1988, pp 1661–1668.
41. RE Schramm, PJ Shull, AV Clark Jr., DV Mitrakovic. EMAT examination for cracks in railroad wheel treads. In: HLM Does Rios, ed. Nondestructive Testing and Evaluation for Manufacturing and Construction. Conference proceedings, Urbana, Illinois, August 9–12, 1988. New York: Hemisphere, 1990, pp 373–380.
42. RE Schramm, PJ Shull, AV Clark Jr., DV Mitrakovic. EMAT for Roll-By Crack Inspection of Railroad Wheels. In: DO Thompson, DE Chimenti, eds. Review of Progress in Quantitative Nondestructive Evaluation, Vol. 8A. New York: Plenum Press, 1989, pp 1083–1089.
43. RE Schramm, PJ Shull, AV Clark Jr., DV Mitrakovic. Crack Inspection of Railroad Wheel Treads by EMATs. In: P Holler, V Hauk, G Dobmann, CO Ruud, RE Green, eds. Nondestructive Characterization of Materials. Proceedings of the 3RD International Symposium, Saarbrucken, FRG, October 3–6, 1988. New York: Springer-Verlag, 1989, pp 421–428.
44. Spotting the Defects in Railroad Wheels. Business Week (September 28, 1974), p 44.
45. JV Cowan, GD Cowan, JG Cowan. U.S. Patent 3,812,708, May 28, 1974.
46. HJ Salzburger, W Repplinger. Ultrasonics international 83, conference organizers Z Novak and SL Bailey. Kent, England: Butterworth & Co., 1983, pp 8–12.
47. HJ Salzburger, W Repplinger, W Schmidt. Deutsche Gesellschaft für zerostorungsfreie Prüfung. Jahrestagung 25, Berichtsband 10, Teil 1, pp 51–60, 1987.
48. RE Schramm, AV Clark Jr., DV Mitrakovic, Y Cohen, PJ Shull, SR Schaps. Tread Crack Detection in Railroad Wheels: An Ultrasonic System Using EMATs. In: NISTIR 3967, Report No. 22. Washington, DC: U.S. Department of Commerce, May, 1991.
49. IR Frielinghaus. Ultrasonic Testing of Shafts and Axles. [Application notes.] Krautkramer Branson, [n.d.]. SD 276 (72) Cr.
50. J Krautkramer, H Krautkramer. Ultrasonic Testing of Materials. 4th ed. Berlin: Springer-Verlag, 1990, p 561.
51. J Krautkramer, H Krautkramer. Ultrasonic Testing of Materials. 4th ed. Berlin: Springer-Verlag, 1990, p 16.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

4

Magnetic Particle

Arthur Lindgren*
Magnaflux, Glenview, Illinois

Peter J. Shull
The Pennsylvania State University, Altoona, Pennsylvania

Kathy Joseph
The Pennsylvania State University, University Park, Pennsylvania

Donald Hagemaier
The Boeing Company, Long Beach, California

4.1 INTRODUCTION

If you are working with a ferromagnetic material, with possible defects at or near the surface, *Magnetic Particle Inspection* (MPI) is one of the most economical NDE methods to use.

Once you have decided to use MPI, you will have to make three further decisions:

- Which magnetizing method to use
- Which type of current to use
- Whether to use wet or dry particles

This chapter explains the factors that influence these decisions (1–8).

* Retired.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

4.1.1 Technique Overview

MPI is a very simple method: magnetize the sample, and simultaneously (typically) flow finely divided ferromagnetic particles over the surface. Any defects in the material will affect the magnetic field in the sample, and thus can attract magnetic particles to the edges of the defects. This means that magnetic particles can outline surface and near surface defects, and therefore can be used as an indicator for such defects.

Since the sample must be magnetized, MPI is limited to parts that are easy to magnetize (i.e., have high relative permeability of 100 or more). MPI cannot be used on nonferrous parts such as copper, brass, aluminum, titanium, austenitic stainless steel, or any of the precious metals.

4.1.2 History

As with most advances in science, Magnetic Particle Inspection was developed mostly by accident. In 1918, Major Hoke, working in the Bureau of Standards, noted that the metallic grindings from hard steel parts being ground on a magnetic chuck often formed patterns corresponding with cracks in the surface of the parts. Hoke patented his discovery, but did little to standardize the process.

Major Hoke's patent was read by Dr. DeForest from the Massachusetts Institute of Technology, who had been commissioned to determine the cause of failures in oil-well drills. DeForest developed the "circular" magnetizing concept so that defects in all directions in a part could be located. In 1929, DeForest formed a partnership with F. B. Doane of Pittsburgh Testing Laboratory and purchased Hoke's patent. Doane then developed the concept of controlled types and sizes of magnetic particles. By 1934 the partnership became the Magnaflux Corporation. The technique of suspending particles in a liquid was added in 1935 at the Wright Aeronautical Company in Paterson, New Jersey.

The first major advance in MPI came in 1941, when Robert and Joseph Switzer found that fluorescent magnetic particles can provide greater sensitivity with fewer particles. Other major advances include (a) the development of multidirectional magnetization, which reduced the number of separate tests required on a specimen, and (b) paste-on artificial defects for assuring system performance. Other than these advances, MPI has basically remained the same since the mid-1940s, although both the applications and the detection methods have been further refined.

MPI is successful and popular partly because the birth of MPI technology corresponded with a great need for nondestructive testing throughout the world. The 1930s and 1940s saw astonishing growth of a great many industrial systems whose failure could result in great tragedy. Airplanes, ships, submarines, automobiles, and nuclear-energy, to name a few, are major industrial products

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

that use MPI. In fact, it is now known that MPI was used to test various parts of the reactor and the containers of uranium fuel for the Manhattan Project.*

Today MPI, used throughout the world in near innumerable applications, has proven to be a very strong and flexible nondestructive method, and should continue to be a vital part of industry safety for many years to come.

4.1.3 Advantages/Disadvantages

Magnetic Particle Inspection provides many advantages with few disadvantages. It is generally the best and most reliable method for finding surface cracks, especially very fine and shallow ones. MPI tends to be simple to operate, and produces indications that are simple to interpret. There is almost no limitation on size or shape of the part being tested. MPI also works if the defect is filled with a foreign material, or if the sample is covered with a nonmagnetic coating such as thin coats of paint or plating. A skilled operator can detect the depth of most cracks with high accuracy. Unlike penetrants, MPI can also detect limited subsurface defects to a maximum depth of approximately 6.35 mm. MPI costs little, and lends itself to automation. Finally, it is very forgiving: minor deviations from optimal operating conditions often still produce reasonable indications.

The greatest disadvantages of MPI are (a) that it can be used only on ferromagnetic parts and (b) that it can detect only surface and near-surface cracks. In addition electrical arcing and burning may occur at the contact points during testing and the object often must be demagnetized after the process, which is time-consuming. Finally, many parts require individual handling.

Table 4.1 reviews the advantages and disadvantages of MPI.

4.2 THEORY

This section reviews the basic principles of magnetism that you need in order to understand crack detection with magnetic particles: basic magnetism; magnetic fields and their effects on materials; magnetization; hysteresis; and leakage fields.

4.2.1 Basic Magnetism

Anyone who has played with refrigerator magnets or magnetized checker pieces knows that permanent magnets either attract or repel other permanent magnets. Also, permanent magnets can attract and magnetize other pieces of metal, as can be seen with a magnet attracting an iron nail (Figure 4.1a). The ability to attract or repel other magnets or pieces of metal is referred to as the *magnetic effect*. It has been found by experiment that all bar magnets have two points of concentrated

*The Manhattan Project was a top-secret project for the development of the first atomic bomb.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

TABLE 4.1 Advantages and Disadvantages of MPI

Advantages	Disadvantages
Accurate and reliable	Requires ferromagnetic material
Simple to operate	Only detects surface-breaking and near-surface cracks
Indications are produced directly on the surface of the part	For maximum sensitivity the surface should be thoroughly cleaned and dried
Little training needed for operators	Demagnetization is often necessary
Almost no limitation on size or shape of the part being tested	Exceedingly large currents are sometimes required for the testing of very large castings and forgings
Works well through thin coatings of paint, or other nonmagnetic coverings such as plating	Can heat and burn highly finished parts at the points of electrical contact
Detects cracks filled with foreign material	Individual handling of parts for magnetization is often necessary
Provides some crack depth information	Contact with the surface is sometimes required
Low cost	Some parts require multiple inspections
Forgiving of mistakes	
Some subsurface sensitivity	
Lends itself well to automation	

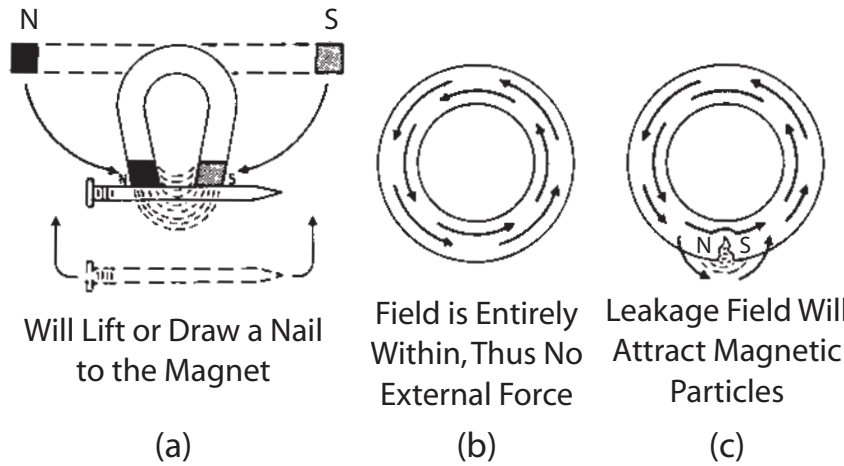


FIGURE 4.1 (a) Magnetic force attracting a ferromagnetic material (a nail). (b) Circular magnetic (no poles). (c) A small break in a circular magnet produces a leakage field beginning on a north pole and ending on a south pole.

Copyright © Marcel Dekker, Inc. All rights reserved.

magnetic effect; these points are referred to as the *north* and *south* poles of the magnet. These poles react like opposite charged particles: opposite poles attract, and like poles repel. Unlike electric charges, there is no way for an object to have a single pole. If you cut a permanent magnet into two pieces, the resulting magnets contain two poles of one-half strength. If you bend the bar magnet into a circle, then the magnet has a uniform magnetic effect and no poles (Figure 4.1b). But a small break in the circle will cause the magnetic field to leave the circle at a north pole and return at a south pole (Figure 4.1c). Therefore, it is common to state that all magnetic poles exist in pairs.

4.2.2 Magnetic Field

Consider a bar magnet surrounded by iron filings on a plate. If you give the plate a small tap to overcome friction, the iron filings will align due to their attraction to the bar magnet (Figure 4.2). Note that the filings map out lines between the north and south poles of the permanent magnet. These lines are referred to as the *magnetic field lines* (sometimes called *magnetic induction lines* or *magnetic flux density lines*) and are represented by the letter *B*. The magnitude of the magnetic field depends on the density of the lines. *B* is measured in SI units of tesla (T) and CGS units of gauss (G). Through convention, the lines outside the magnet start at the north pole and end at the south pole, and inside the magnet start at the south pole and end at the north pole. Note that the lines do not cross each other—this characteristic is common to all magnetic field lines. Finally, note that the density

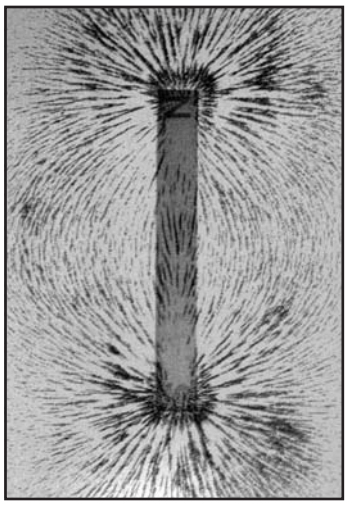


FIGURE 4.2 Iron fillings depicting the magnetic field of a bar magnet.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

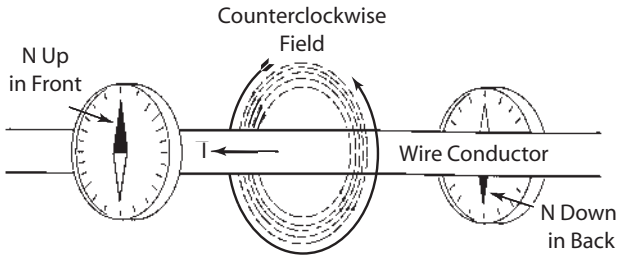


FIGURE 4.3 Direction of a magnetic field surrounding a current carrying wire indicated by the orientation of a needle on a compass.

of the lines, and thus the magnitude of the magnetic field, increases as one approaches the poles.

In 1820, a simple experiment was conducted that proved that permanent magnets are not the only things that can create magnetic field lines. In this experiment, a compass was placed in proximity to a current-carrying wire. When a strong current was induced in the wire, the compass needle deflected. No matter where the compass was placed around the wire, the needle always deflected in the same direction relative to the current (Figure 4.3). The relationship between the magnetic field and the current depends on the magnitude of the current, the orientation of the wires, and the distance to the wires; but the direction is always perpendicular to the wire and circles the wire. A common way to determine the direction of the magnetic field from a current-carrying wire is called the *right-hand rule*: If you point your right thumb along the current, your fingers are curled in the direction of the magnetic field (Figure 4.4).

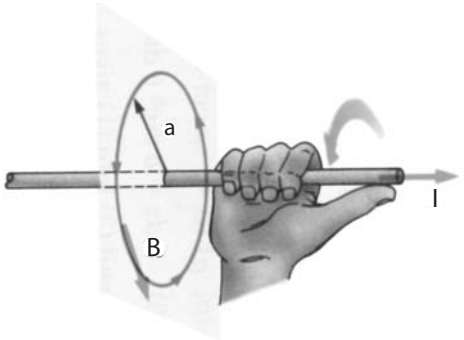


FIGURE 4.4 ''Right-hand rule'' for determining the direction of the magnetic field from a current carrying wire.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

The general equation for the magnetic field from a current-carrying wire is complicated, so consider instead these three simple cases: a long straight wire, a circular loop, and a solenoid.

As discussed above, if a current is excited in a wire, a magnetic field is induced around the wire. For a long straight wire, the magnetic field from this wire is:

$$B = \mu_0 \frac{I}{2\pi a} \tag{4.1}$$

where I is the current in the wire, a is the distance to the surface of the wire, and μ_0 is a constant called the magnetic permeability of free space. The permeability will be discussed in more detail in Section 4.2.3.

In a loop of wire, the magnetic field lines align as shown in Figure 4.5. This alignment creates a large magnetic field inside the loop. The magnitude of the magnetic field in the center of the loop is written as:

$$B = \frac{\mu_0 I}{2R}$$

where I is the current and R is the radius of the loop.

A coil of conducting wire with more than one turn is generally called a *solenoid*. When current passes through this wire, the magnetic fields look similar to stacking loops of wire. The magnetic fields inside and outside the loops add, and the field between the individual loops tend to cancel. For a long and tightly wound solenoid the magnetic field inside is nearly constant, and the magnetic field outside resembles the magnetic field from a bar magnet (Figure 4.6). For such a solenoid, the magnetic field inside becomes:

$$B = \mu_0 NI \tag{4.2}$$

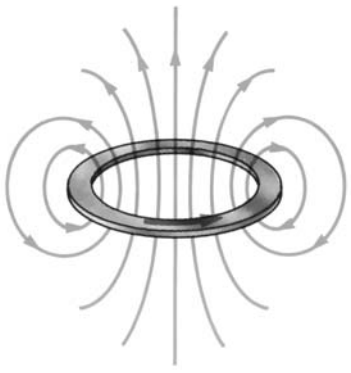


FIGURE 4.5 Magnetic field from current-carrying loop.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

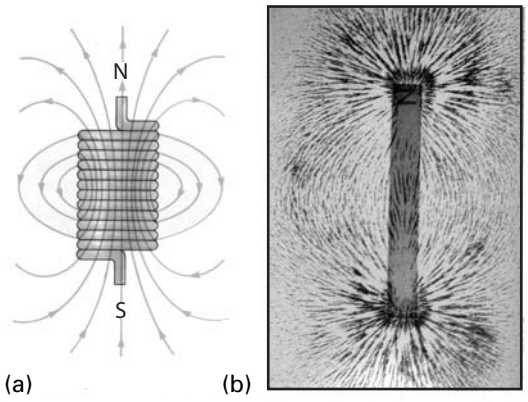


FIGURE 4.6 Comparison of the magnetic fields from (a) a solenoid and (b) a bar magnet.

where N is the number of turns per unit length of the solenoid. Note that this equation is independent of the size of the solenoid. Due to the simplicity of this equation and the ease of creating and controlling the magnitude of the magnetic field, in industry solenoids are used more frequently than permanent magnets to create magnetic fields.

4.2.3 Magnetic Fields in Materials

Magnetic Dipoles and Magnetization

A solenoid creates a magnetic field that is similar to that of a permanent magnet. This is a simple interpretation of how electrons oscillating around a nucleus can create a magnet. Consider the classical representation of an atom, with an electron orbiting a nucleus in a circle. In this case, the orbiting electron is similar to a small current in a small loop, or loops of wire. Far from the atom, this orbiting electron looks like a tiny magnet. Since they have two poles like a magnet, these "little magnets" are called *magnetic dipoles*.

The magnetic effect of the dipoles can be added like vectors, so that two in the same direction will add and two in opposite directions will cancel out. The total magnetic effect from the dipoles divided by the volume of the object is called the *magnetization* (M). Note that since the magnetization is the magnetic effect per unit volume, it does not depend on the size or shape of the object.

Since all materials have electrons, why isn't every object a magnet? The reason is that dipoles tend to align in a perpendicular direction to each other (like two magnets with north/south sides touching), so that the total magnetization is often zero.

Copyright © Marcel Dekker, Inc. All rights reserved.

Effect of External Fields on Materials

If you place an external magnet near an object, the magnetic field inside the object depends on the magnetic field due to the external field B_e and the magnetic field due to the magnetization of the object B_m, or

$$B = B_e + B_m \tag{4.3}$$

where B_m is proportional to the magnetization, M. Unfortunately, due to the definition of the dipole and thus of the magnetization M, the units of B and M are not the same. Therefore, a constant, μ_0, was defined so that $B_m = \mu_0 M$. μ_0 is called the *permeability of free space*. In SI units, $\mu_0 = 4\pi \times 10^{-7} N/A^2$, where N is newtons and A is amps. CGS units were specifically developed to make these equations simpler, so that $\mu_0 = 1$.

To simplify Eq. (4.3), a new variable H was introduced, so that $B_e = \mu_0 H$, and thus:

$$B = \mu_0(H + M) \tag{4.4}$$

The parameter H is called the *magnetic field strength*. H represents an externally applied field and is independent of whether a magnetic material is present or not. M accounts for a material's ability to be magnetized. B is the resultant magnetic field density from the combined effects of an external field and a material's degree of magnetization.

In cases where the object's magnetization is solely due to an external field, a large group of substances displays a linear relationship between H and B so that

$$H = \chi M \tag{4.5}$$

where χ is a dimensionless constant called the *magnetic susceptibility*. Combining Eqs. (4–5) results in:

$$B = \mu_0(H + M = \mu_0(\chi M + M) = \mu_0(1 + \chi)H$$
$$B = \mu_0 \mu_r H \tag{4.6}$$

Magnetic flux density = materials ability to be magnetized
× externally applied field

where μ_r is called the *relative permeability* of a substance; permeability represents the ease with which a magnetic field is created in an object. Note that Eq. (4.5) is true only if there is a linear relationship between H and M. Independent of the validity of Eq. (4.5), however, Eq. (4.6) is used to define permeability, i.e.:

$$\mu_r = \frac{B}{\mu_0 H} \tag{4.7}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Remember that μ_0 is a constant which equals $4\pi \times 10^{-7} N/A^2$ in SI units (where B is in Tesla) and is 1 in CGI units (where B is in gauss). In the laboratory, we usually measure the external field ($\mu_0 H$) by measuring the total field when the sample is removed (in air). If we measure B in air and then through the material, then $B(\text{air}) \approx B(\text{vacuum}) = \mu_0 H$, and Eq. (4.6) can be rewritten as:

$$\mu_r = \frac{B(\text{material})}{\mu_0 B(\text{air})/\mu_0} = \frac{B(\text{material})}{B(\text{air})} \tag{4.8}$$

Eq. (4.8) is correct for both SI and CGS units.

4.2.4 Types of Magnetization

Materials in general fall into three categories:

1. Paramagnetic
2. Diamagnetic
3. Ferromagnetic

Paramagnetic substances have a small positive permeability, so that an applied magnetic field, H, will create magnetism in the substance. These substances obey Eq. (4.5), so that when the external field is removed, the magnetization is zero, or when $H = 0$, $M = 0$. *Diamagnetic* materials, on the other hand, have a small negative permeability. Therefore, an applied magnetic field will create magnetism in the opposite direction in the substance. Again, removing the external field removes the magnetization.

Ferromagnetic materials are vital to MPI, as they are substances that can be made into permanent magnets. All ferromagnetic materials are composed of areas where the dipoles align together, called *domains*. In a nonmagnetized substance, the domains are randomly arranged, so that the total magnetization is zero. When an externally applied field is applied, the domains rotate to align in one direction and the substance is magnetized (Figure 4.7). Unlike paramagnetic and diamagnetic materials, however, ferromagnetic materials remain magnetized when the external field H is removed. This continued alignment after removal of the external field is called *retentivity*. Because of retentivity, Eq. (4.5) is not correct for ferromagnetic materials. But we still use the term permeability, defined as B/H (Eq. (4.6)). In this case, permeability is not constant, and care should be used with this variable, particularly when using standard permeability tables. Iron, cobalt, nickel, and gadolinium are examples of ferromagnetic materials. Figure 4.8 is a diagram of stainless and corrosion-resistant steels, distinguishing between ferromagnetic (diagrammed as "magnetic") and nonferromagnetic (diagrammed as "non-magnetic").

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

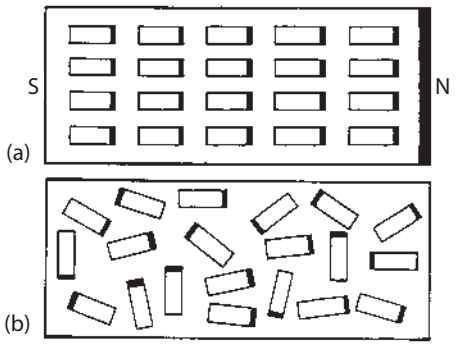


FIGURE 4.7 Magnetic domains in ferromagnetic material: (a) ordered to make a magnet and (b) randomly ordered.

4.2.5 Magnetic Hysteresis in Ferromagnetic Materials

If an external field is applied to a ferromagnetic material, the domains inside the material can align so that even when the external field is removed, the material retains a total magnetism. Therefore, if you graph B versus H before applying an external field and again after removing it, the initial point ($B = 0$, $H = 0$) is different from the final point ($B \neq 0$, $H = 0$). In other words, the magnetization of the substance depends on both the current external field and the previous external fields applied. The difference between the original path and the return path is called *hysteresis*, which literally means "lagging behind." To understand ferromagnetism, you must fully understand hysteresis.

Consider a ferromagnetic material with no magnetization, $B = 0$. If you apply a small external field H, the material will become magnetized. At this point, the process is still reversible: if you remove the external field, the internal magnetization returns to zero.

On the other hand, if you continue to increase the external field, the total magnetization increases until the maximum number of domains are aligned and the material has reached a maximum magnetization (Figure 4.9a). This is called the *saturation point*. If you now reduce the external field to zero, the ferromagnetic material retains magnetization. The amount of magnetization the material retains is called the *residual magnetism* or *retentivity* (Figure 4.9b).

If the material retains magnetism, an external field in the opposite direction is now needed to nullify the magnetization. The amount of external field needed is called the *coercive force*. If you continue to apply an external field but in the opposite direction, the material begins to be magnetized in the new direction, until it again reaches the saturation point—in the new direction (Figures 4.9c,d).

Copyright © Marcel Dekker, Inc. All rights reserved.

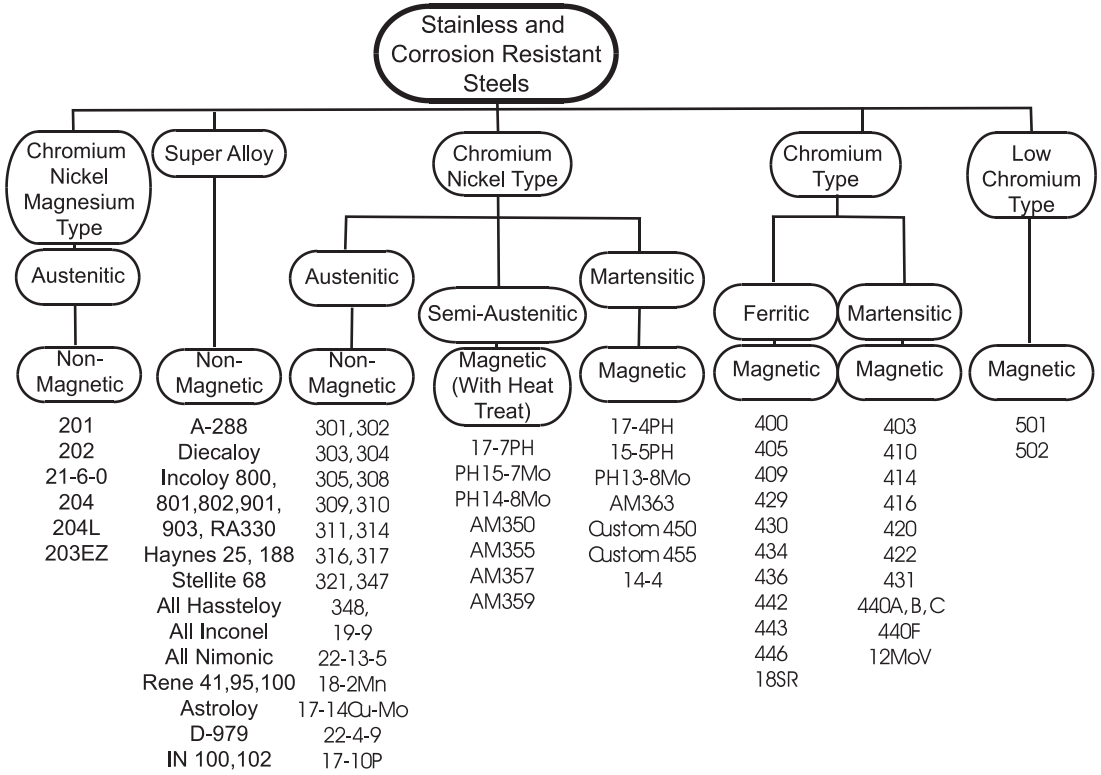


FIGURE 4.8 List of ferromagnetic (''magnetic'') and nonferromagnetic (nonmagnetic) corrosion resistant steels.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Copyright © Marcel Dekker, Inc. All rights reserved.

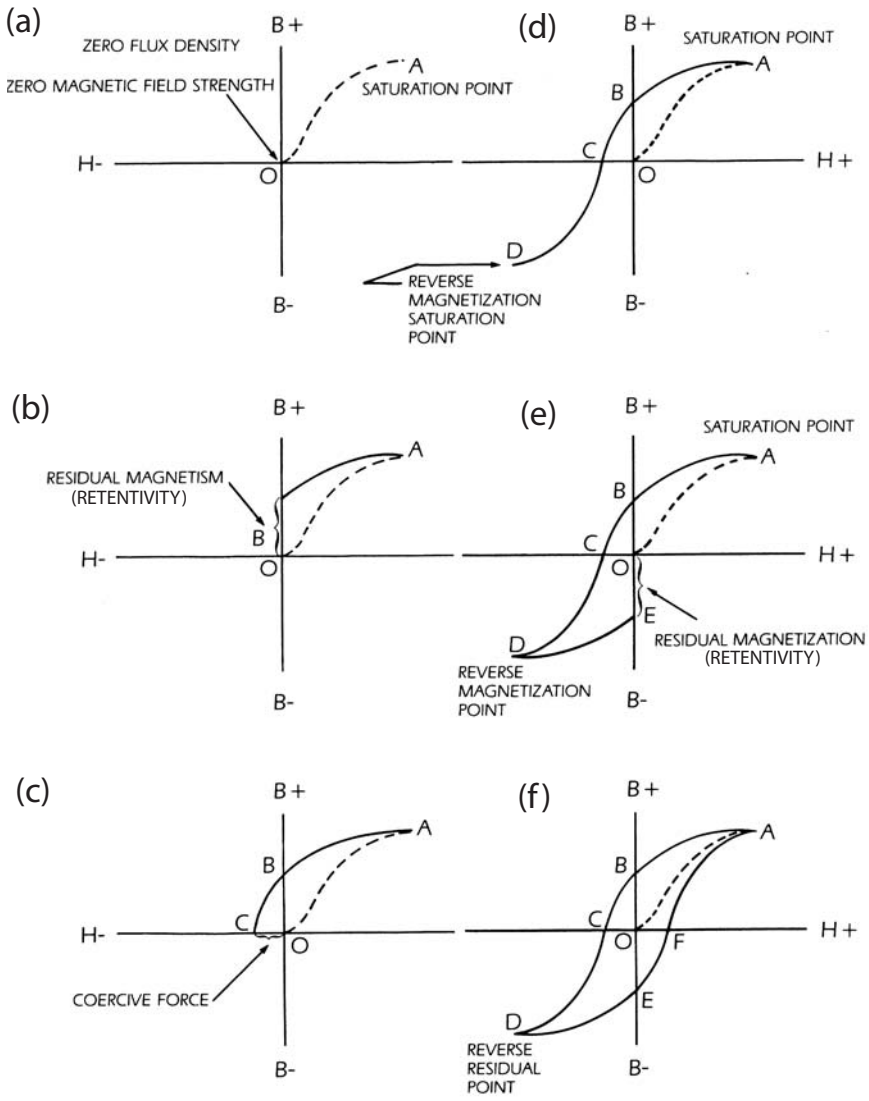


FIGURE 4.9 Diagram of hysteresis loop for ferromagnetic material.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

If you now reduce this external field (i.e., in the new, opposite direction), the ferromagnetic material again retains a magnetization even when the external field is zero. But since the external field was in an opposite direction, the residual magnetization will be negative (Figure 4.9e). A positive external field would be needed to either nullify the magnetic field or create a positive magnetic field (Figure 4.9f).

This loop—from positive saturation point to negative saturation point and back again—is called the *hysteresis loop*. The shape of the curve is a function of the material used, so that materials that have high retentivity (often called *hard ferromagnetic* materials) have wide shapes (Figure 4.10a) and materials with low retentivity (often called *soft ferromagnetic* materials) have narrow shapes (Figure 4.10b).

Let us examine the permeability in this process. Remember that relative permeability is defined from Eq. (4.7) as:

$$\mu_r = \frac{B}{\mu_0 H}$$

so the permeability is proportional to the slope of a B/H curve. Clearly, then, permeability is not constant. But when a small field is applied to a nonmagnetized

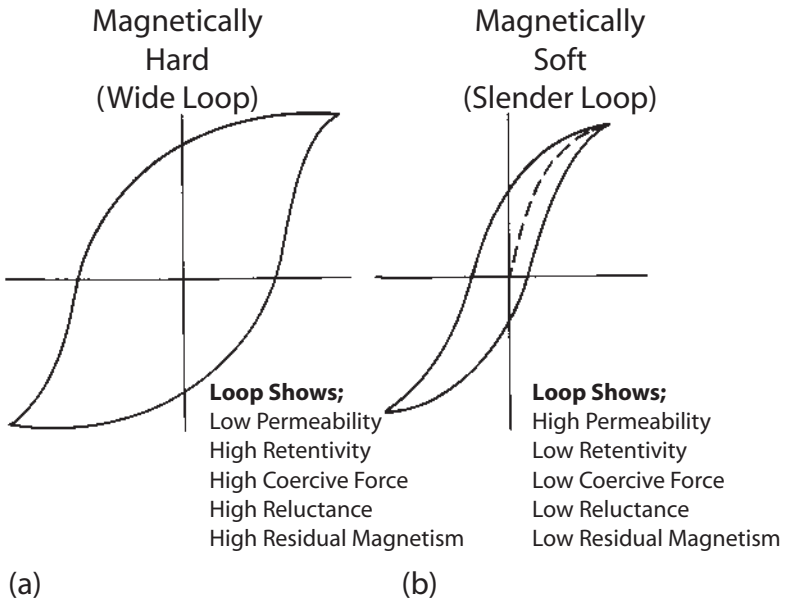


FIGURE 4.10 Hysteresis loop for (a) a magnetically hard and (b) a magnetically soft ferromagnetic material.

Copyright © Marcel Dekker, Inc. All rights reserved.

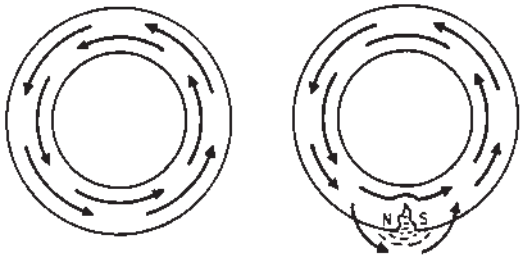


FIGURE 4.11 Magnetic field in a circular ring (no poles) and the effect of a break in the ring creating a leakage field with a north and a south pole.

sample, the material does have a small, positive permeability, similar to that of paramagnetic materials.

4.2.6 Leakage Field

Consider a circular bar with a permanent magnetic field (Figure 4.11). Since it is a circle, the bar has no poles. But if a small cut is made in the sample, a tiny north and south pole are created—meaning that an object with no externally applied field now has a small external field. This new field is called a *leakage field*.

The very small north pole on one side of the discontinuity and south pole on the other side are capable of attracting finely divided magnetic particles to form an outline of the imperfection. (Only the poles attract particles.) In effect, the particles provide a high permeability bridge for the magnetic field lines to travel across.

But a defect will create a leakage field only if the crack is perpendicular to the magnetic field. If a crack is at an arbitrary angle to the magnetic field, then its component perpendicular to the field will create a leakage field, while the other component will be undetected. In general, defects at less than 30° from the magnetic field are undetectable. The defect also must be at or near the surface of the material in order to create the leakage field. A scratch will not disturb the field, and cannot be seen in MPI (Figure 4.12).

4.3 TECHNIQUES/EQUIPMENT

The specific MPI technique and equipment to use in a given situation depends greatly on the character of the interrogated part (shape, material properties, condition of the surface, etc.), the location and type of flaws expected to be

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

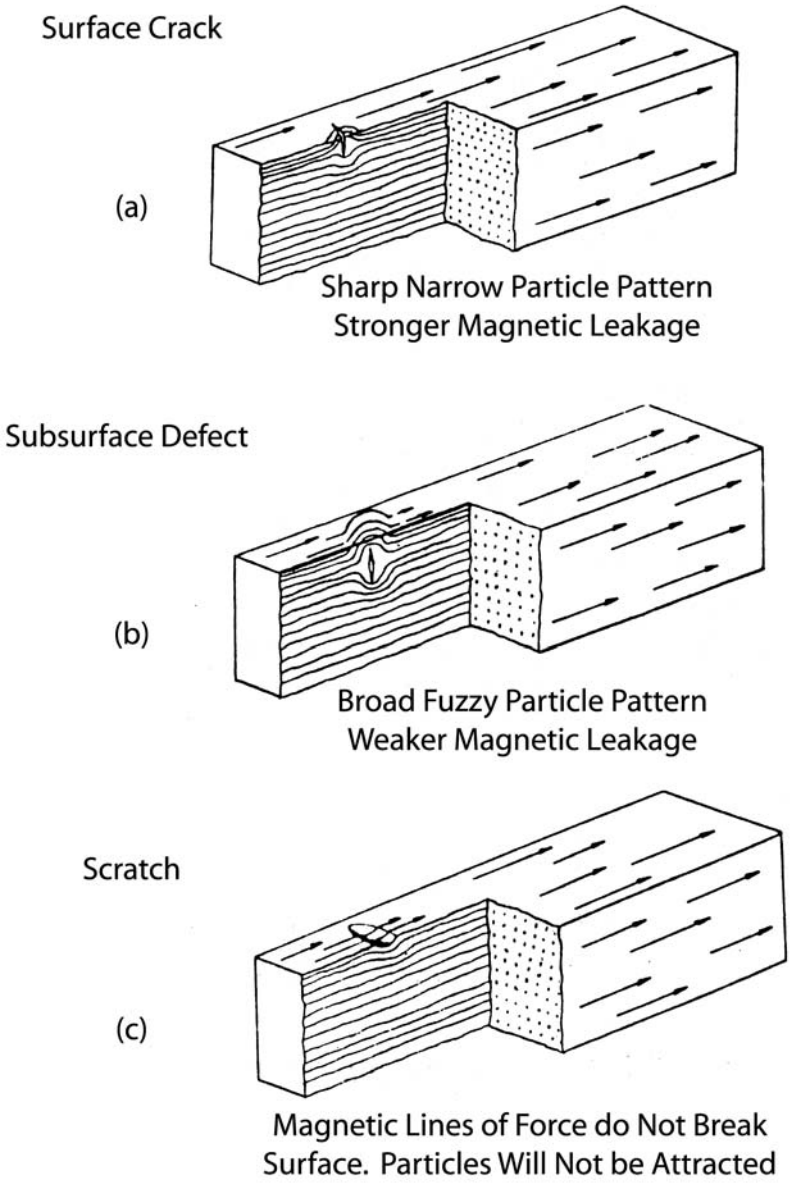


FIGURE 4.12 Diagram to demonstrate the leakage field from (a) a surface crack, (b) a subsurface defect, and (c) a scratch.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

detected (surface, subsurface, seams, inclusions, etc.), and the inspection environment (as part of a production line, maintenance within a plant, or field inspection). Although there are a host of permutations, MPI offers only four basic parameters:

1. Direction of magnetization: within the specimen, a magnetic field may be produced (a) circularly, (b) longitudinally, or (c) a combination of these two.
2. Type of magnetization current: (a) AC currents provide rapid (60 or 50 Hz) alternating changes in the direction of the magnetic field, creating high mobility of the magnetic particles, although the depth of penetration is limited; (b) DC currents (including HWDC, 1FWDC, and 3FWDC) provide high penetration but low particle mobility.
3. Method of field generation: the magnetic field may be generated by (a) a current flowing through the part (current injection), (b) an external source, such as a coil, within which the part is placed (magnetic field injection), or (c) magnetic induction (indirect current injection).*
4. Type of magnetic particles: (a) dry and (b) those suspended in a liquid (wet).

As is common in engineering, the advantages of one technique are sometimes the disadvantages of a competing technique, and vice versa; it is necessary to optimize competing effects in order to produce high-quality surface and near-surface inspections of ferromagnetic parts. The many types of magnetizing systems are confusing at first, but remember that there are really only a few methods, with many variations. The methods are summarized in Table 4.2.

4.3.1 Direction of Magnetization

In Section 4.2.6, you learned that magnetic particles are attracted only to the magnetic poles of a magnetized part. At the poles, the magnetic field leaves the part, travels through the surrounding medium (typically air) and returns into the part at the opposite pole. When a flaw creates north and south poles, the field leaving the part is called a *leakage field*. This leakage field attracts the particles and indicates a flaw.

Because cracks tend to be long and very narrow, a significant leakage field is created only when the magnetizing field is perpendicular to the crack length. In this case, the path of least resistance (reluctance) for the magnetic field is a path through the air, instead of around the flaw within the part. In order to inspect for cracks in any orientation, we use the two basic magnetization processes—

*While the terminology of current and field injection is not common in the profession, the concepts are useful in understanding the relationships among the parameters, and so these methods of magnetic field generation will be presented throughout the discussion.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

TABLE 4.2 Types of Magnetizing Methods

Name	Field created	Description	Works well on	Disadvantage
Coil/solenoid	Longitudinal	Circular wires surrounding object	Most parts	Very little subsurface sensitivity
Surface yoke (kay-jay)	Longitudinal	U-Shape "magnet" on surface	Most maintenance applications (welds, crane hooks, etc.)	Slow Little subsurface sensitivity
Connected Yoke (p-jay)	Longitudinal	U-Shape "magnet" with shape connected	Most maintenance applications (welds, crane hooks, etc.)	Slow Little subsurface sensitivity
Head shot	Circular	Current flows through sample	Most parts	Can cause burn marks
Central connector	Circular	Current flows in a nonmagnetic bar inside hollow sample	Hollow tubes Adds inside diameter sensitivity	None
Induced current Magnetization	Circular	Magnetization by magnetic induction requires no part contact	Many shapes where contact must be avoided	Requires special fixture and sometimes a core
Prods	Circular	Current flows between prods placed on sample	Large objects/ welds	Slow and can cause burn marks
Multidirectional magnetization	Circular and longitudinal	Magnetizes part in two or more directions in a single operation	Most parts	Requires special magnetizing unit
Threaded cable	Circular	Current flows in cable inside aperture in sample	Small attachment holes in sample	None

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016



Copyright © Marcel Dekker, Inc. All rights reserved.

longitudinal and circular—that create mutually perpendicular fields. In general, a longitudinally magnetic field is imposed along the part axis and will typically produce north and south poles at the ends of the part. In contrast, a circular magnetic field will generally remain within the part. Flaws interrupting either of these fields create additional poles at the site of the leakage field.

Inspecting for flaws in arbitrary directions requires at least two separate inspections. But a multidirectional magnetizing method (where fields pointing in different directions are applied alternately to the part) can decrease the number of inspections.

In practice, there are many magnetizing techniques for creating either longitudinal or circular fields in a part, depending on the shape and size of the sample. Most of the methods are just variations on a set of basic methods, enumerated below.

Longitudinal Magnetization

A longitudinal magnetic field is created in a part by either (a) magnetization from a current flowing in an external solenoidal coil, or (b) passing a magnetic field through the part (magnetic field injection) by contacting it with a yoke (Figure 4.13). In all cases, the magnetic field leaves the part. In the case of a solenoidal coil, a north and a south pole are created at opposite ends of the part. For direct field injection, the magnetic field travels through the part, entering and leaving at the legs of the yoke to complete the required magnetic field loop (Figure 4.13c)—magnetic field lines must close upon themselves.

The simplest way to create longitudinal magnetization is with a solenoidal coil (Figure 4.13a,b). According to the right hand rule (see Section 4.2.2), when a current flows in the coil, the field will be circular and perpendicular to the flow of the current, i.e., the magnetic field will be along the axis of the coil. The part is placed in the magnetic field of the interior of the coil. In this common configuration, *nonrelevant poles* are created at the ends of the part; these poles will attract magnetic particles, creating extraneous (nonrelevant) indications, which can mask flaw indications near the ends. Quick-break circuitry has been specially designed to reduce the effects of these nonrelevant poles.

Circular Magnetization

A circular magnetic field is created in a part either by (a) passing a current through the part (direct current injection), or, if the part is hollow, by (b) a current flowing in a rod threaded through the part, or by (c) magnetic induction (indirect current injection). (Magnetic induction will be discussed in a separate section.) In general, the resulting magnetic field stays within the part unless it encounters a flaw oriented perpendicular to direction of the magnetic flux.*

*Part geometry can cause the field to leave the part. For example, a part might contain an intentional narrow slit that might create a low-reluctance path outside of the part.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

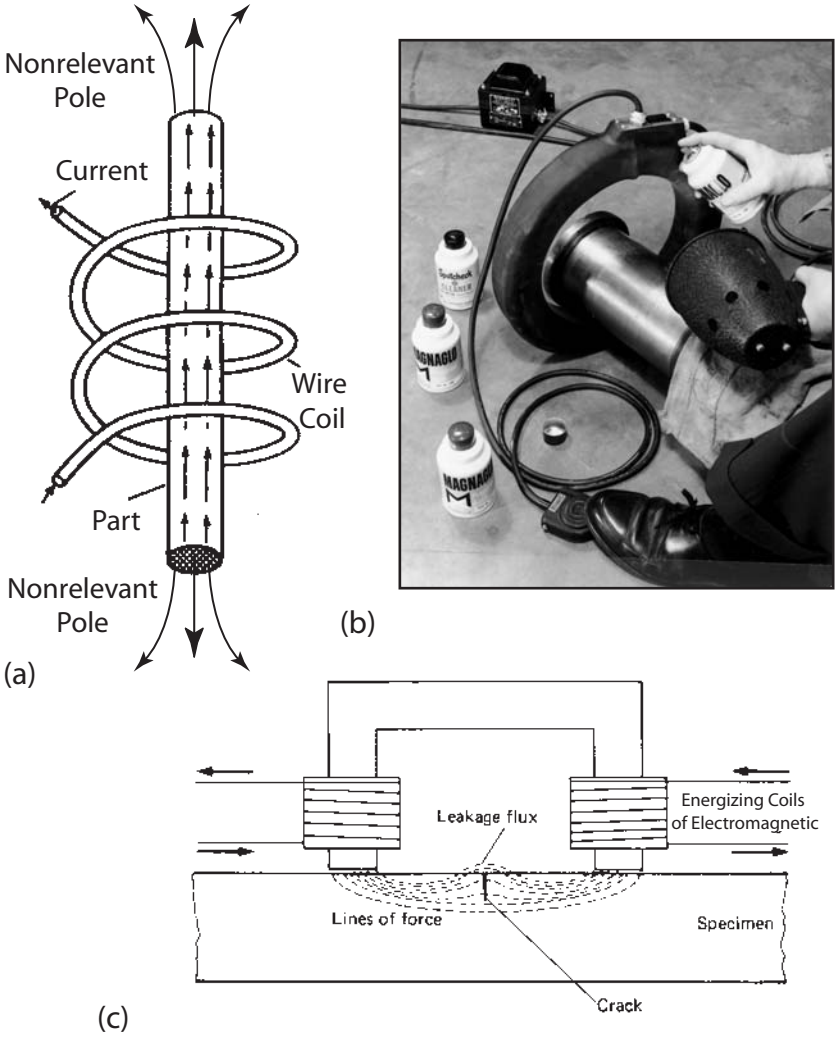


FIGURE 4.13 Methods of creating a longitudinal magnetic field for MPI. (a) A solenoidal coil; note the nonrelevant poles at the ends of the part. (b) A coil in use. (c) A yoke creating a magnetic field in a part.

The simplest way to create circular magnetization is by passing a current through the length of the specimen (Figure 4.14). According to the right hand rule, a magnetic field will be circumferential and perpendicular to the direction of current flow. This geometry means that the circular magnetic field is sensitive to

Copyright © Marcel Dekker, Inc. All rights reserved.

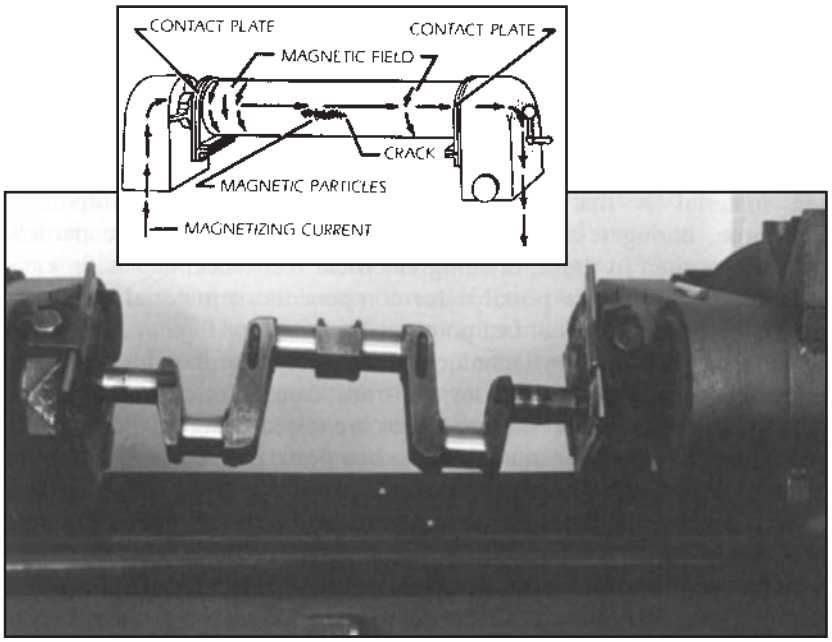


FIGURE 4.14 Diagram of a headshot passing current through a part which creates a circular magnetic field in the part.

cracks running parallel to the direction of the current flow (longitudinal cracks). Note that the sample geometry is not limited to tubular objects. Figure 4.15a shows a diagram of circular magnetization by passing current through a toroidal (ring-shaped) part. Test setup and results for a toroidal part are shown in Figure 4.15b, where longitudinal cracks are indicated. (Note the darkened background to facilitate viewing of the fluorescent indications.)

For hollow parts such as gears and short tubes, current flowing in a conductive bar that is threaded through the part opening generates a circular magnetic field within the part. The conductive bar is called a *central conductor*.

Comparison of Longitudinal and Circular Fields

The advantages of longitudinal fields are often the disadvantages of circular fields, and vice versa. For example, circular magnetic fields tend to be stronger than longitudinal fields, a clear advantage for circular magnetization. However, this implies that longitudinally generated fields are easier to demagnetize than those circularly magnetized. The advantages and disadvantages of the two fields are summarized in Table 4.3.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

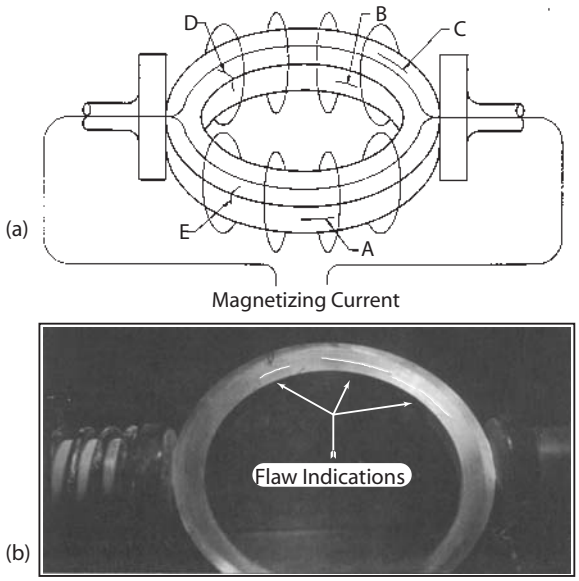


FIGURE 4.15 MPI of a circular part using a head- and tail-stock unit: (a) diagram and (b) photograph of an actual test showing flaws. Note, the darkened background to facilitate viewing of the fluorescent indication.

TABLE 4.3 Comparison of the Advantages and Disadvantages of Longitudinal and Circular Magnetic Fields in MPI

Longitudinal magnetic field	Circular magnetic field
Advantages	
Fields are easy to generate	Higher field strength than longitudinal
Method never requires electrical contact (no chance of arcing or burning)	Magnetic field is wholly contained within the part
Rapid processing of long and small parts	More penetrating than longitudinal
Easy to demagnetize part	
Disadvantages	
Lower field strength than circular	Sometimes difficult to remove circular DC fields
Commonly produces nonrelevant poles at the part ends	Electrical contact is often necessary
Very little subsurface sensitivity	

Copyright © Marcel Dekker, Inc. All rights reserved.

When choosing one field orientation over the other, remember that the maximum sensitivity occurs when the flaw is oriented in the direction of the current flow.

Multidirectional Magnetization

Multidirectional magnetization combines the directional advantages of two or more magnetic fields, in different orientations, imposed on a part in rapid succession. This method saves time, since a single test can inspect in multiple directions. But it is also less sensitive than single-direction magnetization, which allows more time for an indication to build up.

Multidirectional magnetization can be either dynamic or static. The dynamic methods combine an AC signal along one axis with either another AC signal or a DC signal oriented along another axis. Assume we have two AC fields applied to a part such that the magnetic fields are perpendicular to each other. If the two signals are 90° out of phase in time, the vector sum of the two fields will rotate through 360° (Figure 4.16a). This *rotating vector method* inspects all flaws lying in a plane perpendicular to the rotating field. If an AC and a DC field are combined, the vector will sweep out an arc similar to that of windshield wipers (Figure 4.16b). This *swinging field method* attracts particles to all flaws lying perpendicular to any point on the arc. In the static method (used mostly on very large parts such as steel castings), multiple DC fields are rapidly and sequentially applied in different directions.

An example of multidirectional magnetization is a system that combines an exterior solenoid encircling the part and a current passing through the part. The solenoid creates a longitudinal magnetic field along the long axis of the part, thus detecting the transverse cracks. The direct current creates circular fields that detect longitudinal cracks. For this process to work most effectively, the magnetization from the solenoid should be about the same as the magnetization from the direct current. Multidirectional methods are always used with the continuous wet method (see Methods of Application in Section 4.3.4).

4.3.2 Type of Magnetization Current

All MPI methods rely on a current flowing in a conductor to generate the interrogating magnetic field. Five different types of current are available: direct current (DC), alternating current (AC), half-wave direct current (HWDC), single-phase–full-wave direct current (1FWDC), and three-phase–full-wave direct current (3FWDC). The type of current used depends primarily on the depth of the defect from the surface, not the crack size. AC currents provide a highly concentrated field at the part's surface for detecting surface and very near surface defects. DC currents, which penetrate deeper, provides both moderate surface and subsurface sensitivity. Table 4.4 summarizes the defect-depth–sensitivity for the

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

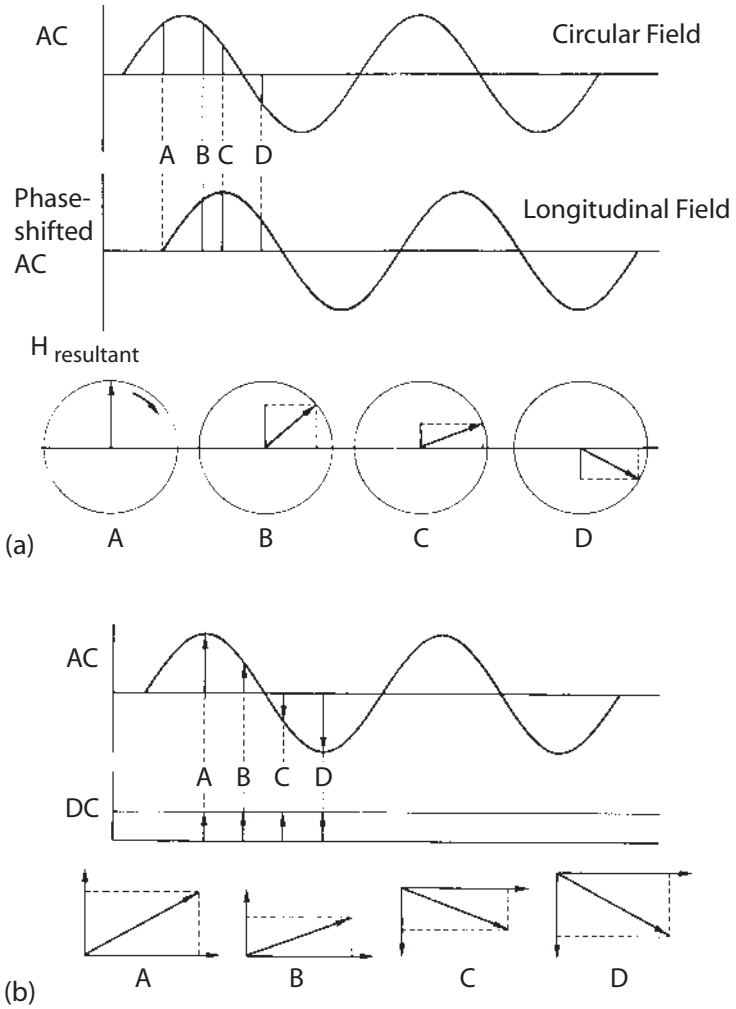


FIGURE 4.16 Multidirectional magnetization currents: (a) combining two AC signals that are out of phase by 90° (the rotating-vector method) and (b) combining an AC and a DC signal (the swing field method).

various current types. Estimates of depth-of-crack–sensitivity are given for steel with $\mu_r > 500$.

The application of the current is sometimes called *a shot*.

Copyright © Marcel Dekker, Inc. All rights reserved.

TABLE 4.4 Relative Penetration Sensitivity: Current Type Versus Wet or Dry Method (with approximate penetration values for mild steel)

	AC	HW	DC (1FWDC & 3FWDC)
Penetration	Poor	Excellent	Good
Wet method	Surface and near surface to 0.250 mm	Surface and near surface to 0.650 mm	Surface and subsurface to 1.3 mm
Dry Method	Surface and near surface to 0.250 mm	Surface and subsurface to 6.35 mm	Surface and subsurface to 1.3 mm

Direct Current

Direct current is a current that is constant over time, as in a battery. Although direct-current methods were very popular in the early days of magnetic particle testing, they are almost never used today since battery units are much too expensive to maintain.

Although shot durations vary, they typically last about 1/2 second.

Note: Many practitioners commonly, although incorrectly, refer to HWDC, 1FWDC, or 3FWDC as DC. Even though this is technically incorrect, for convenience we will also refer to these pseudo-DC techniques simply as DC, with the understanding that we are not referring to true DC.

Alternating Current

An alternating current (AC) is a current that changes direction over time. In most MPI equipment, the AC signal is approximately sinusoidal at a frequency of 60Hz (50Hz in European equipment). Because the magnetic field within the ferromagnetic material is created by the alternating electrical current, the polarity of the magnetic field also alternates. This change in magnetic field changes the polarity (north-south orientation) of a leakage field. This alternating polarity "shakes" the particles, thus increasing their mobility and their attraction to the leakage field. (This effect is particularly noted when using half wave current and dry particles for inspection.)

Besides increasing particle mobility, a time-varying magnetic field also induces eddy currents within the material. Eddy currents cause the magnetic field to decay exponentially with its depth into the part, thus limiting the interrogation

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

depth to near-surface. The depth δ is the skin depth, where the magnetic field is 37% of its maximum (surface) value, and is defined as

$$\delta = \frac{1}{\sqrt{\pi\sigma\mu f}} \tag{4.9a}$$

where δ is in meters, σ is the conductivity and is in units of 1/ohm m, f is the frequency in Hz, and μ is the permeability, which is $\mu_0\mu_r = 4\pi \times 10^{-7}\mu_r N/A^2$. A more common version of the equation for this depth is

$$\delta = 1981\frac{1}{\sqrt{\sigma\mu_r f}} = 1981\frac{\sqrt{\rho}}{\sqrt{\mu_r f}} \tag{4.9b}$$

where δ is in inches, δ is in units of 1/Ohm cm, $\rho = (1/\sigma)$ is the resistivity and is in units of Ohm cm, f is the frequency in Hz, and μ_r is the relative permeability (has no units). Note that Eqs. (4.9a) and (4.9b) are the same equation with different units. The skin depth is usually considered the maximum depth to determine defects at a set frequency. In MPI, the high permeability of ferromagnetic materials generally limits the depth of penetration for AC signals. For typical materials and frequencies, this depth is on the order of millimeters. For steels with $\mu_r > 500$, detection of defects at depths of more than 0.5 mm is considered unreliable with an alternating current/magnetic field. For more details on skin depth and eddy currents, see Chapter 5.

AC current is generally used for detecting very-near-surface or surface-breaking flaws. This type of MPI is very sensitive to such flaws, particularly when used with wet particles 1/10 the size of dry particles. (For an explanation of wet and dry particles, see Types of Particles in Section 4.3.4.)

Half-Wave Current (HWDC)

HWDC, despite its cumbersome name, is simply an AC signal passed through a rectifying diode that acts as a one-way valve for the current. Current flowing in the opposite direction is removed—i.e., *half*-wave rectification (Figure 4.17); thus HWDC can also be considered a "pulse." HWDC combines the benefits of DC and AC: high penetration (large skin-depth), as with DC; and high particle mobility (caused by the magnetic field fluctuating as the current is pulsed), as with AC.

HWDC is commonly used in combination with dry particles for multipass weld inspection locating relatively large flaws: thermal cracks, slag inclusions, root cracks, undercuts, and incomplete fusion.

Combining HWDC with the wet method reduces particle mobility. In the wet method, the flow rate of the particle suspension past the part greatly influences the particle mobility. Nonetheless HWDC, combined with the comparatively small particle size of the wet method, produces good results for near-surface and surface-breaking flaws. This combination is commonly used to

Copyright © Marcel Dekker, Inc. All rights reserved.

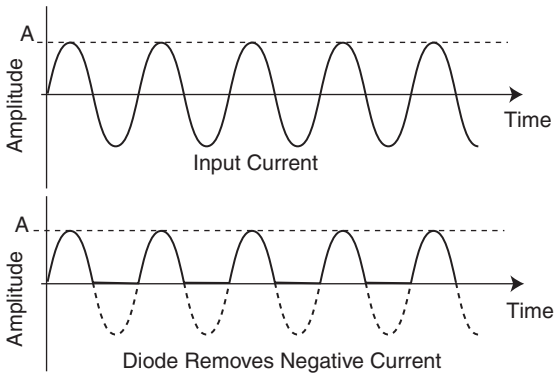


FIGURE 4.17 Diagram of a Half-Wave (Rectified) Direct Current (HWDC).

detect fine surface flaws such as forging laps, fatigue cracks, and even grinding cracks closed under compressive loads.

Single-Phase–Full-Wave Direct Current (1FWDC)

Single-phase–full-wave direct current (1FWDC) is simply full-wave rectified AC. In the half wave rectified current, the negative current is removed, leaving only half of the original current. But when an AC signal is full-wave–rectified, the polarity of negative current is reversed, creating a pseudo-positive DC signal with substantial ripple (Figure 4.18). For MPI this form provides the deep penetration of a DC signal while retaining reasonable particle mobility from the AC ripple.

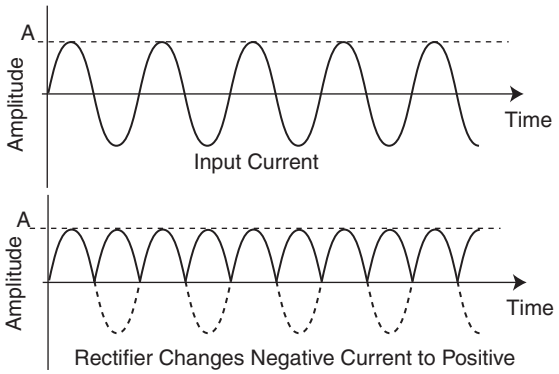


FIGURE 4.18 Diagram of a Single Phase Full-Wave (Rectified) Direct Current (1FWDC).

Copyright © Marcel Dekker, Inc. All rights reserved.

Because 1FWDC has lower particle mobility than HWDC, it is generally used with the wet method, where the particle mobility is primarily controlled by the fluid flow rate and not the form of the current. Consequently, 1FWDC is seldom used with dry particles, whose mobility depends on the form of the current.

Three-Phase–Full-Wave Direct Current (3FWDC)

Like 1FWDC, 3FWDC undergoes full-wave rectification, where the negative current is converted to a positive current. A three-phase current is composed of three single-phase lines that are progressively out of phase by 120° (Figure 4.19). 3FWDC simulates a DC signal quite well, with only a small AC ripple (about 15%). Thus 3FWDC yields significant penetration depth with low particle mobility.

3FWDC has similar sensitivity to 1FWDC. But 3FWDC works better on large parts, which require high currents (>6000 amperes) to generate the necessary level of magnetic field—too much for single-phase circuits.

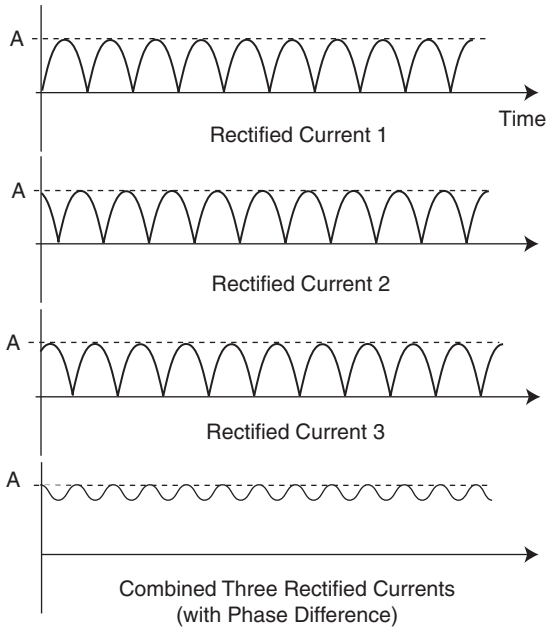


FIGURE 4.19 Diagram of a Three Phase Full-Wave (Rectified) Direct Current (3FWDC).

Copyright © Marcel Dekker, Inc. All rights reserved.

Current Levels

In MPI it is almost always the level of current, not the actual probing magnetic field, that is specified for a given setup and part combination. Although the magnetic field is the important parameter, and formulas relating current to field for a specific measurement are approximate at best, current is easier to measure. Even with the development of accurate magnetic-field meters, exact determination of the field is difficult. But thanks to paste-on artificial flaws, approximating the required field via the current is often adequate. Simply make a ballpark guess at the current, then fine tune the current by directly observing the indications on the artificial flaw. This section estimates current level for two methods for magnetizing the part—a field created by a coil and by passing a current through the part (head- and tailstock bed or prods). The details of the specific equipment are presented in Section 4.3.3.

Current passing through a coil (called a coil shot) creates a magnetic field running in the coil interior and parallel to the coil axis. The formula to approximate the appropriate current level to apply to a coil, when the coil has a part resting directly on the coil's inside diameter (i.e., not supported in the center of the coil), is

$$I = \frac{K}{N(L/D)} \tag{4.10}$$

where I is the current, K is a constant (45,000), L is the part length, D is the part diameter or diagonal, and N is the number of turns of the coil (typically five turns). This formula is valid only for $2 < L/D < 15$, and is independent of coil diameter, provided that the part diameter is less than $1/10$ of the coil diameter (i.e., has a coil fill factor less than 0.1). For greater fill factors, the equation must include both the coil and part diameters.

The formula states that the amount of current required is directly proportional to the part's diameter and inversely proportional to its length. In other words: the larger the diameter, the more current is required; and the longer the part, the less current is required.

It seems reasonably intuitive that for a fixed magnetic field, current would need to increase as the diameter increases. But the idea that less current is required as length increases may demand some explanation. When a part is magnetized in a coil, magnetic poles appear at the ends of the part. These poles are formed such that they decrease the magnetizing force within the part, an effect called *self-demagnetization*. Because the poles are localized to the magnet ends, their effect on the overall magnetization diminishes as the part length increases. Note, however, that after $L/D > 15$ the equation is no longer valid, and increasing the length does not continue to reduce the current required to produce a given magnetization.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

For a head- and tailstock bed, current is passed through the specimen, circularly magnetizing it (a head shot). To maximize the ratio of magnetizing field to applied current, the resultant magnetic field should be at the "knee" of the magnetic hysteresis curve—the point just prior to saturation. The rule of thumb for many years was between 12 and 40 A/mm. Today, through the use of accurate magnetic field meters and the use of artificial defects, this very wide range has been reduced to 12–20 A/mm.

4.3.3 Magnetic Field Generation Equipment

In MPI, magnetic fields are always created by a current flowing through a conductor. If the conductor is the part itself, the current is simply injected into the part, creating a magnetic field inside. If the part is placed in a separate conductive coil, one of two things can happen. Either:

A magnetic field is created within the coil, into which the part is placed, and the magnetic field must enter and leave the part.

Current flowing in the separate coil induces currents within the part (as an electromagnetic transformer) and creates a magnetic field within the part (electromagnetic induction). Here the field may or may not leave the part.

Recall that if the method of magnetic field generation requires the field to leave the part, creating nonrelevant poles, the magnetization is longitudinal; if the magnetic field does not leave the part (except at flaw locations), the magnetization is circular.

A variety of devices can produce interrogating magnetic fields from a current-carrying conductor. These devices accommodate a wide variety of specialized part restrictions and geometries, flaw sizes and orientations, and operating environments. They include coils, prods, yokes, head- and tailstock beds, central conductors, and induction rings.

Coils

Coils can produce longitudinal or circular magnetic fields, without the potentially damaging effects of electrical connections to the part. The most common coil is the fixed multiturn solenoid (Figure 4.13a,b). When the coil is energized, a large interrogating magnetic field is produced inside the solenoid. This geometry allows parts to be inspected rapidly, either by manually placing parts in the coil or by automatically feeding them through the coil.

How long should a coil be? What diameter should it have? What geometry restrictions are there on the part? Ideally, the coil would be longer than the part, so that the magnetic field remains within the part, leaving only at flaw sites and at the ends. The coil diameter would be nearly equal to the part diameter. (The ratio of

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

part diameter to coil diameter is called the fill factor.) In practice, however, machines have to be flexible enough to handle a variety of part lengths and diameters; this flexibility is more important than the marginal gains of the ideal setup. Additionally, ample space is required for particle application and inspection. This means we must compromise the ideal setup in order to gain access to the part. To increase the length of uniform field without sacrificing workability, systems sometimes employ two coils separated by a distance.

The primary geometry restriction is the ratio of part length to diameter. If the ratio is too small (on the order of 2:1) the magnetic field created by the nonrelevant poles at the ends of the part masks flaw indications. For smaller ratios, pole pieces made from ferromagnetic material can be added to the ends of the part, to increase its effective length, which increases the uniformity of the field within the part.

Coils are also made by a multiturn wrap of a flexible conductive cable around a part. When threaded through a hollow objects, these coils produce a circular field without electrical contact (Figure 4.20).

Prods

Prods are a pair of hand held conductors (similar to stick welding grips) connected to a pair of flexible cables placed in direct contact with the part, completing a high-current low-voltage circuit (Figure 4.21a). This direct-contact method produces a circular magnetic field within the part. The prods are constructed with contacting tips designed for efficient contacting with the part, and with handles for positioning. Maintaining the tip quality can minimize the surface damage that often results from electrical arcing to, and burning of, the part at the contact point during operation.

Prods are designed for two-person or single-person operation. In two-person operation (which is faster), the inspector manipulates one of the prods and applies the magnetic particles while a second operator (assistant) holds the second prod in position. In single-person operation, the two prods are combined on a single handle, allowing the free hand to administer the particles (Figure 4.21b). In this single-person dual-prod arrangement, the handle design allows the prod spacing to be adjusted for the particular application.

The main advantages of prods are their portability and their adaptability to complex and large test pieces. Portable power units equipped with prods are commonly used in maintenance and in field inspection.

Yokes

Yokes produce a longitudinal magnetic field. They are constructed with a solenoid coil wrapped around a high-magnetic permeable material, typically shaped in the form of a **U** (Figure 4.22a). A magnetic field is generated within the **U** by energizing the coil. Both feet of the magnetized **U** are placed directly onto

Copyright © Marcel Dekker, Inc. All rights reserved.

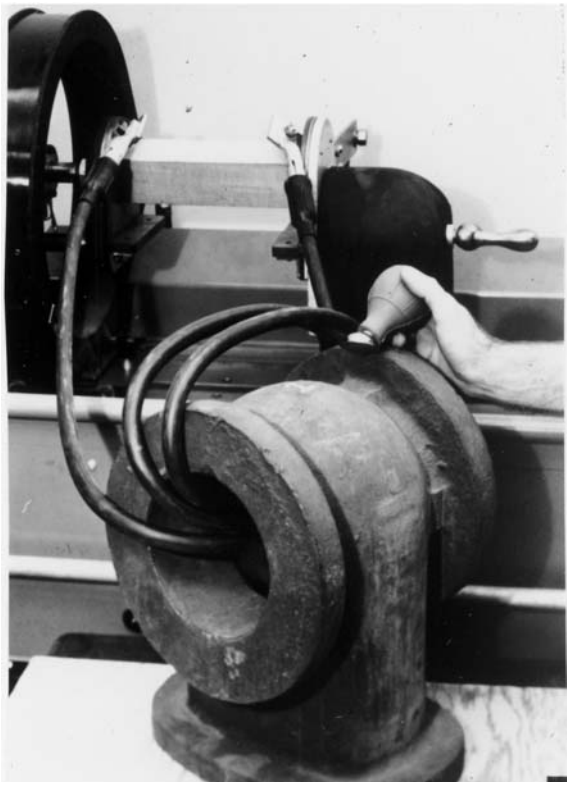


FIGURE 4.20 Testing of a hollow part using a flexible cable attached to a head- and tailstock unit. The part is circularly magnetized.

the part, creating a low-reluctance return path for the magnetic field through the part. Although yokes require, for maximum sensitivity, direct contact with the part, surface damage does not occur, since no current flows through the part.

Yokes owe their popularity to their flexibility and portability. They operate with either AC or DC currents, allowing for sensitivity to both surface and near-surface defects. To accommodate different part geometries, many yokes have articulating legs (Figure 4.22b). The high portability of yokes makes them ideal for maintenance and field inspections.

Head- and Tailstock Beds (Wet Horizontal MP Units)

Operating on the same principles as prods, the head- and tailstock bed passes a current through the part (head shot), producing a circular magnetic field within

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

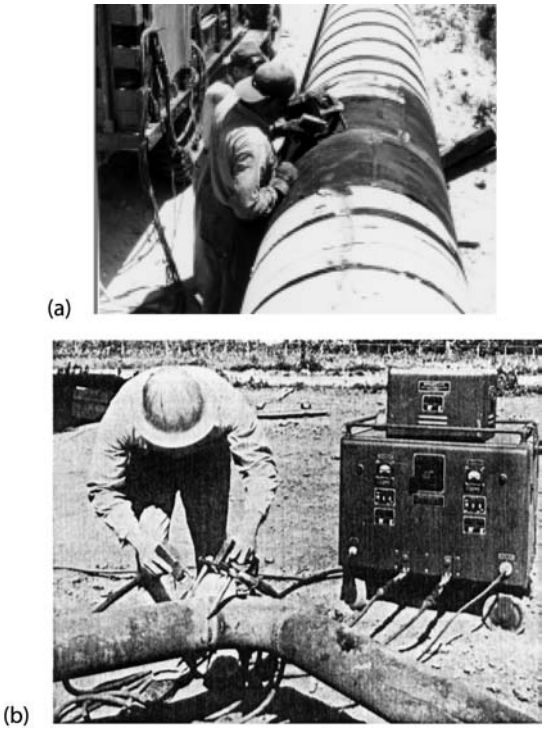


FIGURE 4.21 MPI using (a) two technicians on a pipeline weld and (b) one technician operating duel prods on a refinery weld.

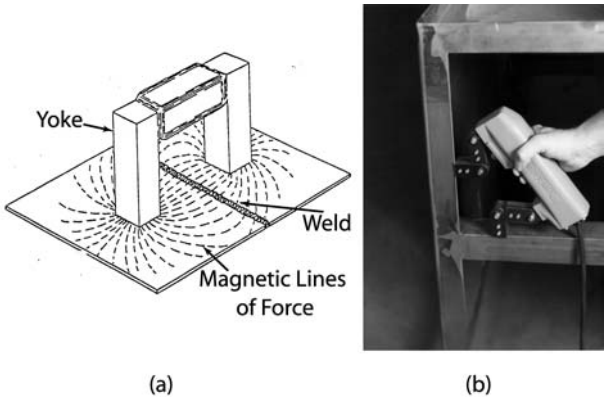


FIGURE 4.22 (a) Diagram of yoke and (b) photograph of articulated yoke.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

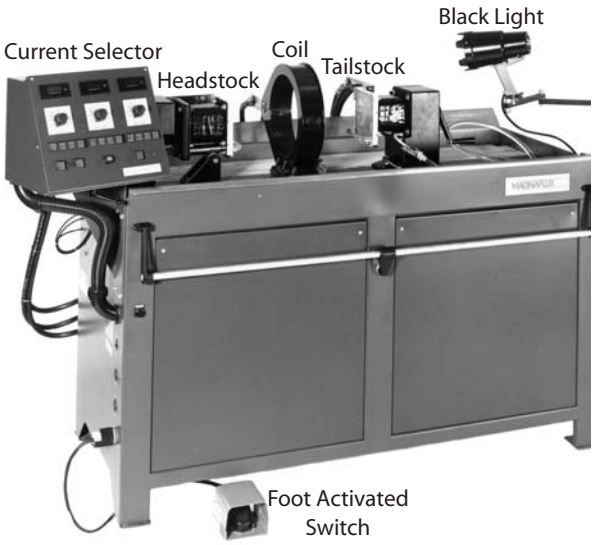


FIGURE 4.23 General purpose multidirectional magnetizing unit with a coil and a headstock/tailstock arrangement—wet horizontal unit.

the part (Figure 4.23). (In the industry, head- and tailstock beds are called *wet horizontal MP units*.) The head- and tailstock act as the current contact points, and the part is clamped between the two, completing the electrical circuit. Arcing and burning are common, although not as common as with prods. To reduce these problems, the contacts are made of replaceable lead sheets.

Most head- and tailstock beds are equipped with a 12- or 16-in. coil for use with the wet method, and with a current selector switch. This combination is common on production lines because many parts require both longitudinal and circular magnetization. For head shots, a rail-mounted coil can be positioned out of the way in a recess at the headstock. By adding a timing circuit that sequentially activates circular and longitudinal magnetization, wet horizontal MPI machines become multidirectional magnetization inspection systems. These nonportable units are commonly used in production lines, and they facilitate the use of fluorescent wet particle inspection using the continuous method.

Central Conductors

Central-conductor systems thread a current-carrying conductor through hollow parts to create a circular magnetizing field within the part. Figure 4.24 shows a common setup, using a head- and tailstock bed with a central conductor placed between the contacts.

Copyright © Marcel Dekker, Inc. All rights reserved.

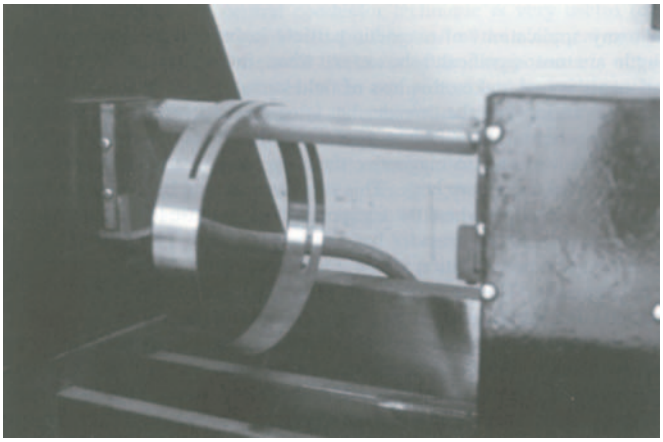


FIGURE 4.24 A ring-shaped object inspected on a head- and tailstock unit using a central conductor.

MPI using central conductors ensures magnetization of both the interior and the exterior of the part in the circumferential direction, ensuring total inspection for longitudinal cracks except where the central conductor might contact the part. Because no current passes through the part, there is no potential for surface damage from arcing or burning.

A variation of the central conductor (Figure 4.25) threads a flexible central conductor through an opening to inspect for transverse and radial cracks.

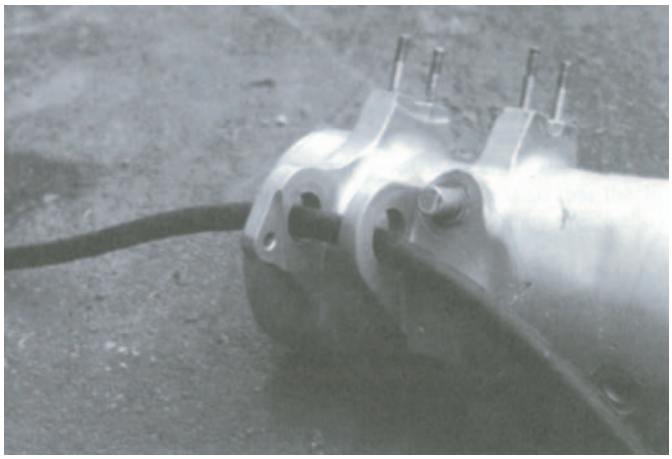


FIGURE 4.25 Mounting brackets inspected with a flexible central conductor.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Induction Rings

Induction rings produce a magnetic field by electromagnetic induction. To understand induction, consider a conductive circular part lying concentric to an encircling coil (Figure 4.26). Energizing the encircling coil produces a primary magnetic field that is perpendicular and circumferential to the current flow direction. When the current is abruptly stopped (or changed), an induced current appears in the conductive part. This current produces a secondary magnetic field within the part. The directional relationships among the energizing and induced currents, and the primary and secondary magnetic fields are detailed in Figure 4.26. This phenomenon is a result of electromagnetic induction and Lenz's Law and detailed in Section 5.2 of Chapter 5.

The induction method produces circular magnetization of the part. To improve the induction efficiency and thus sensitivity, a central core is added. Typical cores should have the following properties:

They should be made of low retentivity, high permeability steel.
They should be laminated, minimizing the flow of eddy currents within the core itself—eddy currents would slow the collapse of the applied field.
They should be long enough to make the L/D ratio of the core-part combination at least 3.

Figure 4.27 shows automated inspection of bearing races using magnetic particle induction rings with a central core.

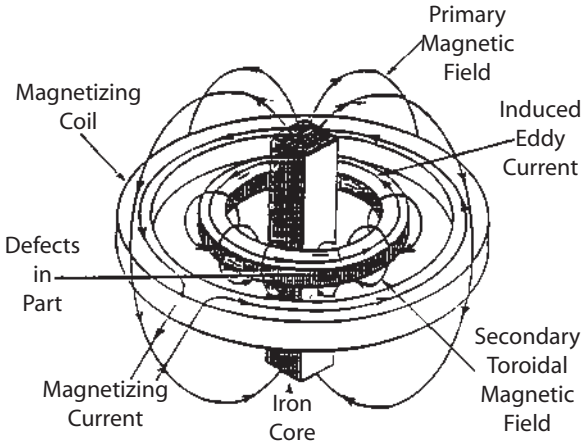


FIGURE 4.26 Diagram of MPI magnetic induction method. The diagram shows the primary (applied) and secondary (induced) currents and magnetic fields.

Copyright © Marcel Dekker, Inc. All rights reserved.



FIGURE 4.27 MPI assembly line inspection using magnetic induction method.

Quick-Break Circuitry

Quick-break circuitry employs electromagnetic induction fields to minimize leaky fields. What are leaky fields? Current levels in Section 4.3.2 introduced self-demagnetization for longitudinal fields, where the magnetizing force on a part decreases near the poles. We recall that the flux lines must close in upon themselves, following the path of lowest reluctance. Ideally, the magnetic field would travel entirely within the high-permeability (low-reluctance) ferromagnetic material until the flux reached the end of the part, creating a pole. Then, once outside the part, the magnetic flux would loop back and close in upon itself at the opposite pole at the other end of the part.

In reality, however, the magnetic field leaks from the part all along the magnet. This leaky field is very small near the center of the length of the magnet, increases slowly in the direction of the poles, and finally increases rapidly very near the poles (Figure 4.28a). Note that the magnetic flux always leaves the part *normal* to the part surface, even at the poles. In MPI inspection, a *leaky* field attracts the magnetic particles just as the leakage field (from a flaw) does. Near the poles, then, the strength of the leaky field will cause particles to mask small flaw indications. (Note this effect becomes important when the leaky field (normal field) exceeds five times the leakage field (tangential field).)

The quick-break circuit rapidly removes the applied current to the energizing coil, and therefore rapidly collapses the (primary) magnetic field. (Remember

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

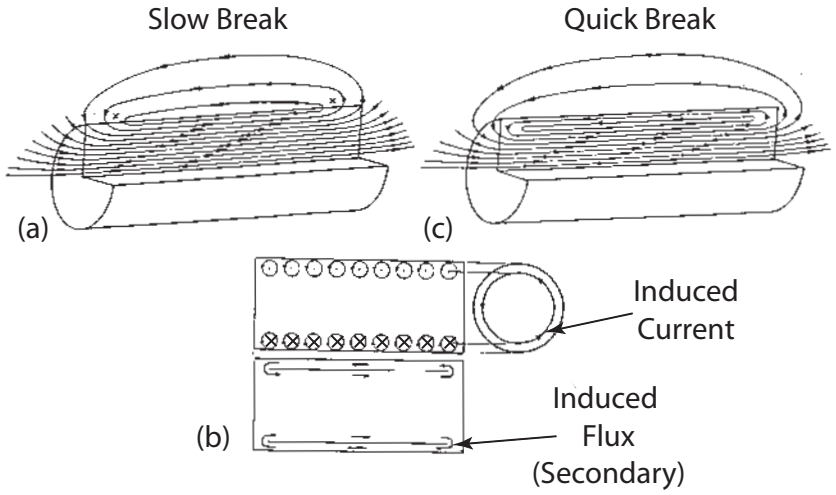


FIGURE 4.28 Diagram of how quick-break circuitry reduces the affect of nonrelevant poles on the surface of a cylindrical part. (a) Magnetic field without quick-break circuitry, (b) induced surface currents, and (c) magnetic field with quick-break circuitry.

that this device works only with longitudinal fields excited by a coil.) If the primary field changes rapidly, currents are induced in the surface of the conductive part, circumferential and perpendicular to the primary field (Figure 4.28b). This circumferential current, flowing in the part, produces a secondary magnetic field that flows axially along the part. The total field is the sum of the primary and the secondary magnetic fields. The result of the quick-break circuit is shown in Figure 4.28c, where the leaky field has been dramatically reduced, particularly near the poles. This method can be thought of as creating a current sheet on the part perimeter that shields and thus keeps the magnetic field within the part.

Note that although we have used terms that might indicate circular geometries, quick-break circuitry is not restricted to cylinders. Also note that quick-break methods slightly decrease the residual magnetic fields at the poles of the part.

4.3.4 Magnetic Particles

At the heart of MPI are the fine magnetic particles that create an indication at a leakage field caused by a flaw in the magnetized sample. The parameters surrounding these particles are

Copyright © Marcel Dekker, Inc. All rights reserved.

1. Type of particle (dry or wet)
2. Method of application (continuous or residual)
3. Viewing method ("color contrast" using ambient light, or fluorescent)

Your choices will depend on the type of magnetization current, equipment used, working environment, and part and flaw specifics.

Types of Particles

MPI uses chemically treated iron oxide particles, small in size and varying in shape. Their small size increases their mobility, enabling them to form an indication at the flaw leakage field. Because the particles have different sizes (all of them are small, but some are smaller than others), they can detect large as well as small defects. It's not a good idea to attempt to recover and reuse dry particles: the smaller ones are usually lost in the attempt, and with them the sensitivity to small flaws.

Not only their size but their varying shapes help particles detect flaws. When particles are attracted to the site of a leakage field, they pile up, forming an indication. If the particles were all spherical, the pile would collapse like a stack of bowling balls when the force was removed. But if the pile is made up of different sizes and shapes of particles, they will interlock and stay together—like a mixture of bowling balls and bowling shoes.

Particles are selected for their high magnetic permeability and low retentivity. Because of these factors, particles on a magnetized part are attracted to leakage fields only. (Otherwise, the particles would become permanent magnets themselves, and be attracted to each other and any ferromagnetic material—useless for locating leakage fields.)

Magnetic particles may be applied to the surface of the part either as a dry powder (dry method) or in a liquid suspension (wet method). Dry particles, which are approximately 10 times as large as wet particles, are primarily used in conjunction with portable systems such as prods and yokes to detect subsurface and medium to large surface flaws. They are typically applied using a powder spray bulb or automatic powder blower. The smaller wet particles are more mobile, thus more sensitive for detecting fine surface and near-surface flaws. They are applied in a continuous flow system or sprayed onto the part.

Magnetic particles come in a variety of colors, to contrast with the part. The most popular colors are

Dry: red, gray, and black
Wet: fluorescent

Red is the most popular (dry) choice and is commonly used on new welds, pipeline welds, and preventive maintenance application. Gray is mainly used for inspecting gray iron (the shades of gray are different, and the contrast is good).

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Dry black particles provide higher sensitivity, but the low contrast limits their use. To improve contrast, parts may be sprayed with a water-washable white paint prior to inspection.

Most MPI uses particles impregnated with fluorescent dye to enhance visibility. (Fluorescence in a darkened environment can produce a contrast of 1000:1 or higher; by contrast, it is said that the contrast between pure black and white is 25:1.) Today, an MPI inspector using fluorescent particles on an automotive or aerospace production line can view up to 500 parts per hour as they pass by on a conveyor line (Figure 4.29). Fluorescent particles come in large and small sizes for detecting large or near-surface flaws or small surface-breaking flaws, respectively.

In maintenance and field inspection, where ambient light is difficult to control, inspectors mainly use dry particles with portable systems.

Methods of Application

Magnetic particles are applied using either the *continuous method* or the *residual method*. The continuous method, used on most assembly lines, applies particles to all surfaces of the part while the magnetizing force is at its highest level. This is rather easy to accomplish when using dry particles: the inspector simply positions the yoke or prods on the part, energizes the circuit, lightly sprinkles the dry powder on the surface of interest, stops the application, and then allows the current to stop.

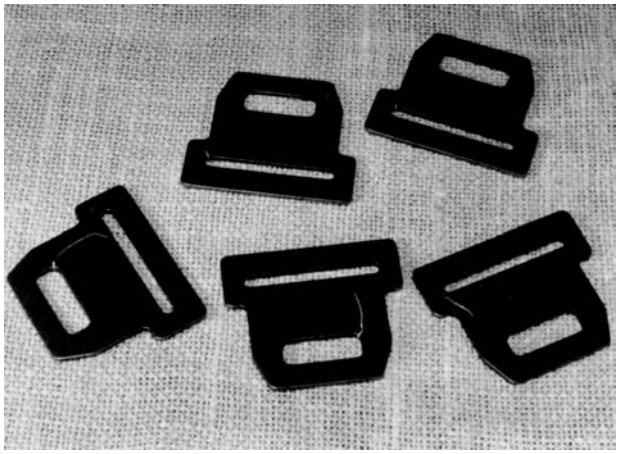


FIGURE 4.29 MPI inspection of seat belt buckles for stress cracks.

Copyright © Marcel Dekker, Inc. All rights reserved.

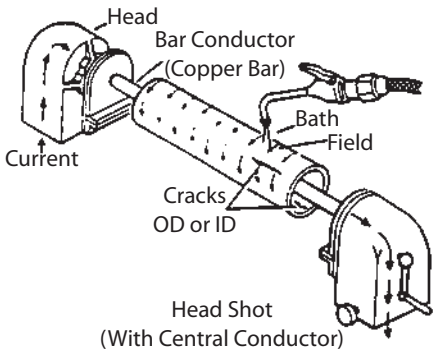


FIGURE 4.30 Application of wet particles on a part inspected with a central conductor on a head- and tailstock MPI unit.

With wet particles, the inspector's task is a bit more difficult. To assure good coverage, the inspector usually must wet the part over all surface areas first (Figure 4.30). Then, just as the nozzle is diverted, the magnetizing button is pressed. The current duration on most wet magnetic particle units is set for about 1/2 second. If the bath application is timed correctly, it is flowing on all surfaces of the part while the current is on, and has stopped flowing and is not even dripping once the current has stopped. But exact timing can be difficult with large or complex-shaped parts. To ensure coverage at the time of peak current, the inspector may press the magnetizing button several times during and after the bath application. This technique works, but it takes longer, and under certain circumstances it can overheat the part.

Fortunately, magnetic particle testing is very forgiving. Most parts inspected are retentive to some degree. Once an indication has formed, a minor flow of liquid over it will not wash it completely away. Larger flaws retain at least some of the indication formed. Allowing some bath to flow after the current has stopped is called the *sloppy continuous method*. Many parts, in many plants, are processed this way every day.

The *residual method* allows the inspector to apply the particles any time after the magnetizing force has been removed. For this method to work, the residual field retained by the part must be large enough to develop a large leakage field at the sites of all flaws of interest. Generally, the residual method is used with wet particles and is less sensitive then the continuous method. However, the simplicity of the residual method explains its common use on highly retentive materials such as bearings. Additionally, in the special case of inspection for cracks under chrome plating, the residual method can be more sensitive than the continuous method.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

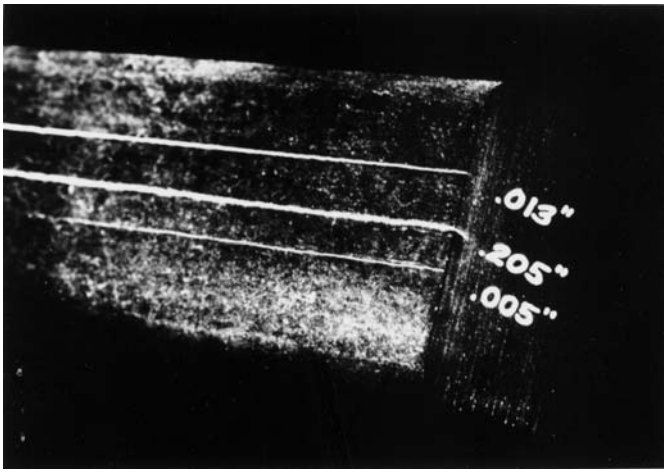


FIGURE 4.31 Brightness indication of seam depth.

Viewing Methods

Except for a few specialized systems, human operators view all indications. This feature alone creates a host of inspection issues: operator fatigue, persistence of vision, adaptation of the eye to the dark, and normal visual acuity. These issues are addressed in Chapter 2, Liquid Penetrant. Operators generally view indications in one of two ways: in visible light, or in fluorescent light.

Dry particles are viewed using visible light. Typically, the lighting is simply ambient. But the light intensity at the part should be at least 100 ft candles. Apart from ambient lighting, the color contrast between the particles and the part can be used to make indications more visible. White contrast paints are also available to increase indication visibility.

Viewing indications with (wet) fluorescent particles requires illumination from a black light source in a darkened environment.* The primary requirement of a black-light system is that it produce the proper level of illumination at the part (approximately 1000 $\mu W/cm^2$). Typically, the light intensity is monitored daily using a black light meter. Many of these meters also monitor the level of ambient visible light.

Under optimal lighting conditions an operator can determine the flaw depth with reasonable accuracy (Figure 4.31). (Note that the pileup of the particles at the defect, and thus the visual intensity of the indication, is predominately a function of the depth of the crack.)

*Black lights pose a variety of maintenance concerns (not addressed here).

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

4.3.5 Demagnetizing the Sample

After MPI, many parts retain some magnetic field which must be removed before additional processing or before placing them in service. Examples where demagnetization is necessary are

Plating, painting, or post cleaning—chips and small particles difficult to remove
Machining and grinding—clinging chips effect finish quality
Electric arc welding—the arc is magnetic and can result in poor weld quality
Moving parts (like gears)—increased friction
Magnetic sensitive instruments—reduced accuracy

(Note, even parts containing only a circular field must be demagnetized if they are to be ground, drilled, or machined.)

To understand demagnetization, consider magnetization, as described by the $B-H$ curve in Figure 4.9. In this curve, B is the magnetic field density in the object and H is proportional to the external magnetic field. When an external field is applied to a ferromagnetic material, the material retains magnetism even when the external field is removed. If, however, an external field is applied that changes direction and slowly decreases in intensity, then the hysteresis loop decreases in size to a point, and the part is demagnetized (Figure 4.32).

Using a solenoid is the simplest way to demagnetize a part. Energizing a coil with an alternating current creates an alternating magnetic field. The magnetic field from the solenoid decreases with increasing distance from the coil; when a part is slowly passed through the solenoid, the magnetic field within the part changes direction and slowly decreases in intensity. (Figure 4.33 depicts an AC demagnetization coil in operation.) 90% of all parts magnetized in MPI are demagnetized using a simple AC coil. For best results, the magnetic axis of the part should be parallel with the coil's axis. This method is not successful on large heavy parts magnetized using direct current.

Current contact methods (head- and tailstock beds) can also be used to demagnetize a part. This method requires that an alternating current pass through a sample and slowly decrease in intensity to zero. If a low (1 Hz) frequency is used, then the skin depth will be high; large samples can be demagnetized this way.

Yokes are sometimes used for demagnetization where portable equipment is necessary. The part is demagnetized by setting the yoke to create an alternating field and then placing the yoke on the part, moving the yoke, while still in contact with the part, in circles, and then moving it to an edge of the part and slowly removing it.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

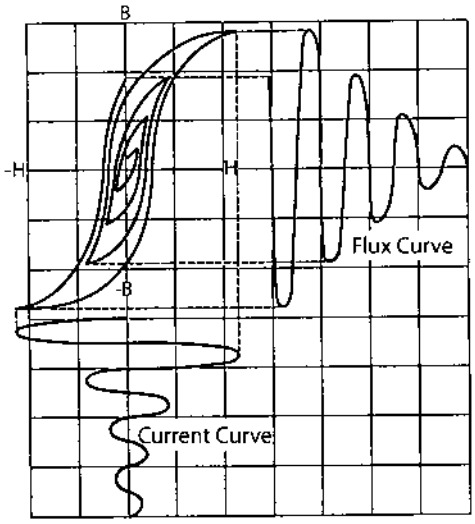


FIGURE 4.32 Diagram of the demagnetization process on hysteresis loop.

4.3.6 Condition of Parts

As with any NDE method where the flaw indication is viewed on the surface, the quality of the surface will affect the results. MPI is quite forgiving of minor surface imperfections, such as light coatings of oil, rust, or thin protective coatings. To maximize sensitivity, however, the surface should be smooth and free of dirt, oils or other solvents, and corrosion. Common cleaning methods include detergents, solvents, vapor degreasing, grit blasting, and grinding (see Chapter 2). Unlike penetrant testing, smearing over of cracks by grinding or grit blasting does not severely affect the sensitivity.

Surface roughness can create a uniform magnetic leakage field over the surface. Such a field creates a constant low-level indication over the rough surface, reducing its contrast against a flaw indication—like looking at stars during the daytime instead of at night.

Flaws can be detected under nonferrous protective coatings such as paint or cadmium plating of less than 0.075 mm. Should a circular field be required and it cannot be obtained by inducing the required current into the part, a nonconductive coating, such as paint, must be removed at the points of contact. Thin ferrous coatings, such as nickel, less than 0.025 mm will not interfere with inspection.

Copyright © Marcel Dekker, Inc. All rights reserved.

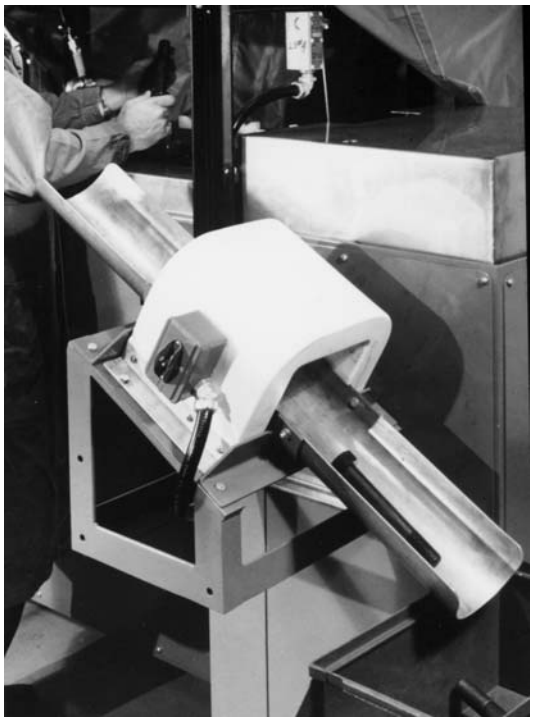


FIGURE 4.33 Demagnetization coil in operation.

4.4 INSPECTION AIDS

When MPI is used repeatedly, it may fail to consistently detect a given type of flaw. There are three ways to avoid this failure:

Controlling particle suspension
Controlling magnetization and system performance
Avoiding nonrelevant indications

The inspection aids described in this section each address one or more of these issues.

4.4.1 Controlling Particle Suspension

MPI is said to be very forgiving of less than optimum operating procedures. Even if too little current is used or the bath is not applied exactly right, the larger defects will usually be found anyway.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Still, bath maintenance is very important. If you are using the wet method with fluorescent particles (the most common method), you can miss more defects from poor bath maintenance than from using the wrong amount of current.

Bath failure has three common causes:

1. Particle concentration becomes too low or too high
2. Particles lose their fluorescence
3. A contaminant in the bath brings particles out of suspension

Changes in bath suspension are not easy to observe visually: a fluorescent particle bath always looks good under black light, since it is flowed over a part. Only a centrifuge tube reading can indicate contamination or a change in particle concentration. A centrifuge tube is a transparent settling beaker with a narrow graduated stem at the base (Figure 4.34). The centrifuge tube is filled to a specific volume from a properly agitated bath, using the nozzle used for applying particles to the part. (Passing the particles through a demagnetization coil helps separate



FIGURE 4.34 Centrifuge tube used to test particle concentration.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

clumped particles before they are placed in the stand.) After a settling period, the height and condition of the settled particles are observed. With production work, taking a reading before every eight-hour shift is a must.

Changes in Particle Concentration

Changes in particle concentration in a suspension bath can affect the particles ability to detect flaws. You can monitor changes by using a centrifuge tube:

Fill the tube to the 100 mL line with fluid from the application nozzle. *Never fill the tube directly from the bath reservoir.*
If you are performing this procedure at the beginning of the day, run the circulation pump for at least 10 minutes.
Pass the tube through a demagnetization coil and place it in the settling stand.
After a 1/2-hour (water suspension) or one-hour (oil suspension) settling period, record the particle volume.

The acceptable operating range for fluorescent particles is 0.15–0.25 mL. Below 0.10 mL, the particle concentration is too low to detect small defects; above 0.40 mL, background fluorescence may mask small defects.

Add particles or carrier fluid as needed.
Check the tube under black light for evidence of layer separation, bands, or striations.

Loss of Fluorescence

Fluorescent particles are particles impregnated or coated with a fluorescent dye at manufacture. The continued agitation of the bath suspension not only breaks down the larger particles to smaller ones, but also removes the dye from their surface (Figure 4.35). Normally this action is not a problem, since particles routinely leave the bath during inspections and are replaced by new particles each time a reading indicates low concentration. Loss of dye on the particles is indicated by seeing a separate, bright fluorescent solid top layer of particles in the stem. Heavy magnetic particles fall out first, followed by the lighter particles of pigment.

Contaminations

The most common contaminants are:

Water when using an oil suspension bath
Oil when using a water suspension bath
Excessive wetting agent used to improve part wetting

In an oil suspension bath, magnetic particles begin to agglomerate if 0.5% water is present. The clumped particles act as large particles, reducing particle

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

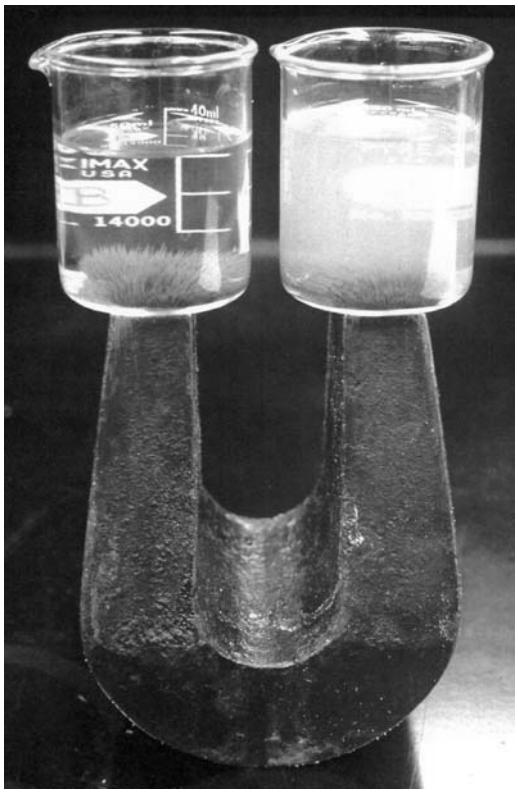


FIGURE 4.35 Photographs of particles in suspension. The beaker on the left holds magnetic particles of good quality and the beaker on the right holds magnetic particles that have separated from their fluorescent dye. The magnet has attracted all particles to the base of the beaker. Under black light, the clear suspension on the left indicates no pigment loss from the particles. The beaker on the right shows a fluorescent cloud in the liquid indicating loose, nonmagnetic dye floating in the liquid.

mobility and thus flaw sensitivity. Particle clumps also yield false concentration readings from the centrifuge tube. Worse, the clusters of particles begin coming out of suspension, float on top of the bath reservoir, and stick to the sides of the tank or to parts being inspected. Such a condition usually calls for replacing the entire bath, after cleaning the tank and piping.

Water suspension baths become contaminated if 0.5% oil is present. This condition occurs frequently in water baths, since lubricating oils are used where parts are being handled rapidly on a conveyor system. The following unfortunate

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

scenario often occurs on high-production, water-based, suspension systems: to improve part coating, additional wetting agent is sometimes added to the suspension. If too much wetting agent is added, the bath begins to foam. If foam covers the surface of a part being inspected, indications cannot be seen. An inspector sees some foam beginning to form in the reservoir, and counteracts it by adding a foam inhibitor—a liquid made of oil. Too much oil, and the particles start to come out of suspension and float on the surface of the bath. Because of this common pitfall, most high-volume MPI inspections are performed using oil suspensions.

4.4.2 Controlling Magnetization and System Performance

A variety of devices can determine the direction and amplitude of the magnetization field on an inspected part. These methods exclusively measure the fields on the exterior of the part. Many of the test devices not only measure the magnetization field but indicate overall system performance; they do not separate the response of individual parameters. These system-performance test devices typically employ artificial flaws called artifact standards. The most commonly used devices are Hall effect meters, pie gauges, QQIs, Burmah-Castrol strips, Ketos test rings, and pocket field indicators.

Hall Effect Meter

A Hall effect meter is a Gauss meter that measures the amplitude of either an AC or a DC magnetic field (Figure 4.36). In MPI, the Hall meter is used to determine the amplitude of the magnetic field in air parallel to the surface of a test part during part magnetization. The sensor capture area measures the amount of magnetic flux B passing through it. For this reason, the wand is held normal to and as close as possible to the part surface, to avoid reading any *leaking* field that leaves the part normal to the surface. Note this field is not the leakage field associated with a flaw.

Pie Gauge (magnetic penetrameter)

Pie gauges are available in many formats. Typically, the device is made by brazing together six pie-shaped pieces of high permeability, low retentivity material. Then the device is coated with a nonmagnetic material, typically the brazing material. Thus, the device contains six pie-shaped magnetic pieces separated by the distance of the nonmagnetic brazing material. The most common are about 2.5 cm in diameter and 3.2 mm thick (Figure 4.37). The gauge is laid on the part during MPI. The resultant particle indications on the radial, artificial flaws indicate the direction of the field tangent to the part surface and, to some extent, the tangential field amplitude. A pie gauge is sometimes used, along with a test part, to verify overall system performance if a sample part containing

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016



FIGURE 4.36 Gauss meter used to measure the strength of the magnetic field.

known defects is not available. It works better with the dry method than with the wet method.

QQIs (field indicators)

Quantitative Quality Indicators (QQIs) are paste-on artifact standards used to determine both the direction of magnetization and the overall system performance. These artifact standards contain finely etched flaws, and are made of AISI 1005 steel with moderate permeability and very low retentivity—in effect, a thin foil. They come in a variety of sizes, as small as 6.4 mm square by 50 microns thick. The precision-etched flaw is shaped like a crosshair (a circle superimposed on two perpendicular lines) and etched to a depth of 15 microns. QQIs are attached to the part with the flaw side down (i.e., in contact with the part). When a QQI is attached to the surface of a test part and inspected, the response indicates the direction of the field tangent to the surface of the part, as well as overall system performance. QQIs are commonly used in production-line testing to make a system performance sample, a given part tested periodically to verify system performance. Figure 4.38 shows a forging part, instrumented with multiple QQIs, ready to receive a head shot from a head- and tailstock MPI unit. A part with known actual flaws of the type of interest would be preferable, of course, but such specimens are difficult to obtain.

Copyright © Marcel Dekker, Inc. All rights reserved.



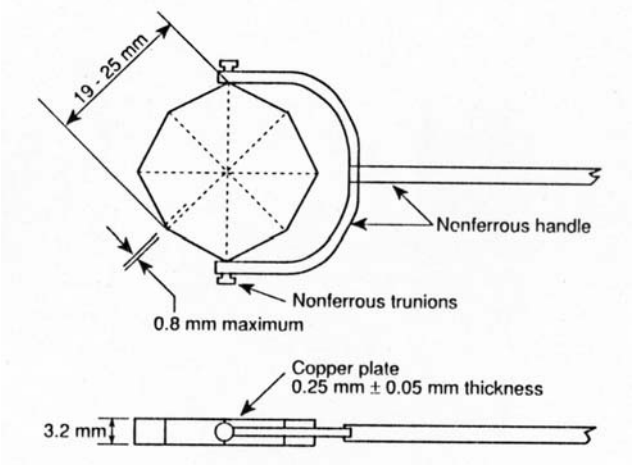


FIGURE 4.37 Pie field indicator used to verify MPI performance.

The threshold field for an indication is about 5 gauss; bright indications form at about 15 gauss. QQIs work better with the wet method than with the dry method, and are a convenient tool for use in balancing the fields in multidirectional magnetizing.

Note: Although QQIs are designed with very low-retentive material, when attached to a highly retentive material the "effective" QQI retentivity can be quite large; residual fields may create indications. This effect occurs only with longitudinal magnetizing fields, not with circular fields.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

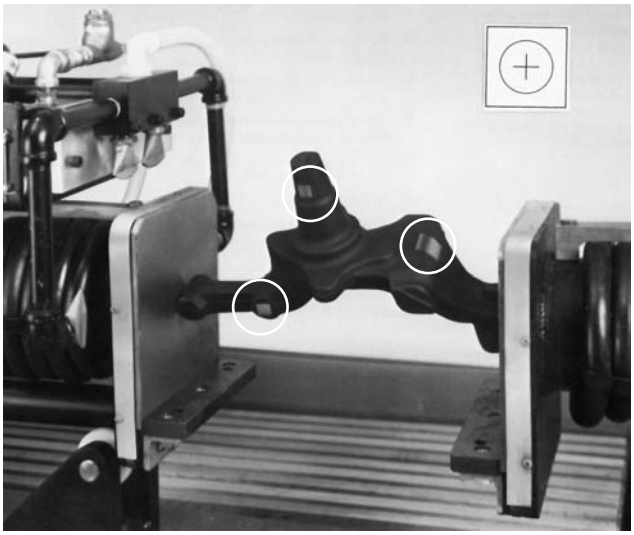


FIGURE 4.38 Forged part instrumented with QQIs to test system performance. The white circles indicate the location of the QQIs.

Burmah-Castrol Strips (field indicators)

Burmah-Castrol strips, another thin-foil form of artifact standard, are steel strips (about 5 cm long, 1.9 cm wide and 0.50 mm thick) used to indicate magnetization field direction. They have moderate permeability but unlike QQIs they are highly retentive. Burmah-Castrol strip flaws are linear, so multiple strips are usually placed perpendicular to each other. They contain three sizes of defects and are available in two sensitivities: aerospace and automotive.

Their high retentivity requires them to be demagnetized between use and makes them unsuitable for balancing multidirectional magnetization units. They work better with the wet method than with the dry method.

Ketos Test Ring

The Ketos test ring is a steel disk, often 125 mm in diameter and 22 mm thick with a 32-mm hole through the center. The ring contains a set of small holes in the surface with increasing distance (1.8 mm per hole) from the outside rim (Figure 4.39). When a Ketos test ring is magnetized using a central conductor, the resultant circular field produces leakage fields on the circumferential surface of the ring at the sites of the series of drilled holes.

The device was first developed in the 1940s as a means of checking the sensitivity of dry magnetic particles. Its ability as a tool for grading the quality of

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

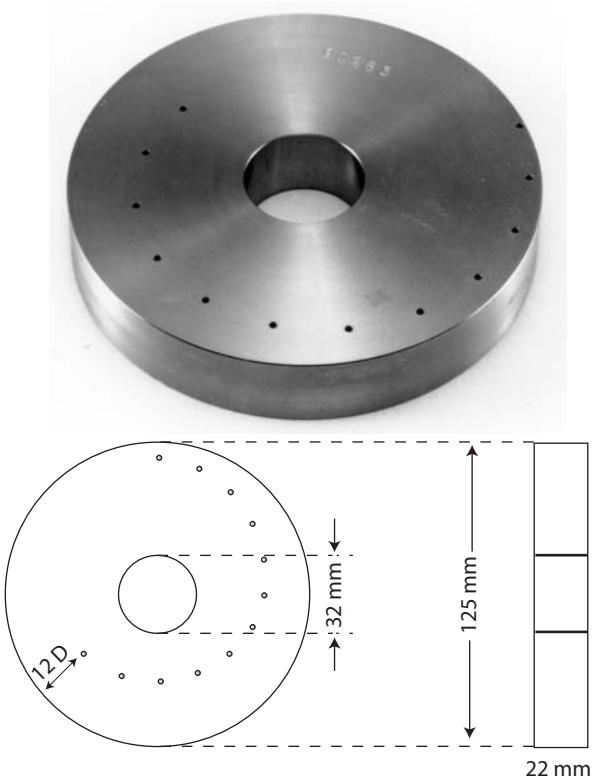


FIGURE 4.39 Ketos test ring used to test magnetic particles.

wet suspensions of magnetic particles is controversial. Most users of MPI think that parts with actual defects, or even paste-on defects, provide a more realistic tool for comparing one type or size of wet magnetic particles to another.

Pocket Field Indicators

Pocket field indicators may be round or square, calibrated in gauss or not calibrated. They are used to determine the magnetic field intensity leaving normal to part, i.e., at a pole. These devices are commonly used during demagnetization of a part.

4.4.3 Avoiding Nonrelevant Indications

An inspector should become familiar with the many types of nonrelevant indications that can appear on parts during MPI. These indications are caused

Copyright © Marcel Dekker, Inc. All rights reserved.

by magnetic field leakage that is not relevant to the strength or useability of the part. These indications commonly result from

The machining or cold working of the part
Attached metals
The shape of the part
External magnetic fields
Dissimilar metals

When a metal part is machined, local permeability changes can create a fuzzy indication. If two pieces are press-fit together, such as a shaft and pinion gear, there will be a sharp indication at the intersection. Also, the boundary zones in welds can create indications at the decarburized (heat-affected) zone.

If two different metals are welded together, the difference in permeability at the junction creates a leakage field, resulting in a fuzzy indication. Figure 4.40b shows a nonrelevant indication in a quality butt weld.

The shape of a part can also cause false indications. If the part has internal splines or drilled holes, the internal structure could reduce the magnetic field and thus create indications (Figure 4.40a). Additionally, if the part has sharp fillets, sharp thread roots, sharp keyways, etc., then the flux lines tend to bridge the gap, creating indications.

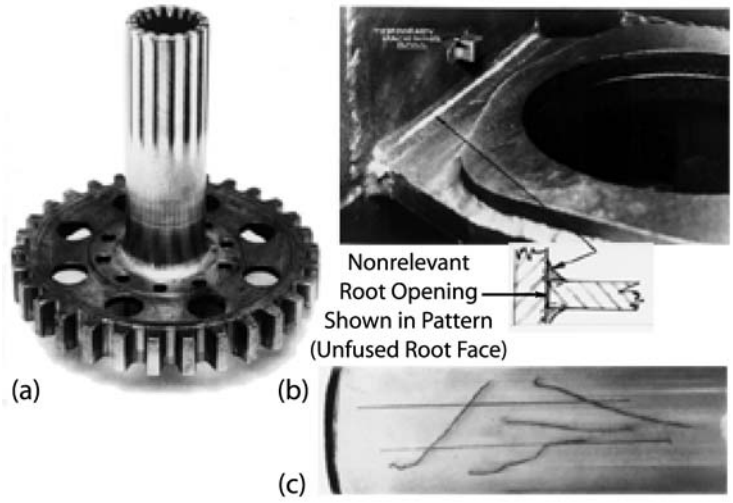


FIGURE 4.40 Examples of false or nonrelevant indications: (a) internal splines, (b) root opening on a butt weld, and (c) magnetic writing.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Finally, a nonrelevant indication can appear if the part has had contact with an external magnetic field outside the MPI system. For example, if a part has had contact with a current-carrying wire, then it contains lines of magnetized areas that interact with the magnetic particles. This phenomenon is often called magnetic writing (Figure 4.40c).

4.5 APPLICATIONS OF MPI

To understand MPI further, consider these specific examples from industry: steel coil springs, welds, and railroad wheels.

4.5.1 Steel Coil Springs

A manufacturer of steel coil springs must decide on a method of NDE to improve quality control. The springs range from 12–60 cm in length, 12–30 cm in diameter, and 1–2.5 cm in the cross section. The springs are formed from bar stock. A variety of cracks often form in the springs during processing and service.

Two types of cracks often appear during the manufacture of these springs: transverse stress cracks due to improper coiling procedures (Figure 4.41), and seams (Figure 4.42) which may later cause fatigue cracks that occur in the coil springs during service at 45° to the long axis (Figure 4.43).

The manufacturer considers several NDT methods. Radiography is too expensive on such a complicated shape and is unable to distinguish surface defects from internal ones. Ultrasonic detection is a possibility when the spring is in the barstock stage, but is too complex to use after coiling. Eddy current is a

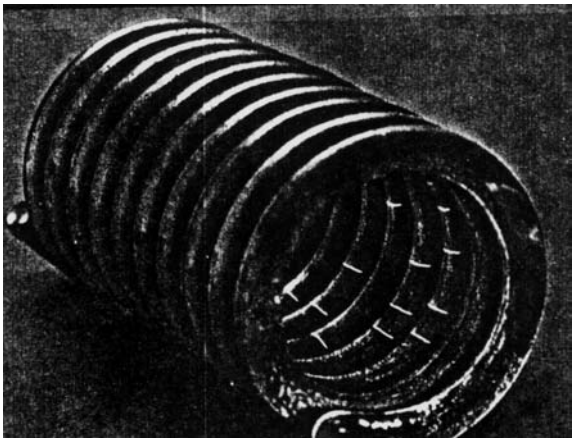


FIGURE 4.41 Transverse stress cracks in coil spring.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

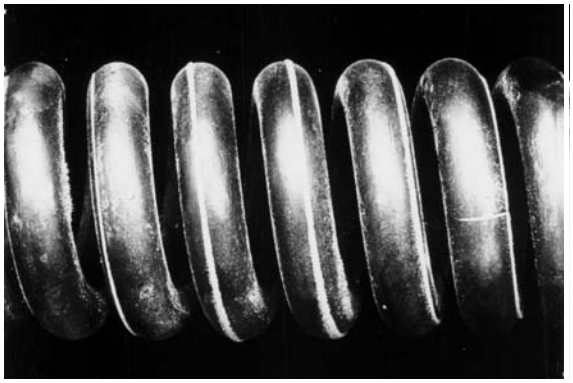


FIGURE 4.42 Deep seams in coil spring.

good method for locating seams before coiling, but would also be too complex to use after coiling. In addition, the many sizes of barstocks involved might make the setup for both ultrasound and eddy current prohibitively expensive and time-consuming. Penetrants could find seams in the barstock and stress cracks in the springs, but would be too costly in both materials and labor. After considering and rejecting all these methods, the manufacturer settles on MPI.

First, all coil springs are inspected for stress cracks (cracks from improper coiling) with a central conductor. As this method cannot be used on a conveyor belt, it is both time-consuming and expensive; so it is used only long enough to



FIGURE 4.43 Fatigue cracks in coil spring (45° from long axis).

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

determine the optimum method of coiling. The inspection is then reduced to a 3% sample—just enough to assure continued control of the coil winding.

Since only deep seams located on the coil springs' inside diameter ever resulted in fatigue failure, sorting out all seams at the barstock stage would be wasteful. Instead, all coil springs are inspected with MPI using a head shot. Since the coil springs are quite retentive, the manufacturer chooses the dry method, applying the particles by an automatic spray on a conveyer line. By adjusting the amount of current applied for the cross-section of the spring, all springs of all sizes are given the same surface magnetic field. The height of the dry powder indications, and thus the depth of the seam, can be judged with an error of less than 0.05 mm. Only the seams considered deep enough to cause failure on the inside diameter of the coil were rejected.

4.5.2 Welds

Weld inspection (Figure 4.44) is critical on many parts, and in some cases is also required by Federal law (see also Chapters 3 and 5). Welds in ferrous materials are commonly inspected with the MPI method. For new weld inspection,



FIGURE 4.44 Pipeline weld MPI using prods.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

subsurface sensitivity is a requirement; circular magnetization by passing a current through the part must be used. The geometry dictates the use of prods. Half-wave direct current (HWDC) is used for locating flaws as deep as 6.35 mm. Prod spacing should be 15–20 cm for best subsurface sensitivity, and the prods should have a current of 100 amps per 2.5 cm of spacing. Note: with prod spacing of more than 20 cm, the inspector starts to lose subsurface sensitivity regardless of the amount of current used.

Most weld defects run in a longitudinal direction along the weld. For maximum sensitivity, the prods should be placed along the length of the weld. (Remember: flaw-detection sensitivity is at a maximum when the current flows in the direction of the length of the crack.) The inspector lightly flows the particles on the work area while the current is on. The prods are stepped along the weld. Inspection lengths are overlapped by about 2 to 3 cm.

Prods are not likely to leave burn marks on a cast product, but arcing is common on wrought steel, particularly if the prod tips are not maintained. To reduce arcing, the prods should be pressed firmly to the part. Additionally, if excessive particles accumulate on the weld at the point of contact, the current level should be reduced.

Repair welds often contain defects that are closer to the surface; these welds are usually inspected using yokes. Because yokes pass magnetic field through the part, the legs of the yoke straddle the weld, allowing the field to encounter the defect at right angles. The yokes are then stepped up the weld. AC fields are used for surface-breaking or very-near-surface flaws, while DC fields provide some depth penetration. Most yokes used for this purpose are lightweight (about 3.4 kg), portable, and usually powered by 115 volts. Some yokes have articulated legs (kay-jays) for inspecting nonplanar surfaces.

4.5.3 Railroad Wheels

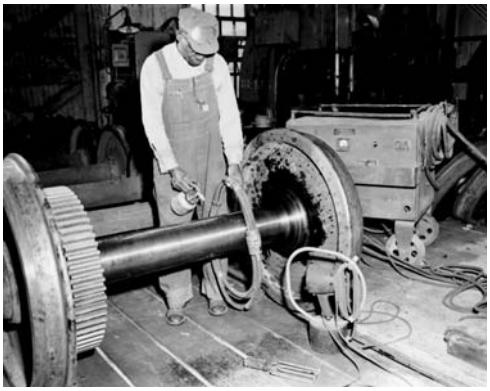
Railroad wheels and axles are commonly inspected with NDE methods, e.g., ultrasonic, penetrant, and eddy current. For MPI of in-service and used wheels and axles, the components are first removed from the train and thoroughly cleaned.

The wheel axle can be examined with a coil (Figure 4.45a); typical axle cracks are transverse to the axle length. The solenoid is in the form of a split coil, so that the operator can easily wrap the coil around the axle (Figure 4.45b). To inspect for defects in the gear teeth of the driving wheels, an ingenious device called a *hair pin coil* was created (Figure 4.46); this device fits between the teeth of the wheels, and works like a central conductor bar between the teeth. For overhaul on used wheels, flexible cable is threaded through the center and wrapped three times, creating a solenoidal coil.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

(a)



(b)



FIGURE 4.45 Railroad wheel axle MPI: (a) actual inspection and (b) detail of split coil.

New wheels are inspected by multidirectional magnetization methods. MPI is not used on installed railroad wheel systems.

4.6 SUMMARY

Although its applications in industry are many and varied, the theory behind MPI is simple. The first, and most important, issue is whether MPI is viable for the defect and the part to be inspected—in other words, whether you are inspecting a ferromagnetic material that may have a surface or near-surface defect. Also, you must decide if the surface coatings will interfere with the necessary inspection sensitivity.

If you decide that MPI can detect the defect in the part, the next question is the type of magnetizing method to use. Often, several methods are possible for a single object, and the choice of method often depends on which type of equipment is easiest to obtain.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

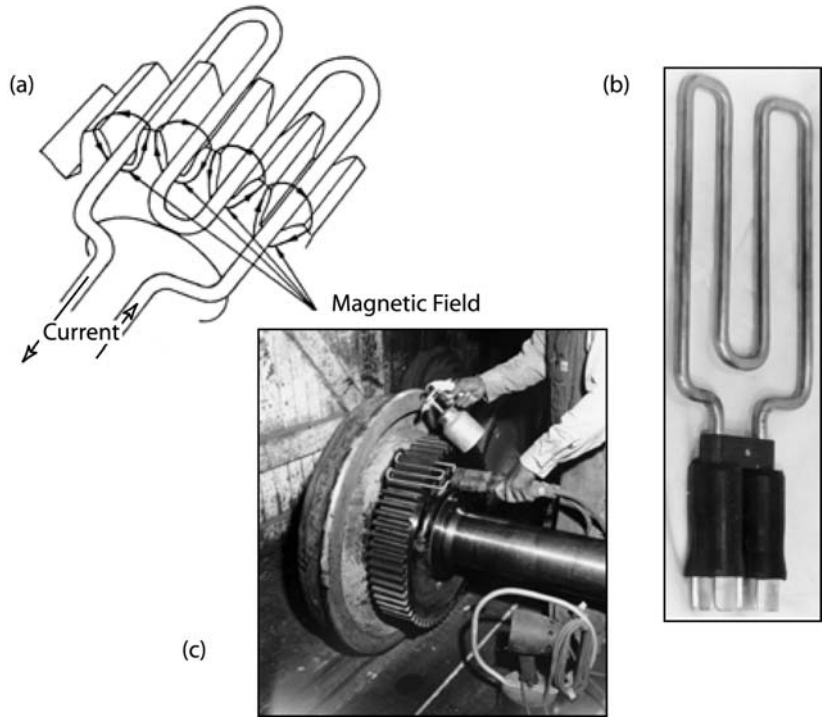


FIGURE 4.46 Specialized coil for inspection of railroad wheel drive gears: (a) diagram of hair pin coil, (b) actual hair pin coil and (c) hair pin coil in use.

The choice of current depends on the depth of the flaw and possibly the size of the part. AC current is used for detailed surface inspection. The choice of HWDC or 1(3)FWDC (single- or three-phase) methods depends in part on the whether wet or dry particles are used. For wet particles, HWDC has near-surface to intermediate penetration and 1(3)FWDC has the maximum penetration. For dry particles, HWDC has the maximum penetration. Finally, fluorescent particles are used with continuous wet method on most production applications. Dry particles combined with portable power sources are used on most preventive maintenance applications, multipass welds, and large castings.

MPI is one of the most economical NDE methods for detecting surface and near-surface flaws on ferromagnetic materials. Although very high currents are used, the very low voltage (8–15 V) makes this method safe for most environments.

Copyright © Marcel Dekker, Inc. All rights reserved.

PROBLEMS

Introduction

1. Describe the basic theory of MPI.
2. Is MPI useful for detecting surface cracks in railroad wheels? In ceramic parts? Why or why not?
3. Name the three important historical advances in MPI.
4. What are the two biggest limitations of MPI? The two biggest advantages?
5. In what way is MPI superior to the dye-penetrant method? In what way is it inferior?

Theory

9. Describe how an object can be magnetized without attracting or repelling nearby magnets.
10. By looking at the iron filings in Figure 4.2, describe the areas of high magnetic field and low magnetic field. How can you distinguish the two? Can you tell from the picture which end is north and which is south?
11. Can a current-carrying wire attract a nail? Why or why not?
12. Can a current-carrying wire cause a nail to rotate? Why or why not?
13. What is the magnetic field 1 ft from the center of a current-carrying wire with a 2 inch diameter and a current of 2 A?
14. How much current is needed to make a magnetic field of 50 tesla with a solenoid with 100 loops and a length of 2 ft?
15. Do all electrons make dipoles? Do all dipoles make magnets? Why or why not?
16. If a bar magnet is cut in two, what happens to the poles? The magnetism? The magnetic field?
17. An object has a magnetic susceptibility of 0.2, what is the permeability of the substance?
18. A sample with a constant magnetic susceptibility is in a solenoid with a current I. If the current doubles, what happens to the magnetic field strength, magnetism, and magnetic field in the sample?
19. What are the distinguishing characteristics of paramagnetic, diamagnetic, and ferromagnetic materials?
20. What is wrong with the statement "iron has a permeability of 1000"?
21. Which has a larger hysteresis loop, a hard or a soft magnet? Why?
22. Which has a larger saturation point, a hard or a soft magnet? Why?
23. Can a ferromagnetic material, a paramagnetic material, or a diamagnetic material have a positive magnetism with a zero magnetic field strength? Describe.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

24. Can a ferromagnetic material, a paramagnetic material, or a diamagnetic material have zero magnetism with a positive magnetic field strength? Describe.
25. Can a ferromagnetic material, a paramagnetic material, or a diamagnetic material have a positive magnetism with a positive magnetic field strength? Describe.
26. Can a ferromagnetic material, a paramagnetic material, or a diamagnetic material have a positive magnetism with a negative magnetic field strength? Describe.
27. Why do samples have to be magnetized in two directions to find all the defects in the sample?
28. If you double the frequency in a AC field, what happens to the skin depth?
29. Describe the four major types of current used in MPI and what types of defects they are best suited to find.

Techniques/Experiment

30. Describe two methods for creating longitudinal magnetization.
31. Describe two methods for creating circular magnetization.
32. Describe a method of multidirectional magnetization that is not mentioned in the text.
33. What happens to the field if two sinusoidal perpendicular magnetic fields are applied if the phase difference between them is zero? Why is this an unwanted effect in MPI?
34. What happens to the field if a constant magnetic field is applied with a perpendicular sinusoidal field? Is this as powerful as two sinusoidal perpendicular fields with a phase of 90 degrees? Why or why not?
35. Name two characteristics that are required of magnetic particles.
36. Why use fluorescent particles? Why use nonfluorescent particles?
37. What is the difference between the dry and the wet method?
38. What is the difference between the continuous method and the residual method?
39. Name three reasons for demagnetizing a sample after inspection.
40. Name two reasons for not demagnetizing a sample after inspection.
41. Why can't you demagnetize a sample by applying a field in the other direction to cancel the field in the sample?
42. Which is easier to demagnetize, a longitudinal field or a circular field? Why?
43. Describe the heating method to demagnetize a sample and explain why it is rarely used.
44. The five-turn 16-in. diameter coil on your DC M.P.I. unit has a maximum output of 3000 amp. You accept a contract for inspecting a variety of bolts and then learn that your customer's specific spec on some sizes require a coil

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

output of 4200 amp. The contract is short and the customer a valued one. What action is best?

(a) Find him another inspection source.
(b) Replace your low output unit with a higher one.
(c) Review the spec with him to obtain a deviation.
(d) Process some of the bolts, two at a time.
(e) Process all of the bolts with a longer (one second) coil shot.
(f) Top tap your power transformer.
(g) Change your unit from 220 V to 440 V operation.

45. Two retentive bars, 1 ft long and 1 in. in diameter are magnetized. One has a longitudinal field, the other, a circular field. To determine which has the longitudinal field:

(a) Lay them side by side (parallel to each other).
(b) Form an "**X**" with them.
(c) Form a "**T**" with them.
(d) Form an "**I**" with them (end to end).
(e) Not possible without additional aid (paper clip, Field Indicator, etc.).

46. For best results, when demagnetizing a coil spring, 1 ft long, 4 in. diameter, 1 in. cross-section, using an AC demagnetizing coil 16 in^2 pass the spring through the coil:

(a) Spring axis parallel to coil axis
(b) Spring axis perpendicular to coil axis
(c) Same as (a) but twist spring going through
(d) Same as (b) but roll it through
(e) Same as (a) but put two springs together, end to end

GLOSSARY

Central conductor: Insulated conducting bar that is placed in a hollow tube to create a circular magnetic field; also known as *threader bar.*

Circular magnetization: Magnetization that points in circles.

Coercive force: Amount of external field needed to nullify the magnetization of a sample.

Continuous method: Method to apply magnetic particles when the magnetic field is highest.

Core, soft metal: Tube of metal that is placed in hollow samples to increase magnetic field from the induced current and the central conductor methods.

Curie temperature: Temperature at which ferromagnetic materials become paramagnetic.

Copyright © Marcel Dekker, Inc. All rights reserved.

Currents

Alternating (AC): Current that changes direction with time, often in a sine-like pattern.

Direct (DC): Current with a constant value.

Half wave (HWDC): Current composed only of the positive current from an AC source.

Single phase–full wave (1FWDC): Current altered from an AC source so that all current flows in one direction .

Three phase–full wave (3FWDC): Combination of three (1FWDC) waves combined to mimic a DC current.

Diamagnetic: Having small negative permeability and no residual magnetization.

Dipole: Smallest magnet in a sample, due to individual electrons in atoms.

Domain: Area in a ferromagnetic material where the dipoles align to make a positive magnetic effect.

Dry method: Method to apply magnetic particles as a dry powder.

Eddy current: Current created in a substance due to a changing external magnetic field; see (Chapter 5 on Eddy Current).

Ferromagnetic: Having large positive permeability and residual magnetization; able to be made into magnets.

Flux-Flo™: Two solenoids connected at each end of a clamp to create longitudinal waves.

Gauss meter: Digital meter to measure magnetic field in air; also known as *Hall effect meter*.

Hard magnets: Objects with high residual magnetization (see Soft magnet).

Hair pin coil: Current-carrying coil that fits between the teeth of the drive-wheel gears.

Hall effect meter: See Gauss meter.

Head shot: Current that flows through a sample.

Hysteresis: Difference between the original path and the return path in a curve; in MPI, often refers to curve in *B–H* diagram.

Hysteresis loop: See hysteresis..

Induced current: Magnetizing method in which current is created in a circular sample when an external field (usually from a solenoid) is removed; method is often used with a core..

Ketos test ring: Disk with small holes in it, to be examined with MPI to determine the quality of the magnetic particles.

Kay-jay: portable yoke with articulating legs to accommodate irregular surfaces.

Leakage field: Field created when a cut or defect is made in a sample with an internal magnetic field.

Copyright © Marcel Dekker, Inc. All rights reserved.

Lenz's law: Physics law that states that if the magnetic field in a closed circuit changes, a current will be induced in the circuit that creates a second magnetic field that opposes the change in the first magnetic field.

Longitudinal: Magnetized along the long axis of a sample.

Magnetic effect: Ability to attract or repel other magnets or pieces of metal.

Magnetic field density (B): Density of magnetic flux per unit area. B is measured in SI units of tesla (T) and CGS units of gauss (G).

Magnetic field strength (H): The magnetic field from an external field divided by the permeability of free space; measured in oersteds.

Magnetic flux density: See Magnetic field.

Magnetic induction: See Magnetic field.

Magnetic susceptibility (χ): Ratio of H over M; a constant χ indicates that the material cannot be made into a magnet.

Magnetic writing: Lines of magnetized areas due to contact with external magnetic fields.

Magnetization (M): Total magnetic effect from the dipoles divided by the volume of the object.

Multidirectional magnetization: Magnetization that has two directions, usually varying between the two directions over time with a swing field.

North/south poles: Points of concentrated magnetic effect.

P-jay: Large C-shaped magnet clamped around a small object to magnetize it.

Paramagnetic: Having small positive permeability and no residual magnetization.

Permeability of free space (μ_0): Constant defined so that to relate the units of B and of distance and current. SI units, $\mu_0 = 4\pi \times 10^{-7} N/A^2$, CGI units $\mu_0 = 1$.

Relative (μ_r): Ease with which a magnetic field enters an object; also the ratio of $B/(\mu_0 H)$.

Polarity: North-south orientation.

Prod: Rod connected to a high-amperage, low-voltage source to create a circular magnetic field.

Quantitative quality indicator (QQI): Small shim with etched face to measure direction and magnitude of magnetic field.

"Quick-break" circuitry: Circuitry to cut off the current quickly; often used in conjecture with the induced current method.

Residual magnetism: Amount of magnetization a material retains after an external field is removed.

Residual method: Method to apply magnetic particles after the external magnetic field has been turned off.

Retentivity: Ability of a substance to retain magnetization after the removal of an external field.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Right hand rule: A common method to determine the direction of the magnetic field from a current-carrying wire; if you point the thumb of your right hand along the current, your fingers will curl in the direction of the magnetic field.

Saturation point: Point at which a sample has reached maximum magnetization due to an external field.

Skin depth: Maximum depth to determine defects at a set frequency.

Sloppy continuous method: Method in which some magnetic particles in liquid are allowed to flow over the sample after the current has stopped.

Soft magnet: Object with low residual magnetization (see Hard magnet).

Solenoid: Coil of conducting wire with more than one turn.

Split coil: Solenoid cut into two pieces so that the operator can easily wrap the coil around the axis.

Swing field: Magnetic field whose direction alternates with time.

Threader bar: see Central conductor.

Toroidal field: Magnetic field created in a ring-shaped object by a current passing through the center of the ring.

Wet method: Method to apply magnetic particles in a liquid suspension.

Yoke: **U**-shaped magnet or **U**-shaped metal wrapped with wire to create a magnet.

Connected yoke: Yoke connected to the edges of a sample (see *P-jay*).

Surface yoke: Yoke placed on the surface of a sample (see *Kay-jay*).

VARIABLES

$\vec{B}$	Magnetic flux density
$\vec{H}$	Magnetic field strength
μ	Magnetic permeability
μ_0	Magnetic permeability in free space
μ_r	Relative magnetic permeability
M	Magnetization
χ	Magnetic susceptibility
ϕ	Magnetic flux
N	Number of turns of a solenoid
ρ	Resistivity
σ	Conductivity
δ	Skin depth
f	Linear frequency
ω	Angular frequency
I	Current
V	Voltage
D	Diameter of the part

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

REFERENCES

1. CE Betz. Principles of Magnetic Particle Testing. Chicago: Magnaflux Corporation, 1967.
2. A Lindgren. Demagnetization Techniques (video tape). L & L Consultants, 1629 Eddy Lane, Lake Zurich, IL.
3. D Lovejoy. Magnetic Particle Inspection, A Practical Guide. London: Chapman and Hall, 1993.
4. Magnaflux Seminar 1987/1988. Chicago: Magnaflux Corporation.
5. Magnaflux Seminar 1993/1994. Chicago: Magnaflux Corporation.
6. Nondestructive Testing Handbook. 2nd ed. Columbus, OH: American Society for Nondestructive Testing, 1989.
7. K Schroeder. Internal Applications Laboratory Report. Chicago: Magnaflux Corporation.
8. R Serway. Physics for Scientists and Engineers. London: Saunders College Publishing, 1992.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

5

Eddy Current

Peter J. Shull*
The Pennsylvania State University, Altoona, Pennsylvania

5.1 INTRODUCTION

Searching for gold, buried treasure, or just the plain old metal sewer line buried somewhere in your yard? Think *eddy current*! Using eddy current (EC) principles, common "metal detectors" that squeal at the presence of metallic objects can detect metal at significant depths within the earth or a concrete wall—without having to "jackhammer" an entire concrete wall or floor.

EC probes measure a material's response to electromagnetic fields over a specific frequency range, typically a few kHz to several MHz for traditional NDE applications. From this response, we can interpret material conditions such as hardness, thickness, presence of corrosion, or defects such as porosity and cracks. The two properties directly interrogated by the EC probe are electrical conductivity (as the term "current" suggests) and magnetic permeability. Thus the method inspects only conductive materials, primarily metals, but also low-conductivity materials such as graphite-epoxy composites.

*Applications submitted by David Mackintosh and Jim Cox.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

EC method is a mature, proven NDE technology with a solid theoretical foundation. It is widely represented in industries such as automotive and aircraft manufacturing, and it is an integral part of inspection and maintenance in the power generation and aircraft industries. Additionally, EC methods are extensively used in process control.

5.1.1 Technical Overview

Eddy current methods interrogate conductive materials through magnetic induction. The probe itself is nothing more than a AC transformer. Before we discuss the transformer nature of the EC probe, remember these basic electromagnetic concepts:

Current flow in a wire generates a magnetic field that encircles the wire (current flow) and points tangentially to the circles.
A magnetic field in proximity to a conductor produces a voltage, or *electromotive force* (EMF), in the conductor. If we assume a closed circuit, a current will flow.

If the wires in either of the two cases are wound to form a long coil (a *solenoid*), the effects are multiplied by the number of turns. Thus we have a mechanism through simple geometry of the wire (coil) to amplify the signal. Note that this is an AC phenomenon; DC fields do not work.

EC measurement consists of four steps:

1. Signal excitation
2. Material interaction
3. Signal pickup
4. Signal conditioning and display

Imagine an EC probe as a pair of coils (Figure 5.1). One coil, the *excitation coil*, is excited with an AC signal; the other, the *pickup coil*, is connected to a voltmeter. The excitation coil produces a (*primary*) magnetic field, part of which passes through (i.e., couples to) the pickup coil, thus exciting an EMF (voltage). The pickup voltage reading will remain constant, assuming a constant drive current and fixed coil locations, until a ferromagnetic or conductive material is brought near the probe and perturbs the field. This change in the magnetic field coupling is reflected in the EMF induced in the pickup coil.

Clearly, ferromagnetic material will perturb a magnetic field, but how does conductive material influence the field? Recall that bringing a conductor into the presence of a magnetic field produced a current in the conductor. Likewise, the primary magnetic field induces currents in the test specimen. These currents must travel in a closed path—generally circular—and are called *eddy currents*. Like all currents, eddy currents must produce a (*secondary*) magnetic field that opposes

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

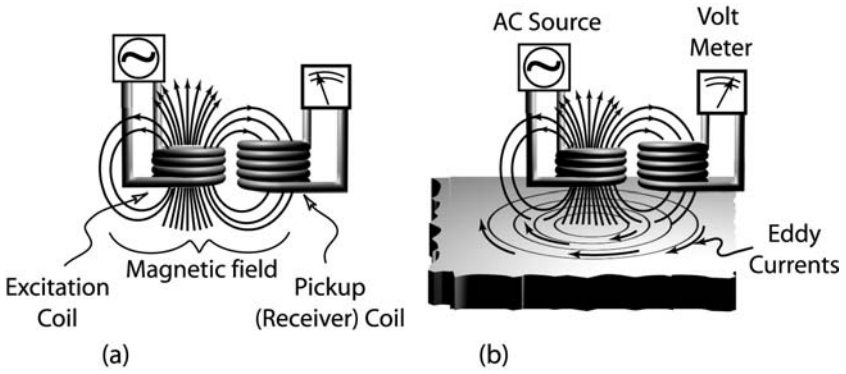


FIGURE 5.1 An eddy current probe (an AC transformer) used to detect the character of conductive materials. Probe response (a) in the absence of a conductive material and (b) in the presence of a conductive material.

the primary field. The pickup coil monitors this reduction of the total magnetic field—primary plus secondary. The magnitude of the eddy currents, and thus of the secondary field, depends on the conductivity of the test material. A perfect conductor would result in perfect primary-field cancelation by the secondary field (assuming complete coupling between EC probe and test specimen).

We can deduce a surprising amount of information by interrogating the conductivity of a material. Conductivity variations occur with material processing, hardness, and temperature. Absence of material—from cracks, voids, coatings, or thinness (such that the interrogating field "sees" through the sheet to the adjoining medium, often air)—implies an apparent change in the overall conductivity from that of a defect-free specimen. Additionally, as eddy current methods are electromagnetic in nature, variations in magnetic permeability will affect the signal response.

In the eddy current method, response signals are displayed in a variety of formats, all of which represent some form of impedance change in the pickup coil. Many application-specific systems simply measure EMF amplitude changes but display the desired parameter, such as coating thickness or hardness. More advanced eddy-current systems measure both the amplitude and the phase. These response signals are displayed on a complex plane called the *impedance plane* (resistance versus reactance).

The above overview concerns *reflected* impedance. There are, however, two specialized probes that do not observe reflected impedance, but instead measure a transmission from one coil to another. The forked and remote field eddy-current probes place the test sample between the excitation and pickup coils.

Copyright © Marcel Dekker, Inc. All rights reserved.

5.1.2 History

The history of EC methods is inextricably linked to developments in electromagnetism. The developments that led to actual eddy-current testing were the products of profound work performed in the 19th century. Before 1820 there was little evidence that electricity and magnetism were linked, although the possibility had been suggested by some. That year, Hans Christian Ørsted experimentally demonstrated that a magnetic needle placed parallel to a current-carrying wire would be deflected. Ørsted's account of the phenomenon states that

> The current itself . . . was the cause of the action and that the "electric conflict acts in a revolving manner," that is, that a magnet placed near a wire transmitting an electric current tends to set itself perpendicular to the wire, and with the same end always pointing forward as the magnet is moved around the wire. . . . The space in which these forces act may therefore be considered as a magnetic field (1).

Biot and Sevart repeated Ørsted's experiment and discovered the law that bears their name. Quantifying the magnitude and direction of the magnetic field surrounding a current-carrying wire, they found that the magnetic field drops off as inverse distance and points perpendicular and circumferential to the direction of the current.

In 1824, Arago experimentally found that the vibration of a magnetic needle was rapidly damped when near a conductor. This curiosity was not explained until 1831, when Joseph Henry in the United States and Michael Faraday in England individually discovered electromagnetic induction. Their experimental apparatus was the now-common transformer with primary and secondary coils on a toroid; one of Faraday's coils is shown in Figure 5.2. The primary coil was connected to a voltaic battery that could maintain, stop, or



FIGURE 5.2 Stamp honoring Michael Faraday, induction coil depicted. (Courtesy Radiology Centennial, Inc.)

Copyright © Marcel Dekker, Inc. All rights reserved.

reverse the current flow. The secondary coil was initially monitored with a magnetic needle placed in parallel to the wire—later, with a galvanometer. Faraday found that when the primary circuit was closed, the needle was momentarily deflected in one direction; when the circuit was opened, the needle was momentarily deflected in the opposite direction. This result clearly demonstrated electromagnetic induction's dependence on time-varying current. Faraday experimented with cores made of iron and wood, finding similar results but with reduced amplitude of the needle deflection for the wooden core.

The existence of eddy currents was first demonstrated by J. B. Foucault in 1830. (Today in France and French-speaking Canada, eddy currents are called *Foucault currents*.) The first documented use of eddy currents in NDE is recounted by R. Hochschild:

> The use of electromagnetic waves for non-destructive testing of metals for flaws and physical properties antedates even the experimental proof of the reality of these waves. In the year that Maxwell died, 1879, at a time when many still doubted his theories and eight years before Hertz demonstrated the existence of electromagnetic waves, D. E. Hughes distinguished different metals and alloys from one another by means of induced eddy currents. . . . Lacking an electronic oscillator, Hughes used the ticks of a clock falling on a microphone to produce the exciting signal. The resulting electrical impulses passed through a pair of identical coils and induced eddy currents in [conductive] objects placed within the coils. Listening to the ticks with a telephone receiver [invented by A. G. Bell two years earlier], Hughes adjusted a system of balancing coils until the sounds disappeared. . . (2).

Hughes commented on the sensitivity of his EC device:

> A millagramme of copper or a fine iron wire, finer than the human hair, can be loudly heard and . . . its exact value ascertained. . . . I have thus been able to appreciate the difference caused [in a shilling] by simply rubbing the shilling between the fingers, or the difference of temperature by simply breathing near the coils. . . (3).

Hughes measured the conductivity of various metals on his induction balance, using copper as a reference standard. This standard, the "International Annealed Copper Standard" (IACS), survives today as a common conductive measure; values are given as a percentage of the conductivity of copper. A relative scale, % IACS, appears in much of the literature on eddy current.

For the next fifty years, no significant advances in EC testing were reported. At the end of the 1920s, EC devices began to appear in the steel industry for measurements on billets, round stock, and tubing. But the limitations of electronic instrumentation (e.g., stable oscillators) allowed no more than some

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

simple sorting applications. Instrumentation and electromagnetic theories developed during WWII, primarily for development and detection of magnetic mines, paved the way for the robust testing methods and equipment that allowed eddy current its entry into mainstream industry.

In the early 1950s, Dr. Federico Förster presented developments that began the modern era of eddy current NDE (4,5). Förster combined precise theoretical and experimental work with practical instrumentation. Clever experiments using liquid mercury and small insulating tabs allowed accurate discontinuity measurements. These experimental results were compared to solutions of Maxwell's equations. Förster produced precise theoretical solutions for a number of probe and material geometries. In a major development in quantitative EC testing, Förster adapted Charles Proteus Steinmetz's complex notation for sinusoidal signals to his phase-sensitive analysis. The EC response was displayed on a complex impedance plane, inductive reactance (energy storage) plotted against real resistance (energy loss). Conventional EC testing and analysis rely on this basic impedance-plane method. In addition to theoretical development proven by clever and detailed experimentation, Förster and his colleagues designed capable measuring equipment. During the 1950s and 1960s Förster's equipment and methods made eddy current an accepted industrial tool. Förster's work has rightly identified him as the father of modern EC testing.

Since the 1960s, significant progress has been made in theoretical modeling, instrumentation, and methodology. Advances have been made in basic understanding of the interaction between electromagnetic probes and specimens, as well as in flaw inversion modeling: obtaining quantitative flaw characteristics by processing the actual probe response through a theoretical model. Hugo Libby pioneered self-balancing systems that use a standard test sample as a reference, as well as a multifrequency EC method used to discriminate between the desired flaw signal and extraneous signal responses such as liftoff or temperature (6–8).

Today, eddy current is one of the most commonly employed NDE methods in industry.

5.1.3 Potential of the Method

EC methods are commonly employed in a vast range of industries: research, manufacturing, power generation, and maintenance. The methods are readily applied to process control, quality control, and in-service integrity inspection. As with many NDE methods, EC sensors respond to only a small number of parameters: conductivity, magnetic permeability, and probe and sample geometry. But this simple set of parameters provides extensive, useful knowledge about conductive and magnetic materials. The following, adapted from Ref. 9, lists some EC sensing uses:

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

1. Measuring the thickness of metallic foils, sheets, plates, tube walls and machined parts from one side only, by noncontacting means
2. Measuring the thickness of coatings over base materials, where the coating and base material have significantly different electrical or magnetic properties
3. Identifying or separating materials by composition or structure, where these influence electrical or magnetic properties of the test material
4. Detecting material discontinuities (which lie in planes transverse to the eddy currents) such as cracks, seams, laps, score marks or plug cuts, drilled and other holes and laminations at cut edges of sheet or plate
5. Identifying and controlling heat treatment conditions and evaluating fire damage to metallic structures
6. Determining depths of case hardening of steels and some ferrous alloys
7. Locating hidden metallic objects such as underground pipes, buried bombs or ore bodies
8. Timing or locating the motions of hidden parts of mechanisms, counting metallic, or objects on conveyor lines
9. Measuring the precise dimensions of symmetric, machined or ground and polished metallic parts, such as bearings and bearing races, small mechanism components, and others

5.1.4 Advantages and Disadvantages

EC inspection is popular for many reasons:

EC readily inspects *metals*—after concrete, the primary structural material in use.*

EC is *noncontacting*—a major advantage over conventional ultrasonics, penetrants, and magnetic particle (Note: EMAT and laser-based ultrasound are also noncontacting), allowing for automated high-speed inspection.

Unlike penetrants and most magnetic-particle methods, EC does not require surface preparation.

The low system cost of EC is a clear advantage over X-ray methods.

EC is one of the few methods used for high-temperature applications.

EC is portable.

EC methods are highly sensitive to a broad range of geometric and material parameters (see previous section).

On the other hand, EC inspection suffers from a number of drawbacks. The most troublesome is that the interrogated material must be a conductor. Although

*Eddy current inspection is used to inspect reinforced concrete, but using only the information obtained from the metallic reinforcing bars.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

TABLE 5.1 Advantages and Disadvantages of Eddy Currrent as a Nondestructive Testing Method

Advantages	Disadvantages
Inspects conductors	Only inspects conductors
No safety concerns	Surface or near surface detection only (specialized forked probes for thin sheet and RFEC probes for tubing are through thickness devices. The latter is capable of inspecting carbon steel tubing thickness up to 6–10 mm thick).
Rapid Inspection	
Sensitive to a large number of parameters related to conductivity, magnetic permeability, and geometry (e.g., flaws, thickness, coatings, hardness, proximity, and edges)	Sensitivity to a wide range of parameters increases complexity of interpretation
Wide operating temperature range.	Sensitive to liftoff variations
Specialized off-the-shelf probes for high temperature applications	Only sensitive to cracks perpendicular to the interrogating surface
Small probe size	
Light weight and portable	
Relatively low cost	
Can be configured in arrays	
Mature technology	

the ability to respond to a large variety of material and geometric properties is a positive trait, it can create significant difficulty in interpreting the response. The EC probe responds not only to the intended material characteristic (conductivity, discontinuities, etc.), but also to unintended signals such as liftoff, and to all of the unwanted material properties related to conductivity and/or magnetic permeability. Separating the response signals, or reducing unwanted signals, requires significant operator training and/or sophisticated algorithms.

Table 5.1 summarizes these and other advantages and disadvantages of EC methods.

5.2 BASIC ELECTROMAGNETIC PRINCIPLES APPLIED TO EDDY CURRENT

In essence, an eddy-current probe is an electromagnetic transformer. This simple fact means that understanding eddy-current systems is, first, a matter of reminding yourself of basic electromagnetic principles and thinking about them in this new context. This section focuses on such principles—magnetic induction and impedance—and relates them to various aspects of EC systems: probes, notation, and different test samples.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

5.2.1 Magnetic Induction (Self and Mutual)

The relationship among current, magnetic field, and voltages provides the basis for EC devices, just as it underlies the operation of such modern conveniences as the radio and television, electric motors, and the even the 110 V transformers used in common household electronics such as answering machines. Simply put, *any current flowing in a wire will induce an associated magnetic field, and any time-varying magnetic field in the presence of a conductive wire will induce a voltage in the wire*. If the wire in the latter case is in a closed loop, current will flow. This reciprocal relationship between current/voltage and magnetic field is called *magnetic induction*.

Current-induced Magnetic Field

Imagine a current flowing in a long, straight wire, and the magnetic field it induces. What does the field look like, and what is its relationship to the current?

Experimentally, we can determine the direction of the field either (a) by sprinkling iron filings on a card perpendicular to the current-carrying wire, or (b) by mapping the field with a compass (Figures 5.3a, b). We observe a magnetic field made up of concentric rings perpendicular to the current flow. The orientation of the magnetic field, indicated by the poles of the compass, depends on the direction of current flow. The direction of the current relative to the magnetic field can be determined through the right-hand rule: if the outstretched fingers of the right hand bend in the direction of the magnetic field, the outstretched thumb will point in the direction of current flow (Figure 5.3c). The circular magnetic field begins and ends upon itself—unlike an electric field, which begins and ends on an electric charge, an electron, or a hole (the absence of an electron).*

Through conservation of energy, the induced magnetic field must decrease with increasing radial distance from the wire—$1/r$ for a long, straight wire. To make up for this decrease, we can increase the magnetic field strength and the magnetic flux density in one of two ways. First, if we shape the current-carrying wire into a loop, we concentrate the magnetic field (strength) within the enclosed area (Figure 5.4). To further concentrate (increase) the magnetic field (strength), we might add loops in the wire, creating the common solenoid. The increase in magnetic field strength, H, is proportional to the number of loops, N. Second, the magnetic field strength is related to the magnetic flux (field) density, B, through the magnetic permeability, μ, as $B = \mu H$ (Section 4.2). The magnitude of μ depends on the material within the solenoid, e.g., the solenoid may have an air core (very low μ) or an

*In a permanent bar magnet, we designate a north and a south pole. This does not imply that there exist positive magnetic charges at one end and negative magnetic charges at the other end, as is the case for electric charges in a charged capacitor (holes and electrons). Despite extensive efforts, no one has observed the existence of a magnetic monopole.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

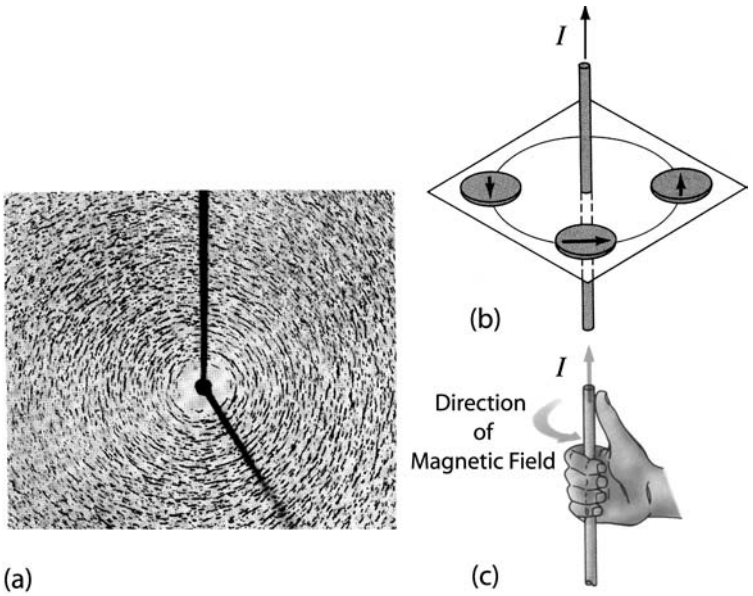


FIGURE 5.3 Orientation of the magnetic field relative to a current flowing in a straight wire. (a) Iron filings sprinkled on a card. (b) Direction of a compass as in Ørsted's (1777–1851) experiment. (c) Pictorial of the right-hand rule.

iron core (very high μ). Therefore, the magnetic flux density can be greatly increased by either increasing the number of turns or the μ of the core material.

The Biot-Savart law describes mathematically the relationship between an impressed current, I, in a conductor and the induced magnetic flux density (Figure 5.5):

$$\vec{B} = \frac{\mu I}{4\pi} \int \frac{d\vec{l} \times \hat{a}_r}{r^2} \tag{5.1}$$

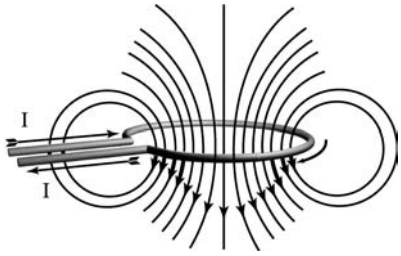


FIGURE 5.4 Concentration of the magnetic field by forming a current carrying wire into a loop.

Copyright © Marcel Dekker, Inc. All rights reserved.

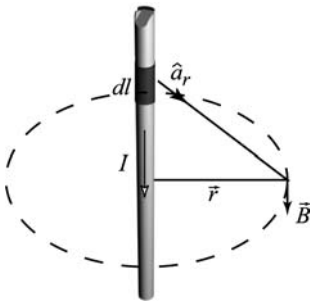


FIGURE 5.5 Magnetic flux density, $d\vec{B}$, developed by the current element, dl.

where r is the distance perpendicular from the wire to the point in space where $\vec{B}$ is to be calculated, and $\hat{a}_r$ is a unit direction vector pointing from a point on the current-carrying wire (source point) to the point in space where $\vec{B}$ is to be calculated. The integral is carried out over the length, l, of the wire. The cross product $\vec{l} \times \hat{a}_r$ states that $\vec{B}$ will be in a direction perpendicular to both the current, which is confined to the wire, and the direction vector $\hat{a}_r$, i.e., $\vec{B}$ will be circumferential to the wire.

Magnetic-field–Induced Current

If we place a conductive wire in a magnetic field, a voltage (EMF) is generated in the wire and a current will flow (provided the wire is not an open circuit). The amount of current that flows is proportional to the magnetic flux,

$$\phi = \int_{\substack{\text{loop} \\ \text{area}}} \vec{B} \cdot \hat{n} dA \tag{5.2}$$

enclosed (captured) by the area of the loop, where $\hat{n}$ is a unit vector normal to the enclosed area. (Note that although this is simply the magnetic flux density ($\vec{B}$) times the loop capture area, $\vec{B}$ is not a constant.) Faraday's law of electromagnetic induction relates the total captured magnetic flux to the induced voltage:

$$V_{\text{emf}} = -N \frac{d\phi}{dt} \tag{5.3}$$

The negative sign states that the resultant current flow due to the induced voltage will flow in such a direction as to oppose the change in magnetic flux. This is a statement of Lenz's law. The current-to-magnetic-flux relationship is

$$\phi = LI \tag{5.4}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

where the constant of proportionality, L, is called *inductance*. The value of L accounts for geometric factors of the wire, such as the shape and size of the loop and the number of turns and permeability of the core material. Thus,

$$V_{\text{emf}} = -L\frac{dI}{dt} \tag{5.5}$$

In a circuit with an ideal inductor in series with a power supply,

$$\begin{aligned} V_{\text{Applied}} &= -V_{\text{emf}} \\ V_{\text{Applied}} &= L\frac{dI}{dt} \end{aligned} \tag{5.6}$$

Self-Inductance

Clearly, electron movement (current) produces a magnetic field, the strength of which is dependent on the geometry of the wire. Because energy is required to produce this magnetic field, the current experiences additional impedance apart from the normal resistance of the wire. This impedance is known as *inductive reactance* and is directly proportional to both the frequency, f, and the inductance, L. The argument for the increase in impedance (inductance) and therefore a decrease in current is a typical transient-event circular argument:

An applied external voltage impresses a current in a wire, which in turn produces an associated magnetic field.

The magnetic field in the presence of the wire *induces* a secondary current (via an induced EMF) that *opposes* the *impressed* current (Lenz's law).

The final current is the sum of the impressed current and the induced current.

This argument exposes two important points. First, the events are circular and act upon themselves. Thus the creation of an EMF in a circuit by virtue of a change in its own magnetic field is called *self-inductance*, L_S. Second, the event is transient; therefore the impedance of self-inductance must be time dependent, confirming Eq. (5.5). Unlike *resistive* impedance, however, where the energy is lost to heat, an inductor (the physical device that creates inductance) stores the energy as the magnetic field and can be retrieved when the impressed current is removed and the magnetic field collapses.

Mutual Inductance and EC Probes

So far we have observed that a current impressed in a loop of conductor (primary coil) produces a magnetic field which extends into space as it loops upon itself. If a second loop (secondary coil) is placed within the magnetic field, a current will flow in this loop proportional to the captured flux (Figure 5.6a).* The creation of

*This concept is the basis for the common transformer.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

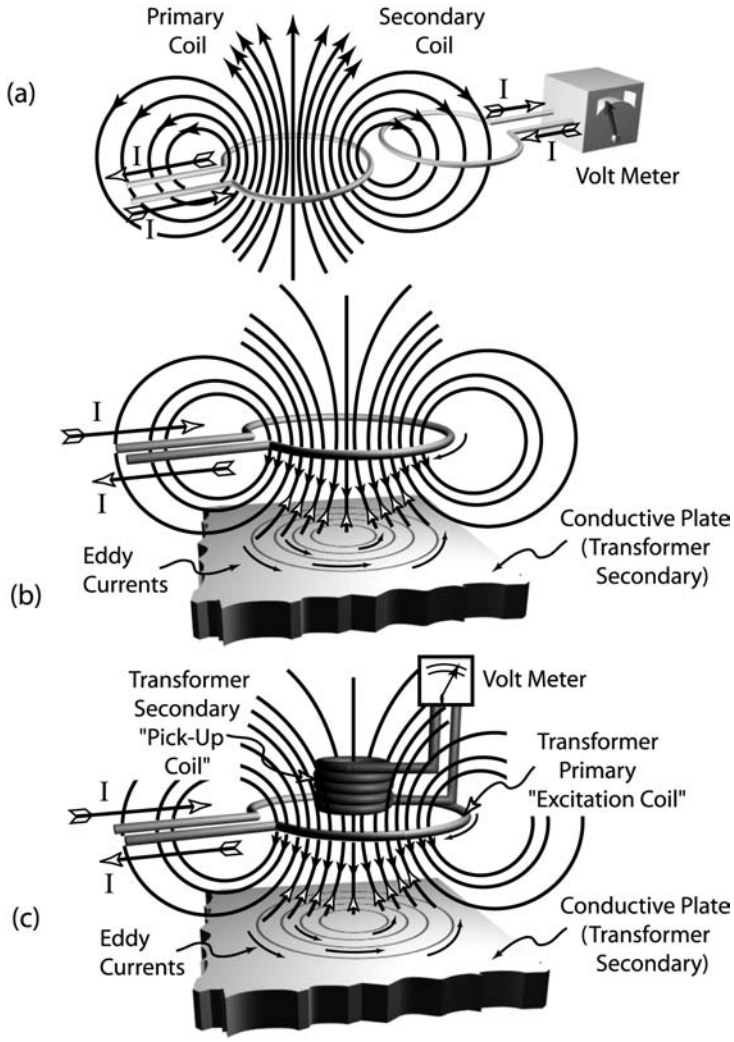


FIGURE 5.6 Inductive coupling between the primary and secondary coils: (a) two wire loops, (b) a wire coil (primary) and a conductive plate (secondary), and (c) two wire coils (a primary excitation coil and a secondary pickup coil) and a conductive plate (secondary). The latter two are common EC probe configurations.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

a current in the secondary (non-generating) loop requires energy extracted from the magnetic field of the primary (excitation) coil. Therefore, there must be a *reflected* impedance to the current flow in the primary coil; this reflected impedance is called *mutual inductance*, L_M.*

Now, what if the second conductor is not a loop but simply a conductive plate, as in Figure 5.6b? The magnetic field will still produce a current flowing in the plate, and thus reflect a mutual inductance into the primary coil; this mutual inductance is the basis for an EC device.

Another common EC probe configuration (Figure 5.6c) utilizes an excitation coil plus a separate monitoring (pickup) coil. In this case, there is a primary coil and two secondary coils (the pickup and the test piece).

5.2.2 An Eddy Current Example

The classic example presented here is taken from an article written in 1954 by Richard Hochschild, entitled "The theory of eddy current testing in one (not-so-easy) lesson" (10). Fundamentally, an EC sensor is an electromagnetic transformer. The controlling parameters for transformer (EC) measurements can be divided into electromagnetic parameters/material properties and geometric parameters (Table 5.2). The electromagnetic properties σ, μ, and f dictate the actual transduction process of magnetic induction and penetration depth of the eddy currents into the test sample. The geometric parameters of the EC coil, the sample, and the distance between the probe and sample affect the magnitude of

TABLE 5.2 Geometric and Electromagnetic Factors that Influence Eddy Current Response

Geometric factors	Electromagnetic factors
Coil parameters: loop area, number of turns, diameter of wire	σ: Electrical conductivity of the sample
Liftoff: Proximity of excitation or pickup coil to the sample (implies coupling efficiency)	μ: Magnetic permeability of the sample
Sample Geometry: Thickness compared to the skin depth, lateral dimensions, shape	f: Excitation frequency
Probe proximity to sample edges	

* Here the term "reflected" in no way implied that eddy current generation is a wave phenomenon. As we shall see in Section 5.2.5, eddy currents propagates by *diffusion*.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

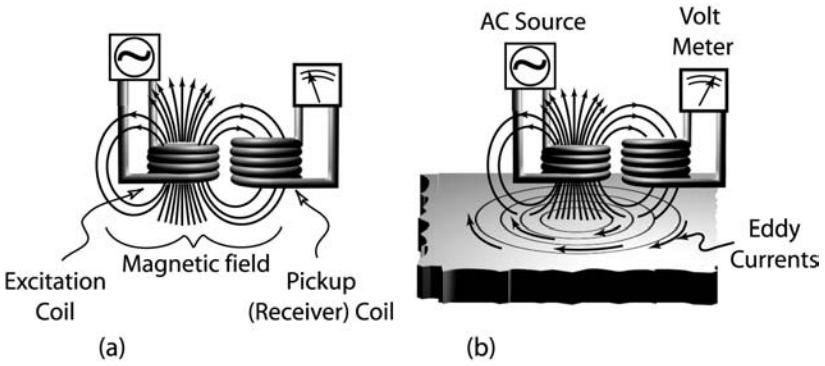


FIGURE 5.7 Eddy current probe (a) in air and (b) in proximity to a conductive sample.

the magnetic field and the flux coupling between the EC probe and the test sample.

The example opens with the construction of an EC probe using two simple coils of conductive wire (Figure 5.7a). Coil 1, the primary (excitation) coil, is connected to an AC source and produces an alternating magnetic field surrounding the coil. Coil 2, the secondary (pickup) coil, is connected to an AC voltmeter. When the two coils are placed in proximity to each other, the secondary coil captures a portion of the flux generated in the primary coil. We monitor this flux linkage (mutual inductance) between the two coils through the voltmeter on coil 2. Assume the coils are fixed in space. When no test sample is present, we label the measured inductance and resistance of the transformer as L_0 and R_0. (An ideal transformer has inductance and no resistance.) ωL_0 and R_0 are called the *impedances in free space*, where $\omega = 2\pi f$. Note that L_0 is the inductance and ωL_0 (ohms) is the *inductive or reactive impedance*.

We now place our simple EC sensor in proximity to a conductive material (Figure 5.7b) and monitor the change in impedance by measuring the change in voltage in coil 2. (The relationship of voltage to impedance will be explained later, Section 5.2.7.) The change in impedance can be explained by the *induction* of eddy currents in the test material, which creates a magnetic field that opposes and thereby reduces the magnetic field in the primary coil, and thus the field coupling to coil 2. Alternatively, we could simply argue that the introduction of the conductive sample intercepts a portion of the flux coupling coil 1 and coil 2.

From these EC sensor fundamentals, we can construct a series of simple experiments to determine the effect of conductivity, sample size (both lateral extent and thickness), material flaws, and distance between the sample and the sensor (liftoff).

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Qualitative Results

As we know, the impedance of the EC probe changes with variations in the induced magnetic flux (in the sample) and the amount of this induced flux that is reflected into the *primary* coil. The reflected field that emanates from eddy currents in the sample is a function of the sample's electromagnetic material properties, its geometry, and its position relative to the sensor.

The reflected impedance from the test sample (in response to its proximity to the EC sensor) is both inductive and resistive. The primary coil produces a magnetic field that induces eddy currents in the test specimen, which in turn produces a magnetic field that opposes the field in the primary coil. This description rationalizes the change in inductive reactance (impedance) of stored energy in the primary coil. But since real conductors have finite conductivity, some of the energy in the eddy currents is resistively lost to heat energy and not returned to the primary coil. Consequently, the presence of a real conductive (and nonferromagnetic, $\mu = \mu_0$) material to the EC sensor causes a decrease in inductance and a concomitant increase in resistance in the primary coil.

Nonferromagnetic Conductive Materials. How does the EC sensor impedance respond to test sample variations in conductivity, σ? Figure 5.8a plots the impedances R versus ωL as conductivity varies. (Z represents the magnitude of the combined impedances of R and ωL, i.e. $Z^2 = R^2 + (\omega L)^2$.)

To explain this graph, let us use the engineering principle of extremes to explore the response. First, assume the sample has $\sigma = 0$ and $\mu = \mu_0$ (nonmagnetic). Hence the sample is a nonconductor, and no eddy currents are excited. The primary coil inductance and resistance are the free space values, L_0 and R_0, and its impedances are ωL_0 and R_0. If we assume the sample is a perfect conductor, $\sigma = \infty$, and $\mu = \mu_0$, then there are no resistive eddy current losses, and again the resistance of the primary coil is R_0. The primary coil inductance depends on how much of the magnetic flux is captured by the sample and reflected back into the primary coil. Let us assume the sample is an infinite half space and 100% of the reflected flux is captured by the primary coil. Now the induced field from the eddy currents is equal and opposite to the driving field in the primary coil. Since the net primary coil field is zero, its inductance must also be zero. The graph in Figure 5.8a summarizes these extremes in conductivity. Given that the primary coil resistance is equal to R_0 at the two extremes in conductivity, we can conclude that between these conductivity extremes the resistance must grow to a maximum and then return to R_0. We argue that as the sample's conductivity increases from zero, eddy currents begin to flow and create resistive losses. At some value of conductivity, however, the increase in EC density outweighs the increase in sample conductivity. This implies that the reflected, opposing magnetic field grows with less resistive losses. Consequently, the primary coil resistance decreases.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

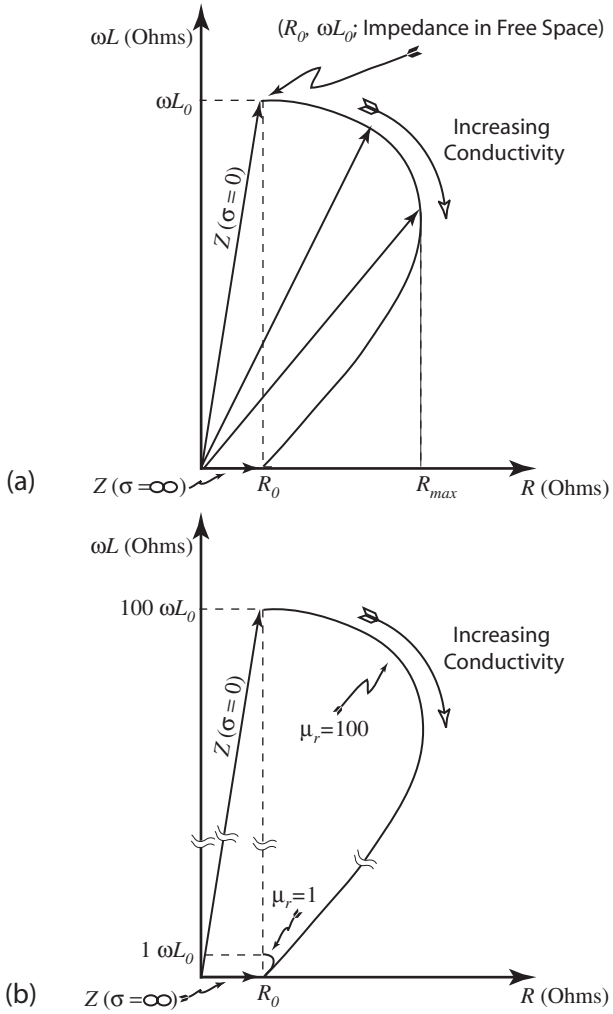


FIGURE 5.8 Response of the inductance and resistance of the primary coil the plate conductivity ranges from to $\sigma = 0$ to $\sigma = \infty$: (a) $\mu_r = 1$ and (b) $\mu_r \gg 1$.

So, how does the primary coil inductance respond to a steady increase in test sample conductivity between $\sigma = 0$ and $\sigma = \infty$? As the conductance increases, the EC density must increase. Therefore the opposing magnetic field also steadily increases, and the primary coil inductance must decrease. But the relationship is not linear with σ. The maximum change in primary coil

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

inductance for a given change in conductance occurs around the peak value in the primary coil resistance.

Ferromagnetic Conductive Materials. Conductive ferromagnetic material causes an increase in the inductance of the primary coil independent of the eddy current effect. We may recall that the common transformer takes advantage of this phenomenon to increase the mutual inductance, and therefore the coupling efficiency, between the coils. A ferromagnetic core in a transformer serves two functions: first, it increases the magnetic flux density, $\vec{B}$, in direct proportion to the magnetic permeability of the material ($B \propto \mu I$); second, it tends to guide the flux within its boundaries. In fact, in order to increase the magnetic flux density and guide the flux towards the sample, the primary coil of many EC probes is wound on a ferrite core.

In review, let us recall the specific relationship among magnetic field parameters ($\vec{B}$ and $\vec{H}$) magnetic material property (μ):

$\vec{H}$—magnetic field strength (magnetizing force), ampere turns/meter

$\mu = \mu_0\mu_r$—magnetic permeability—a material property (constant of proportionality between $\vec{H}$ and $\vec{B}$), henrys/meter

μ_0—magnetic permeability in free space ($4\pi \times 10^{-7}$ H/m)

$\vec{B}$—magnetic flux density, webers/meter^2

The constitutive relationship is

$$\vec{B} = \mu\vec{H} = \mu_0\mu_r\vec{H}$$

$$\text{magnetic flux density} = (\text{material's ability to be magnetized}) \times (\text{magnetizing force}) \tag{5.7}$$

The units of $\vec{H}$ demonstrate the direct relationship between magnetic field strength and current.

Returning to the eddy current problem from the previous section, let the test sample be a conductive ferromagnetic material $\mu_r > 100$. (Mild steel has a μ_r between 300 and 1500, depending on its magnetic history.) For a given EC sensor excitation current, the magnetic flux density, $\vec{B}$, will increase with increasing μ of the test sample. Concomitant with an increase in $\vec{B}$ is an increase in stored magnetic energy represented by an increase in inductance. Therefore, for a ferromagnetic conductive material, the total change in the EC probe inductance from L_0 is the sum of a decrease due to the reflected field from the induced eddy currents plus a large μ increase due to the effect of a large μ (Figure 5.8b).

Although the increase in $\vec{B}$ and guiding effects of a high μ core material in the EC probe are beneficial, the high μ of many conductive materials makes them difficult to inspect. In some cases, inspection of ferromagnetic materials is performed by immersing the sample in a constant magnetic saturation field that

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

is superimposed onto the alternating EC probe excitation. Under this condition the material behaves as if it were nonferromagnetic, i.e., $\mu_r = 1$.

We now have qualitative relationships among the test specimen material properties, σ, and μ, and the EC probe impedance variables, ωL and R. The following sections provide a quantitative, formal relationship among these variables.

5.2.3 Coil Impedance

The simplicity of the eddy current technique is that it is nothing more than an electromagnetic transformer. An EC measurement is performed by monitoring changes in the impedance of this transformer. The typical EC probe (electrical transformer) consists of two or more coils—a primary (excitation) coil and one or more secondary coils. In the previous example, the probe consisted of a primary and two secondary coils—the test sample and a pickup coil (Figure 5.7b).* The primary coil was used for excitation with an impressed current, i.e., a fixed AC current. The secondary coil was used to monitor the mutual or reflected impedance changes caused by the test sample coupling with the magnetic flux generated by the primary coil.

Whether the EC transducer is a single- or multiple-coil probe, there are only two measurable parameters: current and voltage. By measuring current and voltage, we can deduce the resistive and inductive (reactive) impedance of a particular coil. As previously discussed, these values of resistance and inductance are dictated by the entire transformer system, including the mutual induction, L_M, with the test sample. Generally, the EC probe coils are fixed and any change in the resistance and inductance is directly related to the test sample. But other parameters, such as temperature, can also change the probe impedance.

Two representations are commonly used to discuss electrical impedances (or circuit analysis in general): sinusoidal and phasor notations.

Resistive Impedance

In an EC system there are two notions of resistive impedance, or resistance: (a) direct resistance of a wire, and (b) reflected resistance. The first is the common resistance that appears as a material property in any wire or material. The second is the additional resistance found in a coil due to mutual coupling and eddy current losses in the secondary coil(s).† In both cases, the resistance is a function of the material resistivity and frequency of operation. Reflected resistance, however, also depends on the coupling of the reflected magnetic field back into

*Although it functions as one, the sample is generally not considered an EC probe coil.

†To reduce eddy current resistive losses, transformer manufactures make the coupling cores by laminating a thin sheet of ferromagnetic metal or ground iron oxides that is sintered like a ceramic to form ferrites.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

the primary coil. The importance of this point is that one needs to know the coupling factor completely in order to determine resistive properties of the test sample. We will return to this issue later, in Liftoff in Section 5.2.7.

The resistance of the EC probe in free space is designated as R_0, regardless of the number of coils. The resistance of the probe in the presence of the test sample is designated as R. Therefore, the change in resistance caused by the reflected resistance from the sample is $R - R_0$.

Inductive Impedance

Just as it has both direct and reflected resistance, an EC probe (transformer) also has both self- and mutual inductance. An inductor impedes changes in the flow of current. Intuitively we know that the impedance must be a function of frequency as well as the value of the inductor, L, that accounts for geometric factors. Recall that impedance is defined as $Z = V/I$. Assume the impressed current on an inductor is sinusoidal, i.e.,

$$I = I_0 \sin \omega t \tag{5.8}$$

From Eq. (5.5):

$$\begin{aligned} V_{\text{EMF}} &= -L\frac{d(I_0 \sin \omega t)}{dt} \\ &= -\omega L I_0 \cos \omega t = -\omega L I_0 \sin(\omega t + 90^\circ) \end{aligned} \tag{5.9}$$

Therefore, the inductive impedance associated with an inductor is

$$Z = \frac{V_{\text{Applied}}}{I} = \omega L \frac{I_0 \cos \omega t}{I_0 \sin \omega t} = \omega L \cot \omega t \tag{5.10}$$

This representation in time, however, is rarely used. Instead, we generally use either root mean square (rms) values or phasor notation:

$$Z = \frac{V_{\text{rms}}}{I_{\text{rms}}} = \omega L \frac{I_{\text{rms}}}{I_{\text{rms}}} = \omega L \tag{5.11}$$

5.2.4 Phasor Notation and Impedance

Phasor notation represents sinusoidal functions in a convenient form that facilitates mathematical manipulation, i.e., derivatives and integrals. The phasor is a complex number that includes both the magnitude and the phase of sinusoids of the same frequency. They are related through Euler's equation:

$$e^{j\theta} = \cos \theta + j \sin \theta \tag{5.12}$$

Recall that $j = \sqrt{-1}$ and simply implies a counterclockwise 90° phase rotation.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

For example, let $V = V_0 \cos(\omega t)$. If this sinusoid is written using Euler's equation,

$$V_0 e^{j\omega t} = V_0 \cos \omega t + jV_0 \sin \omega t \tag{5.13}$$

then our original voltage, V, is represented by the real part (first term) of the phasor. Figure 5.9 shows graphically the relationship between the phasor and the sinusoidal voltage. A sine function would be represented by the imaginary part (second term) of the phasor. In most applications, a sinusoidal function is represented by a phasor with the understanding that only the real part is used.* Reworking the inductor problem, let

$$I = I_0 e^{j\omega t} \tag{5.14}$$

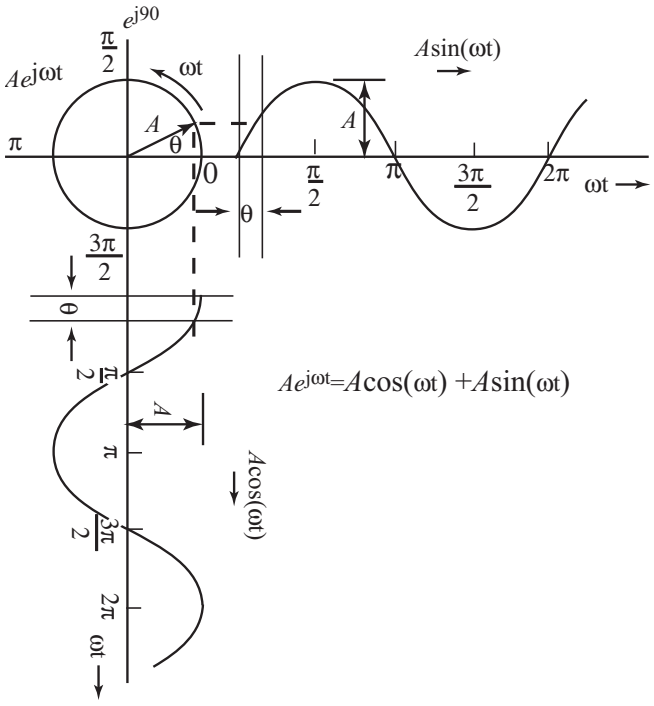


FIGURE 5.9 Relationship among phasors and sinusoids.

*The convention of using the real part of the phasor to represent the sinusoid is arbitrary.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Then

$$\begin{aligned} V_{\text{emf}} &= -L\frac{dI}{dt} = -L\frac{d(I_0 e^{j\omega t})}{dt} \\ &= -j\omega LI = -j\omega LI_0 e^{j\omega t} \\ &= -j\omega L\mathscr{R}e(I_0\cos(\omega t) + jI_0\sin(\omega t)) \\ &= -j\omega LI_0\cos(\omega t) \end{aligned} \tag{5.15}$$

($\mathscr{R}e$ implies the real part and $\mathscr{I}m$ implies the imaginary part.) Alternatively, if we notice that

$$j = \cos(90^\circ) + j\sin(90^\circ) = e^{j90^\circ} \tag{5.16}$$

$$\begin{aligned} V_{\text{emf}} &= -j\omega LI = -\omega L(I_0 e^{j\omega t} e^{j90^\circ}) \\ &= -\omega LI_0\cos(\omega t + 90^\circ). \end{aligned} \tag{5.17}$$

In all versions, the voltage of an inductor leads the current by 90°.

Let us return to the impedance of an EC probe. A simple circuit model would include an inductor and resistor in series, Figure 5.10a. We apply Kirchhoff's voltage law to mathematically represent this series circuit:

$$L\frac{dI}{dt} + RI = V_{\text{applied}} \tag{5.18}$$

The voltage drop across the inductor plus the resistor equals the applied voltage. Let $I = I_0 e^{j\omega t}$; then

$$\begin{aligned} j\omega LI_0 e^{j\omega t} + RI_0 e^{j\omega t} &= V_{\text{applied}} \\ j\omega LI + RI &= V_{\text{applied}} \end{aligned} \tag{5.19}$$

According to Ohm's law, the impedance seen by the voltage source to the flow of current is

$$Z = \frac{V_{\text{applied}}}{I} = R + j\omega L \tag{5.20}$$

What does it mean that the impedance has a real and an imaginary component? Associated with the voltage drop across the real resistive impedance, R, is an energy loss of I^2R to heat, whereas the imaginary inductive reactance or reactive impedance, ωL, stores the energy associated with the voltage drop across the inductor in the form of a magnetic field. As previously discussed, this stored energy can be retrieved by collapsing the field. As indicated by ω times L, the inductive impedance increases with increasing frequency.

Figure 5.10b shows a sinusoidal representation of the current-voltage relationships in an EC coil modeled as a resistor and an inductor in series. I represents the current common to both the resistor and the inductor; V_R and V_L

Copyright © Marcel Dekker, Inc. All rights reserved.

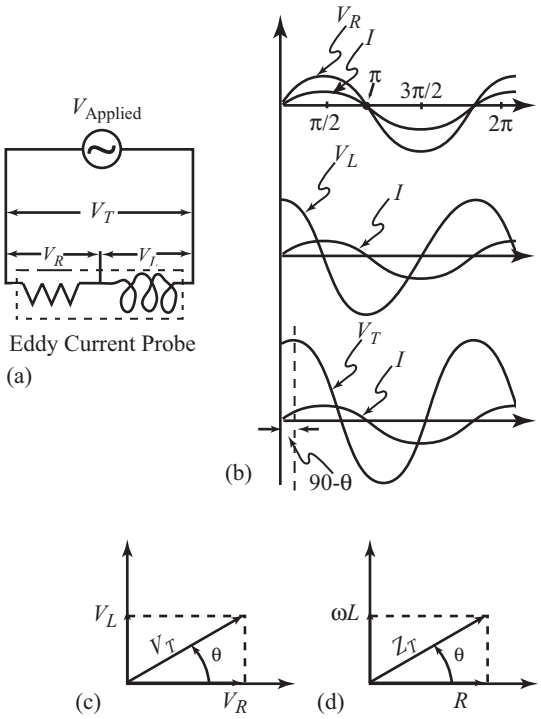


FIGURE 5.10 Eddy current probe modeled by a *R–L* circuit: (a) circuit model, (b) relationship among V_R, V_L, V_T, and I, and (c) phasor representation of V_T and (d) Z_T.

are the voltages measured directly across the terminals of the resistor and the inductor, respectively. The lower graph represents total voltage V_T across the combination of the resistor and the inductor. To show the phasor representation of these relationships, let

$$I = I_0 e^{j\omega t} \tag{5.21}$$

Then

$$\begin{aligned} V_R &= IR \\ V_L &= L\frac{dI}{dt} = j\omega L I_0 e^{j\omega t} = j\omega LI \\ V_T &= I(R + j\omega L) \end{aligned} \tag{5.22}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

This total voltage change across the EC probe (modeled by a resistor and inductor in series) can be plotted on a phasor diagram, where the voltages are represented by their peak values (Figure 5.10c). θ represents the phase shift between the total voltage, V_T, and the current, I. Note that $\theta = 0^\circ$ for V_R and $\theta = 90^\circ$ for V_L. As indicated in the diagram, these voltages are vector quantities—each with a magnitude and a phase angle. Therefore,

$$\vec{V}_T = |\vec{V}_T| e^{j\theta}$$

where

$$\begin{aligned} |\vec{V}_T| &= I\sqrt{R^2 + (\omega L)^2} \\ \theta &= \tan^{-1}\frac{\omega L}{R} \end{aligned} \tag{5.23}$$

Now, that we have the voltage-current (the measurable quantities) relationships, how do we relate these to the impedances from which EC methods determine material properties, geometries, flaws, etc.? Because the current is common to both the resistor and the inductor,

$$V_T = IZ_T \tag{5.24}$$

$$Z_T = R + j\omega L \tag{5.25}$$

Because Z_T is directly proportional to V_T, we can plot Z_T by measuring and knowing I. This is shown in Figure 5.10d, where

$$\vec{Z}_T = |\vec{Z}_T| e^{j\theta}$$

and

$$\begin{aligned} |\vec{Z}_T| &= \sqrt{R^2 + (\omega L)^2} \\ \theta &= \tan^{-1}\frac{\omega L}{R} \end{aligned} \tag{5.26}$$

Note that θ is the same in both plots.

If we included capacitance in a circuit, the capacitive impedance would be determined by the relationship

$$V = \frac{1}{C}\int Idt = \frac{I}{j\omega C} = \frac{-jI}{\omega C} \tag{5.27}$$

$$Z_C = \frac{-j}{\omega C} \tag{5.28}$$

Indicated by the imaginary number, j, the capacitive reactance implies that capacitors are also energy storage elements. (Capacitors store energy in the electric field between two charged conductors.) Although often negligible, an

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

TABLE 5.3 Discrete Electrical Components and Their Associated Energy

Impedance	Real ($\mathscr{R}e$ or imaginary ($\mathscr{I}m$)	Energy form
R	$\mathscr{R}e$	Energy lost to heat energy
ωL	$\mathscr{I}m$	Energy stored in magnetic field
$1/\omega C$	$\mathscr{I}m$	Energy stored in electric field

inductor itself always possesses a parasitic turn-to-turn capacitance as well as resistance.

Table 5.3 summarizes the resistive and reactive impedances and their associated energy. In circuit impedance calculations, these elements may occur in both series and parallel combinations.

5.2.5 Eddy Current Density and Skin Depth (11)

Consider the typical EC setup: a primary/secondary coil in proximity to a good conductor. For the moment, assume the conductor is an infinite half-space. With the excitation of the primary coil, eddy currents appear in the conductor in proportion to the impinging (captured) magnetic flux. The magnitude of the eddy currents induced at any given point in the conductive part is directly proportional to the strength of the magnetic field that reaches that point.

EC density is distributed both laterally (parallel to the surface) and through the depth of the material. In lateral distribution, eddy currents are at a maximum directly beneath the EC probe, and decrease rapidly with increasing lateral distance from the probe. Similarly, the current distribution varies with any tilt of the probe. (Probe tilt is a common problem in EC measurements.)

To approximate the EC density change with depth, we assume the magnetic field is perpendicular to the surface. Using the magnetic field at the surface as a reference, let us look at what happens when the field is converted to eddy currents through the thickness of a conductive material. Producing eddy currents requires energy, which is provided by the magnetic field. As the magnetic field diffuses into the material, the available energy to produce eddy currents decreases.

Another way of looking at change with depth starts from an induction stand point. When eddy currents are induced, they produce secondary magnetic fields. The secondary fields, which oppose the primary field, partially shield the primary field from further penetration into the depth of the material.

Clearly, the EC density must decay with increasing depth. The term *skin depth* describes this decay. The absorption of energy or creation of an opposing field suggests that the skin depth is dependent on the material properties of

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

conductivity (σ), magnetic permeability (μ), and the frequency of oscillation (f) of the magnetic field from the primary coil. Recall that absorption is a function of frequency (Section 3.2.5).

Assume the primary coil has a sinusoidal excitation; the induced EC density, J, at any depth x in the material is given by

$$J = J_0 e^{-(x\sqrt{\pi\mu\sigma f})} \sin(\omega t - x\sqrt{\pi\mu\sigma f})$$

or

$$J_0 e^{-(1+j)(x\sqrt{\pi\mu\sigma f})} e^{j\omega t} \tag{5.29}$$

where J_0 is the current density at the surface. Referenced to J_0, the term $e^{-(x\sqrt{\pi\mu\sigma f})}$ determines the EC decay with increasing depth, and $e^{-j(x\sqrt{\pi\mu\sigma f})}$ denotes the increasing phase lag with increasing depth. We define the skin depth, δ, to be the depth, x, at which the current density has decreased to e^{-1} (36.8%) from the value of J_0 at the surface and the phase has changed by 1 rad (57.29°). Therefore,

$$\delta = \frac{1}{\sqrt{\pi\mu\sigma f}} \tag{5.30}$$

$$J = J_0 e^{-(x/\delta)} e^{-j(x/\delta)} e^{-\omega t} \tag{5.31}$$

Thus at $x = 1\delta$, $|J| = 0.368\, J_0$; at $x = 2\delta$, $|J| = 0.135 J_0$; and at $x = 5\delta$, the current density is less than 1% of J_0. At the depth where $|J| = 0.05 J_0$, 3δ, is called the *effective depth of penetration*. Figure 5.11a plots the decay of the current density as a function of depth into the conductor. The plot axes are normalized current density, J/J_0, and normalized depth, x/δ.*

Figure 5.11a plots the magnitude of the sinusoidally varying EC density as the magnetic field propagated (diffuse) into the material. We also need to consider the phase lag of the currents as a function of depth. The phase angle is typically referenced to the phase at the surface:

$$\theta = x\sqrt{\pi\mu\sigma f} = \frac{x}{\delta} \tag{5.32}$$

This linear relationship of the change in phase with an increase in depth is plotted in Figure 5.11b.

Example: Calculate the phase change in degrees of the eddy current from the surface to a depth of 5δ.

*This skin depth, typically referred to as the *standard skin depth*, was calculated using a plane wave approximation; the magnetic field was assumed normal to the surface. The actual skin depth could be calculated from Maxwell's equations. Nonetheless, the standard skin depth is commonly used to estimate the depth of penetration of eddy currents for a particular material and application.

Copyright © Marcel Dekker, Inc. All rights reserved.

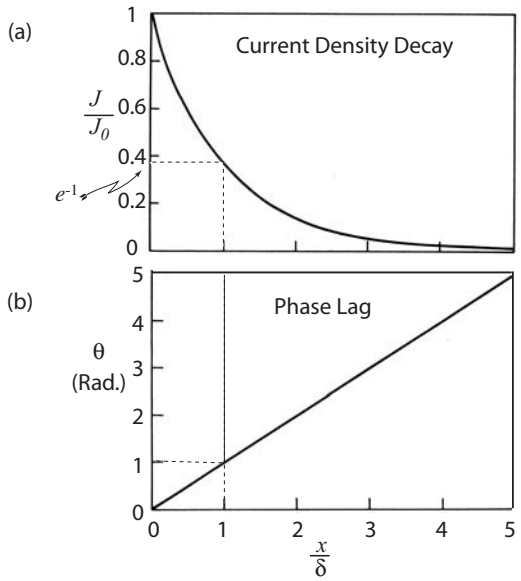


FIGURE 5.11 Eddy current (a) decay and (b) phase lag with depth into a material normalized to the current density induced at the surface.

Solution: From Eq. (32),

$$\theta = \frac{x}{\delta} = \frac{5\delta}{\delta} = 5 \text{ rad}$$

$$\theta = \frac{180^\circ}{\pi} 5 = 286^\circ$$

Therefore, at a depth where the current density has dropped to less than 1% of its value at the surface (5δ), the phase has changed less than one full cycle, 2π or 360°. Consequently, eddy currents are said to diffuse into the material rather than act like waves propagating through the material, i.e., they act more like heat flow than ultrasonic or microwaves.

5.2.6 Maxwell's Equations and Skin Depth (Optional)

We have described the induction of a magnetic field by a current flowing in a coil. This field then propagates as a wave into space and, in our case, is incident onto a conductive medium: the test specimen. The incident magnetic field gives rise to eddy currents, which in turn create a secondary (reflected) magnetic field that opposes the original incident field. This reflected field causes a change in the

Copyright © Marcel Dekker, Inc. All rights reserved.

impedance characteristics of the primary and/or the pickup coils. This section details the mathematical relationship between the incident magnetic field and induced eddy currents as a function of depth within the conductive test specimen, i.e., skin depth. Again, we will assume a plane wave approximation.

James Clark Maxwell developed a set of equations relating electromagnetic phenomena and its commitment interaction with materials. The four Maxwell's equations that relate electromagnetic fields and particles to material properties are

$$\nabla \cdot \vec{D} = \rho \quad (\rho - \text{electric charge}) \tag{5.33}$$

$$\nabla \cdot \vec{H} = 0 \tag{5.34}$$

$$\nabla \times \vec{E} = -\frac{\partial \mu \vec{H}}{\partial t} \text{ (Faraday's law of induction; } \vec{E} - \text{electric field)}, \tag{5.35}$$

and

$$\nabla \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t} \text{ (Ampere's circuital law)} \tag{5.36}$$

the constituative relationships are

$$\vec{B} = \mu \vec{H}$$

$$\text{(magnetic flux density} = \text{magnetic permeability} \times \text{magnetic field)} \tag{5.37}$$

$$\vec{D} = \varepsilon \vec{E}$$

$$\text{(electric flux density} = \text{electrical permittivity} \times \text{electric field)} \tag{5.38}$$

Because the eddy currents are induced by the magnetic field, we will first determine how the magnetic field, $\vec{H}$, decays with increasing depth into a conductor. We begin with Maxwell's fourth equation (Ampere's law),

$$\nabla \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t}$$

$$= \text{conduction current} + \text{displacement current} \tag{5.39}$$

that relates the magnetic field to the induced currents within the medium. $\vec{J}$ represents the conduction current associated with *free* electrons, and $\partial \vec{D}/\partial t$ represents the displacement current associated with *bound* charge, i.e., ionic or polar molecules. In order to solve this equation, we need to write $\vec{J}$ and $\partial \vec{D}/\partial t$ in terms of $\vec{H}$. Combining Ohm's law $\vec{E} = \vec{J}/\sigma$ and $\vec{D} = \varepsilon \vec{E}$ from the constituative equation, we have

$$\nabla \times \vec{H} = \sigma \vec{E} + \varepsilon \frac{\partial \vec{E}}{\partial t} \tag{5.40}$$

We can simplify this equation by applying the vector identity

$$\nabla \times \nabla \times \vec{V} = \nabla(\nabla \cdot \vec{V}) - \nabla^2 \vec{V} \tag{5.41}$$

Thus

$$\nabla \times \nabla \times \vec{H} = \nabla(\nabla \cdot \vec{H}) - \nabla^2 \vec{H} \tag{5.42}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Applying Maxwell's second equation, $\nabla \cdot \vec{H} = 0$, we have

$$-\nabla^2 \vec{H} = \nabla \times \left(\sigma + \varepsilon \frac{\partial}{\partial t}\right)\vec{E} \tag{5.43}$$

Substitution of Maxwell's third equation, $\nabla \times \vec{E} = -\partial\mu\vec{H}/\partial t$, yields

$$-\nabla^2 \vec{H} = -\left(\sigma + \varepsilon \frac{\partial}{\partial t}\right)\mu \frac{\partial \vec{H}}{\partial t} \tag{5.44}$$

where all the variables are in terms of $\vec{H}$. Assuming the magnetic field incident onto the test sample is a plane wave—$\vec{H} = \vec{H}_0 e^{j(\omega t - \vec{k}\cdot\vec{r})}$*—then the first and second derivative of $\vec{H}$ with respect to time are

$$\frac{\partial \vec{H}}{\partial t} = \frac{\partial(\vec{H}_0 e^{j\omega t} e^{-j\vec{k}\cdot\vec{r}})}{\partial t} = j\omega \vec{H}_0 e^{j\omega t} e^{-j\vec{k}\cdot\vec{r}} = j\omega \vec{H} \tag{5.45}$$

and

$$\frac{\partial^2 \vec{H}}{\partial t^2} = j^2\omega^2 \vec{H} = -\omega^2 \vec{H} \tag{5.46}$$

Making these substitutions, we have

$$\nabla^2 \vec{H} = (\, j\omega\sigma - \omega^2\varepsilon)\mu\vec{H} \tag{5.47}$$

Letting $k^2 = j\omega\mu\sigma - \omega^2\mu\varepsilon = j\omega\mu(\sigma + j\omega\varepsilon)$, we recognize the form of the plane wave equation (Helmholtz equation):

$$\nabla^2 \vec{H} - k^2 \vec{H} = 0 \tag{5.48}$$

Note that k is the *propagation constant* that determines both the decay and phase of the eddy currents as they diffuse into a material. This propagation constant is determined solely by the material properties and frequency.

Now that we have the wave equation for a magnetic field, assume the conducting sample is an infinite half-space whose surface is in the yz plane at $x = 0$. Let the magnetic field point in the z-direction perpendicular to the propagation direction x—a *transverse wave*. Therefore, the wave equation for this geometry becomes

$$\frac{d^2 H_z}{dx^2} - k^2 H_z = 0 \tag{5.49}$$

Using the plane wave approximation, let

$$H_z = H_{0z} e^{j(\omega t - kx)} \tag{5.50}$$

*See Section 3.2.1 for details of plane waves.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

where $k = \sqrt{j\omega\mu(\sigma + j\omega\varepsilon)}$ and H_{0_z} is the amplitude of the magnetic field at the surface of the conductor. For a good conductor $\sigma \gg \omega\varepsilon$ (at typical operating frequencies) and recalling that $\sqrt{j} = (1+j)/\sqrt{2}$,

$$k \approx \sqrt{j\omega\mu\sigma} = \frac{1+j}{\sqrt{2}}\sqrt{\omega\mu\sigma} \tag{5.51}$$

and by separating the propagation constant, k, into real (k') and imaginary (k'') components, we have

$$k = k' + jk'' \approx \frac{1}{\sqrt{2}}\sqrt{\omega\mu\sigma} + j\frac{1}{\sqrt{2}}\sqrt{\omega\mu\sigma} \tag{5.52}$$

Thus

$$H_z = H_{0z}e^{-\sqrt{\frac{\omega\mu\sigma}{2}}x}e^{-j\sqrt{\frac{\omega\mu\sigma}{2}}x}e^{+j\omega t} = H_{0z}e^{-k'x}e^{-jk''x}e^{+j\omega t} \tag{5.53}$$

The real exponential term, $e^{-k'x}$, represents the *rate of decay* of the magnetic field with increasing depth into the conductive material while the imaginary exponential term, $e^{-jk''x}$, details the phase progression of the field with increasing depth. Therefore, k' is an attenuation constant and k'' is the wave number (spatial frequency).

Intuitively we know that if the magnetic field, which induces the eddy currents, decays with increasing depth, then the amplitude of the eddy current should also decrease accordingly. Similarly, the phase of the induced eddy current should lag with depth. Verifying this insight begins with Ampere's circuital law for a conductive medium, i.e., $\vec{J} \gg \partial\vec{D}/\partial t$;

$$\nabla \times \vec{H} = \vec{J} \tag{5.54}$$

Continuing with the geometry of the example above, we have the current as a function of increasing depth x, as

$$-\frac{dH_z}{dx} = -J_y \tag{5.55}$$

$$-J_y = -\frac{d}{dx}H_{0z}e^{-\sqrt{\frac{\omega\mu\sigma}{2}}x}e^{-j\sqrt{\frac{\omega\mu\sigma}{2}}}e^{+j\omega t} \tag{5.56}$$

and therefore,

$$\begin{aligned} J_y &= (1+j)\sqrt{\frac{\omega\mu\sigma}{2}}H_{0z}e^{-\sqrt{\frac{\omega\mu\sigma}{2}}x}e^{-j\sqrt{\frac{\omega\mu\sigma}{2}}x}e^{+j\omega t} \\ &= J_0e^{-\sqrt{\frac{\omega\mu\sigma}{2}}x}e^{-j\sqrt{\frac{\omega\mu\sigma}{2}}x}e^{+j\omega t} = J_0e^{-k'x}e^{-jk''x}e^{+j\omega t} \end{aligned} \tag{5.57}$$

J_0 represents the current density at the surface of the conductor. A comparison of the current density (Eq. (5.57)) to the magnetic field (Eq. (5.53)) shows identical decay rates and phase lag with increasing depth.

Copyright © Marcel Dekker, Inc. All rights reserved.

We define the skin depth δ as the depth x at which the amplitude has decreased to $e^{-1} \approx 0.368$ from the level of current density at the surface, i.e., $J_z = 0.368J_0$ and/or $H_z = 0.368H_{0z}$. We recognize e^{-kx} as the amplitude term. Therefore,

$$e^{-1} = e^{-k'\delta} \tag{5.58}$$

$$\delta = \sqrt{\frac{2}{\omega\mu\sigma}} = \sqrt{\frac{1}{\pi f \mu\sigma}} \tag{5.59}$$

5.2.7 Impedance Plane Diagrams

Because an EC sensor is a transformer, we can measure only two quantities directly: voltage and current. But from these two quantities, we can infer all the desired sample information: electromagnetic material properties, part geometry, and part integrity.

We typically present the output of the EC sensor as the ratio of the voltage to the current, i.e., the impedance of the transformer. As the physical parameters (conductivity, magnetic permeability, etc.) of the test sample change, impedance changes. An *impedance plane diagram* maps the changes in EC coil impedance as a function of the test sample variables. This section discusses the basic construction of an impedance plane diagram (plot) as the impedance changes for an EC probe. Changes caused by variations in material or geometric variables of common interest are:

Conductivity σ
Magnetic permeability μ
Liftoff
Sample thickness or layer thickness, and/or
Sample edges

An effective EC probe allows us to discern which parameter has changed. By constructing an impedance plane diagram, we can determine the best operating frequency for this purpose.

The effect of sample discontinuities (cracks, coatings, porosity, bonds, etc.) will be discussed separately when we discuss the detail of an actual inspection (Section 5.4). Because the type of probe affects the shape of the impedance plane, we assume a flat-surface probe in this section, unless otherwise stated. The distinctions for inspection of tubes and rods using encircling and interior probes are explained in Section 5.2.8.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Normalized Impedance

In Hochschild's example (Section 5.2.2), we detailed the sensor's impedance response to variations in sample conductivity using an impedance plane plot of the *real* resistive impedance, R, versus the *imaginary* reactive impedance, ωL (Figure 5.8). While this impedance plane (resistance vs. reactance) format is useful, it requires a new graph for even a simple change in frequency. Instead we can create a single normalized universal curve that normalizes the effects of frequency, primary coil reactance, conductivity, and mean probe diameter.

A universal impedance plot is derived from a nonnormalized plot of a given probe response (Figure 5.12a). Note the change in variable. For convenience the variable for total resistance (coil plus reflected resistance) will be represented by R', and R will now represent the reflected resistance only. Similarly, the total impedance is now represented with a prime. Given the variable change, let

$$Z' = R' + j\omega L \tag{5.60}$$

represent the total impedance for this setup. The first step in normalization is to subtract the free space probe resistance, R_0, from total resistance, R', and identify the reflected resistance as

$$R = R' - R_0 \tag{5.61}$$

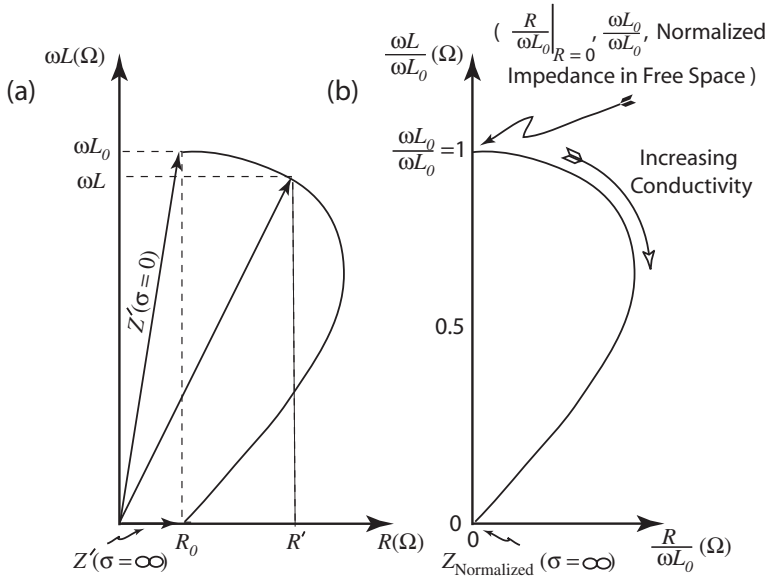


FIGURE 5.12 (a) Nonnormalized impedance plane diagram. Note R' represents the total resistance. (b) Normalized impedance plane diagram.

Copyright © Marcel Dekker, Inc. All rights reserved.

For simplicity, we identify a new impedance as

$$Z = (R' - R_0) + j\omega L = R + j\omega L \tag{5.62}$$

The final normalizing step is to divide Z by the probe's free space reactive impedance:

$$Z_{\text{normalized}} = \frac{Z}{\omega L_0} = \frac{R + j\omega L}{\omega L_0} = \frac{R}{\omega L_0} + j\frac{\omega L}{\omega L_0} \tag{5.63}$$

We now plot a universal EC probe impedance response on the complex plane (Figure 5.12b). The normalized inductive reactance, $\omega L/\omega L_0$, goes from 1 to 0 ohms as, for example, the conductivity of the sample ranges from 0 to ∞ $1/\Omega$m or siemens/m. (The inductive reactance approaching zero ohms assumes that the inductive coupling between the probe and the sample approaches 100%.)

Many texts depict a circular instead of a comma-shaped curve, but this idealization neglects the effects of skin depth. On very thin sheets or thin tubes, however, the circular impedance curve approximation works well.

Test Sample Conductivity and Frequency of Operation

The impedance plane curve simply maps the reflected impedance from the test sample back into the primary coil (or a pickup coil) of the EC probe. As we have seen, the reflected impedance is a function of how much of the induced eddy currents are lost to heat (sample resistance) versus the amount of eddy currents that induce the secondary (reflected) magnetic field that opposes the incident fields. (For this section we assume 100% coupling of the magnetic fields between the probe and the sample.)

In Hochschild's example (Section 5.2.2), we detailed the influence of the test sample conductivity on the reflected impedance in terms of real resistive losses (R') and imaginary reactive energy storage (ωL). (The probe oscillator frequency, ω, is assumed to be constant.) Similarly, when we use the normalized impedance plane diagram, the reflected resistance, R, varies from $0\,\Omega$ to a maximum and then returns to $0\,\Omega$ as the *test* material conductivity σ ranges from 0 to ∞ siemens/m. Concomitantly, the normalized inductive reactance $\omega L/\omega L_0$ varies from 1 to 0 as σ ranges from 0 to ∞ siemens/m. If the material conductivity equals zero, then no eddy currents are generated and the probe impedance is not influenced by the presence of the test sample, $Z_{\text{normalized}} = 0 + j1$ on the normalized impedance plane. At the other extreme, when $\sigma \Rightarrow \infty$ and $Z_{\text{normalized}} \Rightarrow 0 + j0$, there are no resistive losses, and the eddy currents induce a reflected magnetic field equal and opposite to the magnetic field generated by the primary coil, i.e., the magnetic fields cancel each other completely ($\omega L = 0$). Between these two extremes as the test sample conductivity increases from zero, the resistive losses increase as more and more eddy currents are generated. At some point, the increase in EC density for a given increase in

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

sample conductivity outweighs the resistive losses, and R begins to reduce in value, eventually approaching zero as the conductivity approaches infinity.

Interestingly, if the oscillator frequency that drives the primary coil is varied from 0 to ∞, the probe impedance response to a sample with fixed conductivity will be the same as it will be if the conductivity is varied at a fixed frequency. The heuristic argument is that when $\omega = 0$, no magnetic field is generated (no eddy currents) and therefore, $R = 0$ and $\omega L/\omega L_0 = 1$. As $\omega \Rightarrow \infty$, the eddy currents are confined completely to the sample surface and experience no resistive losses; this is equivalent to $\sigma \Rightarrow \infty$. These high-frequency surface-bound eddy currents induce a magnetic field equal and opposite to the excitation field; thus the fields cancel one another and $\omega L/\omega L_0 \Rightarrow 0$. Figure 5.13 shows the impedance change of five metals (titanium, stainless steel, lead, aluminum, and copper) as the EC probe frequency is increased. Note that the impedance trajectory is the same as it would be if the conductivity of the material were increased instead. Table 5.4 lists the resistivity in $\mu\Omega\,$cm and conductivity in %IACS of common materials.

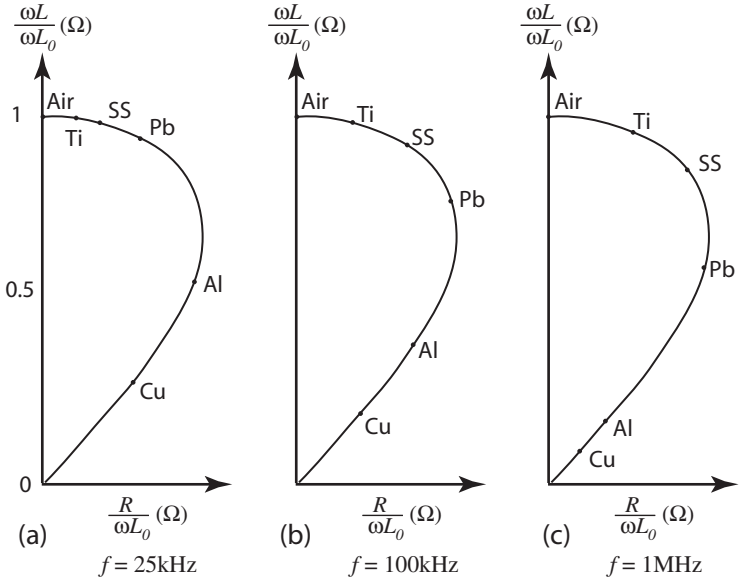


FIGURE 5.13 Effect of EC probe excitation frequency on the impedance of various conductive materials. (a) $f = 25$ kHz, (b) $f = 100$ kHz, and (c) $f = 1$ MHz.

Copyright © Marcel Dekker, Inc. All rights reserved.

TABLE 5.4 Resistivity and Conductivity of Common Materials

Material	Resistivity ($\mu\Omega$ cm)	Conductivity % IACS
Ti-6Al-4V	172.00	1.00
Hastelloy-X	115.00	1.50
Inconel 600	100.00	1.72
Stainless Steel 316	74.00	2.33
Stainless Steel 304	72.00	2.39
Zircalloy-2	72.00	2.40
Titanium-2	48.56	3.55
Monel	47.89	3.60
Zirconium	40.00	4.30
Copper nickel 70-30	37.48	4.60
Lead	20.65	8.35
Copper nickel 90-10	18.95	9.10
Cast Steel	16.02	10.70
Phospher bronze	16.00	11.00
Aluminum bronze	12.32	14.00
Brass	6.20	28.00
Tungsten	5.65	30.51
Aluminum 7075-T6	5.39	32.00
Aluminum 2024-T4	5.20	30.00
Magnesium	4.45	38.60
Sodium	4.20 @ 0°C	41.50
Aluminum 6061	4.10	42.00
Gold	2.46	70.00
Copper	1.72	100.00
Silver	1.64	105.00

Characteristic (Length) Parameter

Because the sample conductivity and excitation frequency exhibit the same influence on the probe impedance, the impedance plane diagram could be plotted as $\sqrt{\omega\sigma\mu_0}$ for a nonmagnetic material (the effect of magnetic materials, $\mu_r \gg 1$, will be addressed in the next section), thus accounting for the variation of either σ or ω on the same curve. Further to complete the normalization of the impedance plane, $\sqrt{\omega\sigma\mu_0}$, which has units of length, is normalized by the mean coil radius $\bar{r}$. $\bar{r}\sqrt{\omega\sigma\mu_0}$ or $\bar{r}^2\omega\mu_0\sigma$ is called the *characteristic length (number)* and is attributed to Deeds and Dodds. Thus the effect of varying $\bar{r}$, σ, or ω traces the same curve on the impedance plane (Figure 5.14). Not surprisingly, the characteristic length is directly related to the skin depth ($\delta = \sqrt{2}/\sqrt{\omega\sigma\mu_0}$). The depth of penetration of the eddy currents into the sample is fundamental to the argument for the

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

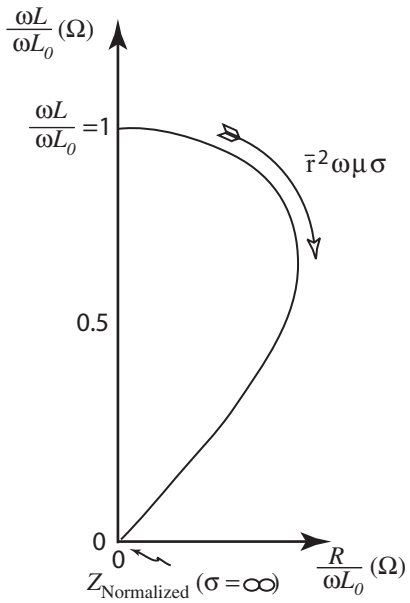


FIGURE 5.14 Normalized impedance plane plotted using the characteristic (length) parameter $\bar{r}^2\omega\sigma\mu_0$.

impedance trajectory. Thus $\sqrt{\omega\sigma\mu_0}$ with its units of length is called the *limiting length*.

A common alternative normalizing parameter for encircling coils is the mean radius, $\bar{a}$, of a tube or bar, i.e., $\bar{a}\sqrt{\omega\sigma\mu_0}$. Note that both $\bar{r}\sqrt{\omega\sigma\mu_0}$ and $\bar{a}\sqrt{\omega\sigma\mu_0}$ have units of dimensionless length—meters/meters.

Ferromagnetic Conductive Material

EC inspection is as common for magnetic conductive materials (primarily iron and nickel based alloys) as for nonmagnetic conductive materials. How, then, does magnetic material influence the impedance plane? Recall that for a given applied magnetic field H, the magnetic flux density B within the material increases proportionally to the magnetic permeability μ of the material, i.e., $B = \mu H$, $\mu = \mu_0\mu_r$. And because the inductance L is proportional to B, testing of a material with $\mu_r > 1$ will increase the reactive impedance such that $\omega L/\omega L_0 > 1$. Note that in the presence of the test specimen with $\mu_r > 1$, the probe inductance (reactive impedance) increases—unlike the EC effect, which decreases inductance. Simply, the magnetic flux density is in phase with the applied field for an increase in μ_r (neglecting hysteresis losses caused by the

Copyright © Marcel Dekker, Inc. All rights reserved.

changing field), whereas the eddy-current–generated secondary magnetic field (ideally) opposes (180° out-of-phase) the applied field. These two effects independently influence the impedance of the EC probe.

In the ideal case, the impedance of the coil would only increase inductively (reactively) and there would be no increase in real resistive losses. But the impedance plane for a solenoidal EC probe encircling a magnetic material shows both a large increase in ωL and an increase in R. Figure 5.15 depicts a family of curves for $\mu_r = 1, 5, 10$ assuming 100% coupling. What mechanism creates these real losses? Figure 5.16 shows the response on B for changes in the applied field H. The area enclosed by the curve represents hysteresis losses. As the applied field changes, energy is required to align the magnetic domains to the field. Therefore an increase in μ_r (slope of the B–H curve) shifts the impedance towards an increase in both ωL and R, as shown in Figure 5.15 with arrows labeled μ_r. The trajectory for decreasing μ_r at a given $\omega\sigma$ will end on the comma curve at the point $a\sqrt{\omega\sigma\mu_0}$.

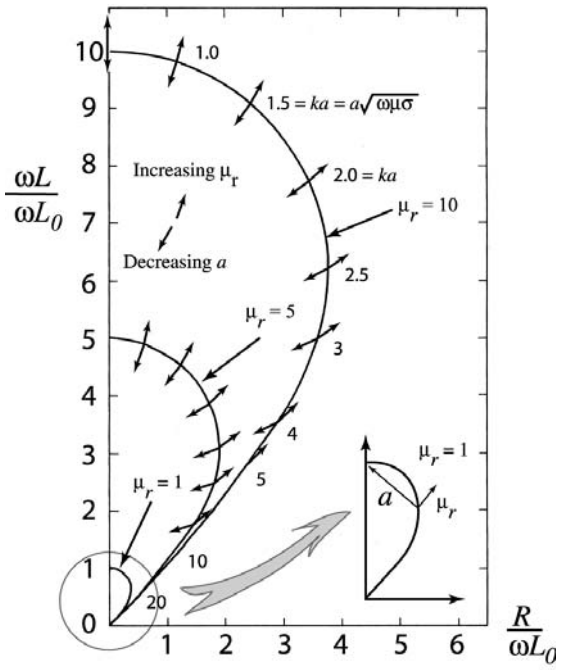


FIGURE 5.15 Impedance plane trajectories for $\mu_r = 1, 5, 10$. (Adapted with permission from HL Libby. Basic Principles and Techniques of Eddy Current Testing. Florida: Robert E. Krieger Publishing Company, 1979, p 55.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Reversible (Effective) Magnetic Permeability. Also depicted in the $B-H$ curve is the importance of the magnetic history of the material (Figure 5.16). The small hysteresis curves represent the hysteresis loop followed during an EC test; eddy currents excite an AC H-field. The slope of these loops is called *reversible permeability*, μ_{rev}. Which loop is traced depends on the DC (static) magnetization of the test material, for example, the application of a DC magnetic saturation field. The intent of a DC saturation field is to align all of the magnetic domains. Such a saturation field would reduce the effect of the skin depth, $\delta = (\sqrt{\pi f \sigma \mu_0 \mu_r})^{-1}$, thereby increasing the depth of EC penetration on ferromagnetic materials. In EC calculations μ_{rev} should be used.

Liftoff (12)

Liftoff describes the distance between the EC probe and the test object. Clearly, as liftoff increases, the inductive coupling from the probe to the test piece will decrease; therefore the test piece's influence on the impedance of the EC probe diminishes (less of the generated magnetic field reaches the sample) and probe impedance approaches ωL_0—its free space impedance.

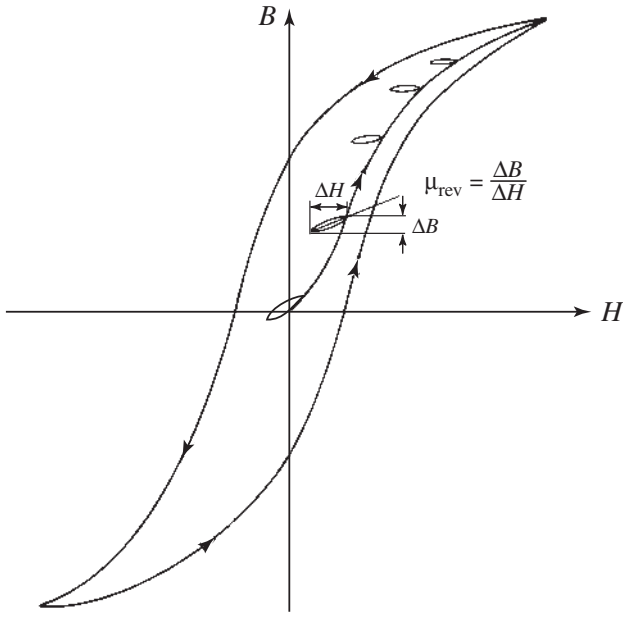


FIGURE 5.16 B–H curve. The small hysteresis loops represent the sinusoidal excitation of the EC probe for different magnetization histories.

Copyright © Marcel Dekker, Inc. All rights reserved.

Figure 5.17 shows the effect of liftoff on the impedance response of the EC probe. Here liftoff l is normalized to the mean coil radius: $l_1 = l/\bar{r}$. The logic for this normalization is that for a given absolute liftoff, the response for a large probe is smaller than that for a smaller radius probe. The dashed lines show the trajectory for decreasing liftoff for fixed values of $\bar{r}^2\omega\mu\sigma$. The solid comma curves represent a constant liftoff as $\bar{r}^2\omega\mu\sigma$ varies form 0 to ∞. The largest solid curve represents 100% coupling, $l_1 = 1.00$, and represents the ideal maximum sensitivity to the variation in material parameters such as conductivity. Note that complete cancelation of the generated magnetic field ($\omega L/\omega_0 L = 0$) can occur only in the ideal case of $l_1 = 1.00$.

Several important observations can be made from Figure 5.17:

1. Impedance changes more rapidly with liftoff when the probe is close to the surface of the part.

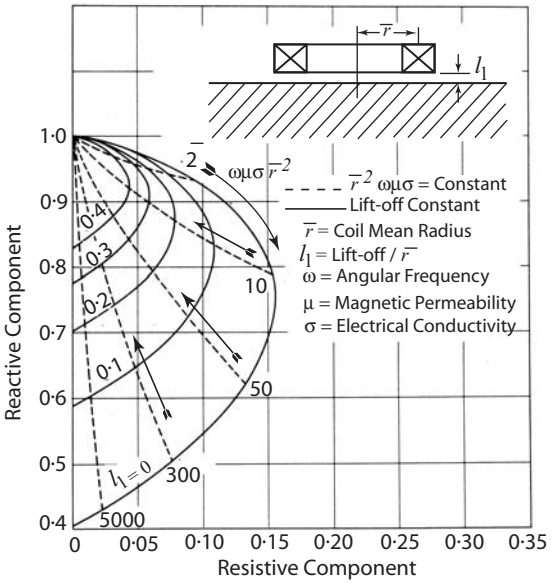


FIGURE 5.17 Effect of liftoff on a flat test piece inspected with an EC coil of mean coil radius $\bar{r}$: Multiple liftoff impedance trajectories are drawn for specific values of $\bar{r}\sqrt{\omega\sigma\mu_0}$ (dashed). Additionally, the trajectories for specific fill factors as $\bar{r}\sqrt{\omega\sigma\mu_0}$ varies are presented (solid). Arrows indicate direction of increasing liftoff. Note that liftoff is normalized to the mean coil radius to account for varying size probes. (CV Dodd. Computer Modeling for Eddy Current Testing. In: RS Sharpe, ed. Research Techniques in Nondestructive Testing. London: Academic Press, 1977, p 449.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

2. Liftoff variations with high conductivity materials cause predominantly reactive (ωL) impedance changes.
3. Small-diameter probes are more sensitive to liftoff than are larger-diameter probes.

The liftoff for EC probes that encircle the test sample is called the *fill factor*. The fill factor, η, is proportional to the amount of the encircling coil that is filled with the test specimen, i.e.,

$$\eta \propto \frac{\text{Diameter of the sample}^2}{\text{Diameter of the coil}^2} \tag{5.64}$$

Edge Effects

As an EC probe approaches the edge of a conductive test sample, the eddy currents must contour to the edge geometry, and a portion of the probe field is no longer intercepted by the sample. The impedance locus of the edge effects for a nonmagnetic material (aluminum) and a magnetic material (carbon steel) are shown in Figure 5.18. The arrow indicates the impedance change as a probe is

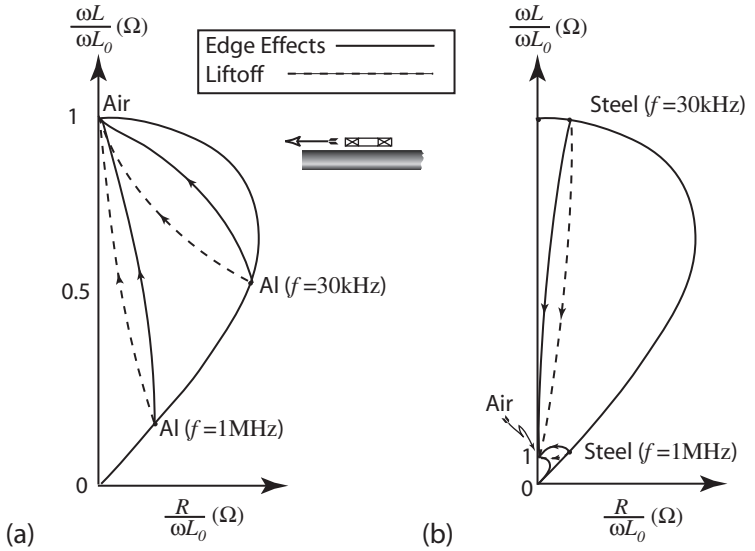


FIGURE 5.18 Edge and liftoff effects on the impedance plane for several frequencies: (a) nonmagnetic material (aluminum) and (b) magnetic material (carbon steel).

Copyright © Marcel Dekker, Inc. All rights reserved.

scanned across the edge of the material. The scan begins with the probe over the sample, and far enough from the edge that it does not influence the probe's coupling magnetic field. The scan ends with the probe in free space, i.e., $Z_{\text{normalized}} = 0 + j1$. For comparison, the figure presents the effects of liftoff and varying excitation frequency. In application, discrimination between edge and liftoff effects increases at lower frequencies for nonmagnetic materials where the angular separation between the two curves is larger and increases at higher frequencies for magnetic materials. Optimization methods for separation of these and other effects are detailed in Eddy Current Measurement Optimization in Section 5.4.1.

Material Thickness

We know that the magnetic field and the concomitant induced eddy currents decrease with increased depth of penetration into the conductive sample. This phenomenon is described as skin depth, δ. When the thickness of the material decreases to the order of several skin depths, a significant amount of the magnetic field penetrates the material. The test material no longer influences this portion of the field that penetrates through the test sample. Figure 5.19a shows the trajectory of the changes in impedance of the EC probe as the (nonmagnetic) material thickness decreases. Five metals (titanium, stainless steel, brass, aluminum, and copper) with different conductivities are shown. The spiral curves represent the impedance locus of variations in specimen thickness. The solid comma curve line represents the impedance locus of conductivity variation on a "thick" specimen, and the dashed line represents changes in liftoff. A typical thickness curve could be generated by scanning a wedge-shaped material. The spiral-shaped response, caused by changes in thickness, begins on the comma curve where any increase in thickness does not change the impedance of the probe. (Recall that at a material thickness of 5δ, the magnetic field that generates the eddy currents has dropped to less than 1% of its value at the material surface.) On the other hand, as the sample thickness decreases, the impedance trajectory leaves the comma curve and spirals towards the probe impedance of free space; at this point there is no sample as the thickness approaches zero. Note that the thickness trajectories for materials with high conductivities have larger spirals than those of the low-conductivity materials. Alternatively, for a given material (constant conductivity), as the frequency increases the thickness spiral will also increase. Figure 5.19b shows the effect as a magnetic material (steel) progressively becomes thinner.

5.2.8 Impedance Plane Analysis for Tubes and Rods (Optional)

Wire, tubing, and rods are commonly inspected using EC methods. (Consider wire as a solid rod.) EC inspection accommodates nearly any tube or rod

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

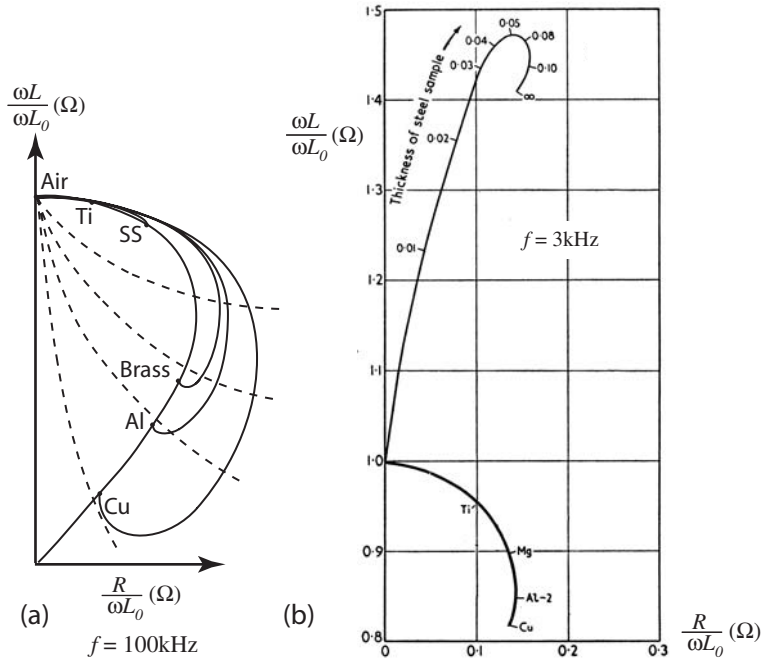


FIGURE 5.19 Material thickness effect on the impedance plane for (a) titanium, stainless steel, brass, aluminum, and copper at $f = 100$ kHz and (b) steel at $f = 3$ kHz. (R Hochschild. Electromagnetic Methods of Testing Materials. In: EG Stanford, ed. Progress in Non-destructive Testing Vol. 1. London: Heywook & Company Ltd., 1958, p 82.)

geometry (circular, rectangular, oval, etc.) and is especially common for heat exchanger tubing. We can inspect tubing and rods with reasonable sensitivity up to about 55 mm in diameter using encircling coils. Beyond 55 mm, we typically inspect with surface probes (13).

Most of the concepts presented here are simply a variation of those presented for surface probes, such as fill factor versus liftoff. We classify tubes and rods into three categories: (a) thin-walled tubing, (b) thick-walled tubing, and (c) solid rods. The EC probes are either encircling or internal (bobbin) probes; encircling probes have the drive and the pickup coils concentric and exterior to the tubes, whereas internal probes lie entirely interior to the inspected tube (see Section 5.3.1). This section emphasizes similarities to impedance plane analysis using surface probes.

Copyright © Marcel Dekker, Inc. All rights reserved.

Fill Factor

Fill factor describes the degree of coupling of the magnetic field between the EC probe and the test part, for both encircling and bobbin probes. Similar to liftoff for a surface probe, fill factor compares the diameter of the probe coils to the diameter of the tube or rod; thus fill factor is defined as

$$\eta = \frac{D_{\mathrm{OD}}^2}{\bar{D}_{\mathrm{Coil}}^2} \tag{5.65}$$

for encircling probes and

$$\eta = \frac{D_{\mathrm{Coil}}^2}{D_{\mathrm{ID}}^2} \tag{5.66}$$

for bobbin probes, assuming the probe and the tube are concentric. The greater the difference between the probe diameter and the tube or rod diameter, the less of the magnetic field reaches the test part.

Characteristic (Frequency) Parameter

For surface probes, we found it convenient to plot the normalized impedance plane, $\omega L/\omega L_0$ versus $R/\omega L_0$, using the dimensionless characteristic parameter $\bar{r}\sqrt{\omega\mu\sigma}$ or $\bar{r}^2\omega\mu\sigma$, which is normalized to the probe mean diameter. This characteristic parameter, developed by Dodd and Deeds, allows a single diagram to display the effects of changes in $\bar{r}$, ω, μ, or σ. Similarly, Libby developed a characteristic parameter for encircling or bobbin probes based on the mean radius of the test piece: $\bar{a}\sqrt{\omega\mu\sigma}$. A more common version of Libby's characteristic parameter based on length—the tube or rod radius—is the *characteristic frequency* f/f_g, developed by Förster based on frequency; f is the operating frequency of the probe, and f_g is the *limiting frequency.* As with the other forms, Förster's characteristic frequency is derived from the solution of Maxwell's equations. Because the geometry is cylindrical, the solution is a Bessel function.* f_g is defined as the frequency at which the argument of the Bessel function equals unity. For a solid rod or a thick-walled tube with an encircling coil,

$$f_g = \frac{2}{\pi\mu_r\mu_0\sigma D_{\mathrm{OD}}^2} = \frac{506{,}606}{\mu_r\sigma D_{\mathrm{OD}}^2}\,(\mathrm{Hz}) \tag{5.67}$$

*Bessel functions have series expansions just as sines and cosines and are no more mysterious.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

where all length dimensions are in meters. The characteristic frequency is often written in terms of resistivity ρ ($\mu\Omega$ cm), f_g (kHz) and D_{OD} (mm):

$$f_g = \frac{5.07\rho}{\mu_r D_{OD}^2}\,(\text{kHz}) \tag{5.68}$$

For a thick-walled tube inspected with a bobbin (internal) probe,

$$f_g = \frac{506{,}606}{\mu_r \sigma D_{ID}^2}\,(\text{Hz}) \tag{5.69}$$

and for inspection of a thin-walled tube with either a encircling or bobbin probe,

$$f_g = \frac{506{,}606}{\mu_r \sigma D_{ID} d}\,(\text{Hz}) \tag{5.70}$$

where d is the wall thickness.

Just as $\bar{r}^2\omega\mu\sigma$ defined the operating point on the impedance plane for surface probes, f/f_g defines the impedance plane operating point for encircling and bobbin probes, e.g.,

$$\frac{f}{f_g} = f\mu_r\sigma D_{ID}^2 \frac{1}{506{,}606} \tag{5.71}$$

Tube Wall Thickness

For EC inspection we classify tubes and rods into three categories: (a) thin-walled tubing, (b) thick-walled tubing, and (c) solid rods. Figure 5.20 depicts the normalized impedance curves for each of these categories. We recall that the impedance plane response for a thin wall material is semicircular. We classify a material as thin when the skindepth is much less that the material thickness. Thus a thin material sees a uniform EC density and no phase lag through the thickness. The thin tube impedance response f/f_g bounds the response for both encircling and internal probes (Figures 5.21 and 22). Such a thin tube, however, is rarely encountered in practice.

Let us look at the response for an encircling probe as the tube wall thickness changes (Figure 5.21). For simplicity, assume a fill factor of unity and a nonmagnetic tube. The probe impedance limits occur when the tube is solid and then when it is considered a thin-walled tube. The two bounding curves are shown as dashed lines. As the tube wall thickness decreases—OD, f, and σ are held constant—the probe impedance traces the dotted curve beginning on the solid tube line and ending on the thin tube semicircle. The different dotted lines represent thickness changes for different values of $f\sigma$. Note that the spiral shape is similar to the response to a surface probe on a thinning plate (Figure 5.19). The solid line in the figure indicates the impedance change for a constant thickness tube as f/f_g increases, i.e., as f or σ increases. Note that as either f or σ decreases

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

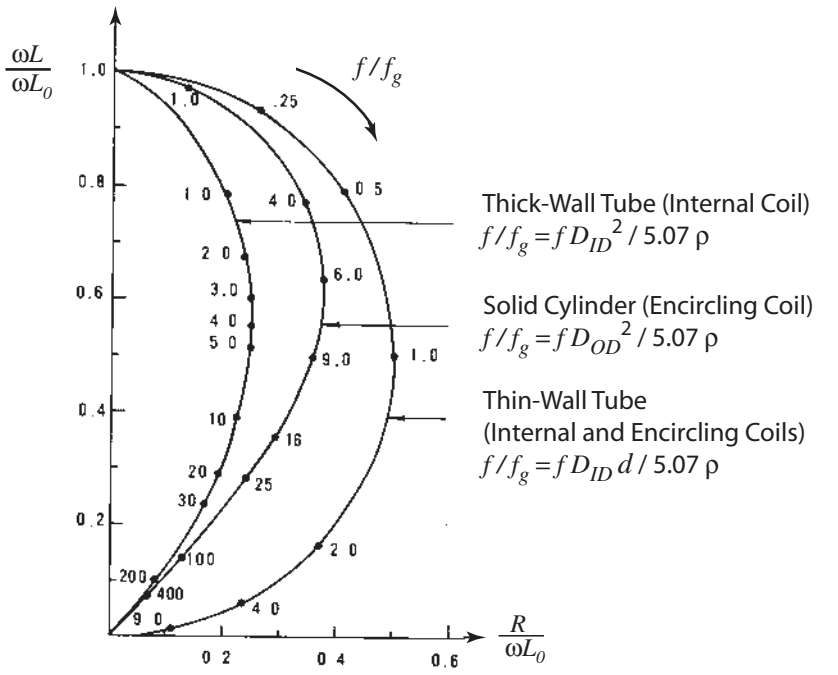


FIGURE 5.20 Normalized impedance plane for a thin-walled, thick-walled, and solid tube. Encircling or internal probe as indicated. (VS Cecco, G Van Drunen, and FL Sharp. Eddy Current Manual, Vol. 1. Ontario: Chalk River Nuclear Laboratories,1981, p 121.)

the skin depth increases, and the tube wall "looks"—to the EC probe—thinner. Thus the penetrating field at low f or σ is approximately uniform through the tube wall thickness, with very little phase shift.

Internal probe impedance response is bounded by the response from a very thick-walled tube, and again bounded by the thin tube curve. The response for any tube diameter must lie between these two impedance limits. Figure 5.22 shows these probe impedance boundaries with dashed lines. We define a tube wall as thick when $d \gg \delta$. Thus the interrogating field does not penetrate the thickness of the tube and the response signal experiences the maximum phase lag difference from that of a thin walled material. The dotted lines in the figure show the response curve as the wall thickness changes—ID, f, and σ are held constant. The solid lines trace the internal probe's response for a constant diameter tube as either f or σ change.

Copyright © Marcel Dekker, Inc. All rights reserved.

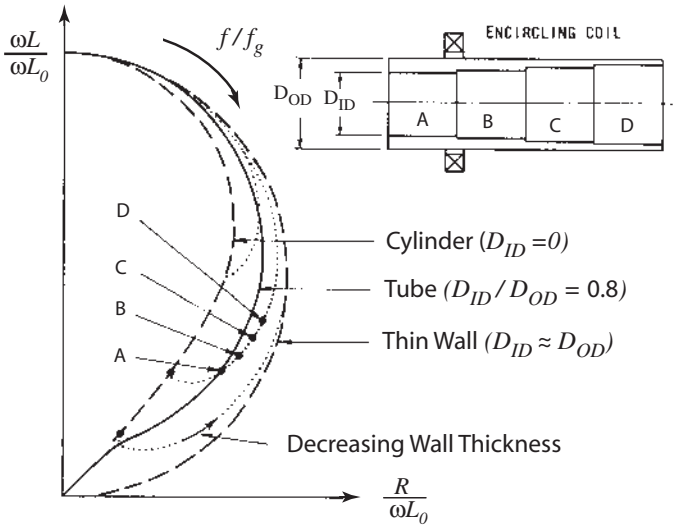


FIGURE 5.21 Normalized impedance plane for inspection of a tube with an encircling coil. (VS Cecco, G Van Drunen, and FL Sharp. Eddy Current Manual, Vol. 1. Ontario: Chalk River Nuclear Laboratories,1981, p 118.)

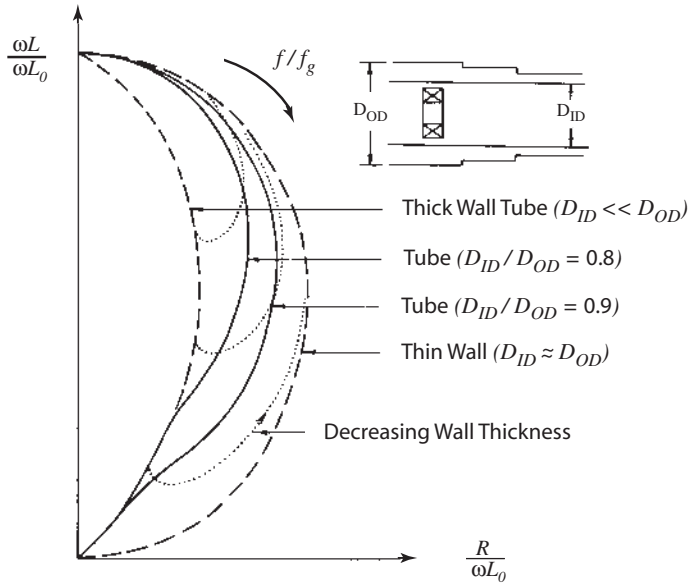


FIGURE 5.22 Normalized impedance plane for inspection of a tube with an internal (bobbin) probe. (VS Cecco, G Van Drunen, and FL Sharp. Eddy Current Manual, Vol. 1. Ontario: Chalk River Nuclear Laboratories,1981, p 119.)

Copyright © Marcel Dekker, Inc. All rights reserved.

As before, a material appearance as thick or thin to the EC probing field depends on the skin depth,

$$\delta = \frac{1}{\sqrt{\pi f \mu_r \mu_0 \sigma}}$$

compared to the physical thickness of the tube. Thus at high frequencies a given tube physical thickness will appear thick to the EC probe, and at low frequencies it will appear thin.

5.3 TRANSDUCERS AND MEASUREMENT EQUIPMENT

Fundamentally, an EC probe is nothing more than a electromagnetic transformer. So the signal from the primary (excitation) coil must be well known and highly controllable. Similarly, the secondary (receiving or pickup) coil output must be measured with a precision equal to the input. We know that the probe impedance is both resistive and inductive (and has a small amount of capacitance). Therefore, our measurement system must be capable of measuring both the magnitude and the phase of the input and output voltage or current. It must also be able to compare the response signal to the drive or to a separate reference signal.

This section explains the common types and configurations of EC probes and reference coils, outlines the fundamentals of EC measurement equipment (in particular dynamic range vs. precision, phase sensitivity, display of response, and resonant circuits), and describes some EC equipment currently in use.

5.3.1 EC Transducers (Probes)

EC probes (1) transduce time-varying current flow into a magnetic field, and (2) transduce magnetic fields into a voltage (EMF).* The most common form of EC probe is a simple coil of fine wire that forms an inductor. Other forms include Hall effect sensors and, less commonly, magnetic recording media. This section focuses only on wire-wound EC probes. But much of the following discussion is general, and would apply to any eddy current sensor.

Absolute and Differential EC Probes

There are two broad categories of EC probe: absolute and differential. The terms refer to the configuration of the pickup coils, *not* the excitation coils. (We recognize, of course, that some coils combine the functions of excitation and pickup coils into a single coil; these are referred to as single coil EC probes.) An

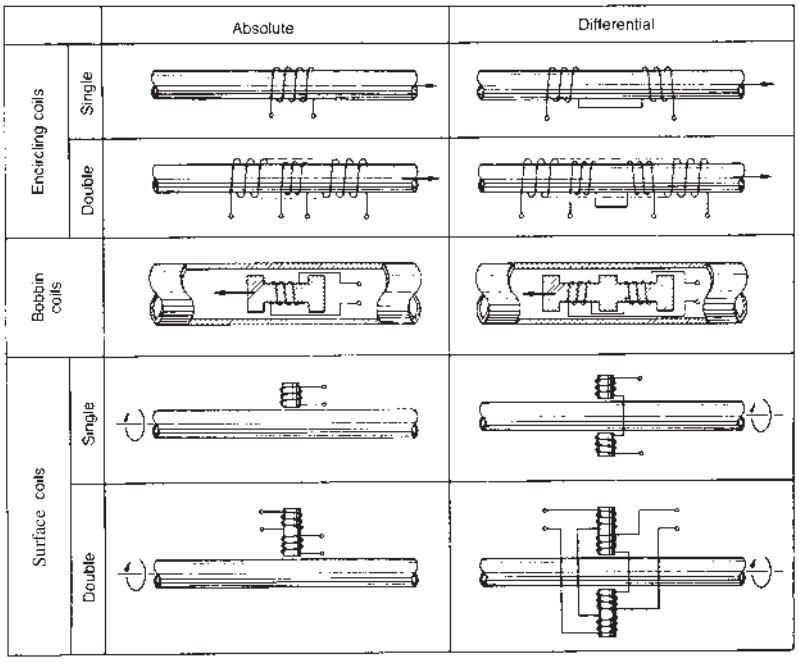
*This EMF produces current flow only if the circuit is a closed loop.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

absolute probe measures changes in the test sample or environment relative to itself; a *differential probe* compares the responses of two probes (Figure 5.23a). Note that the differential pickup coils are wound in opposite directions (hence "difference" or "differential"). The configurations labeled single have the excitation coil and the pickup coil in common (that is, they are the same coil), whereas those labeled double have separate coils. The dashed lines for the double encircling coil indicate that the winding continues but is not depicted. Figure 5.23b shows more realistic schematics of absolute and differential probes. For the double coils, the large outer coil functions as the excitation coil and is used to produce a large uniform, penetrating magnetic field. The smaller dimensions of the internal pickup coils increase the probes' sensitivity to the lateral dimension of discontinuities.

To visualize the difference between an absolute and differential probe, we can scan a tube with wall thinning due to corrosion (a slowly varying parameter) and a crack (a sudden discontinuity). Assume that we are scanning the tube



(a)

FIGURE 5.23 Absolute and differential eddy current probe windings for encircling, bobbin, and probe (surface) coil configurations: (a) general schematic to exemplify the winding configurations.

Copyright © Marcel Dekker, Inc. All rights reserved.

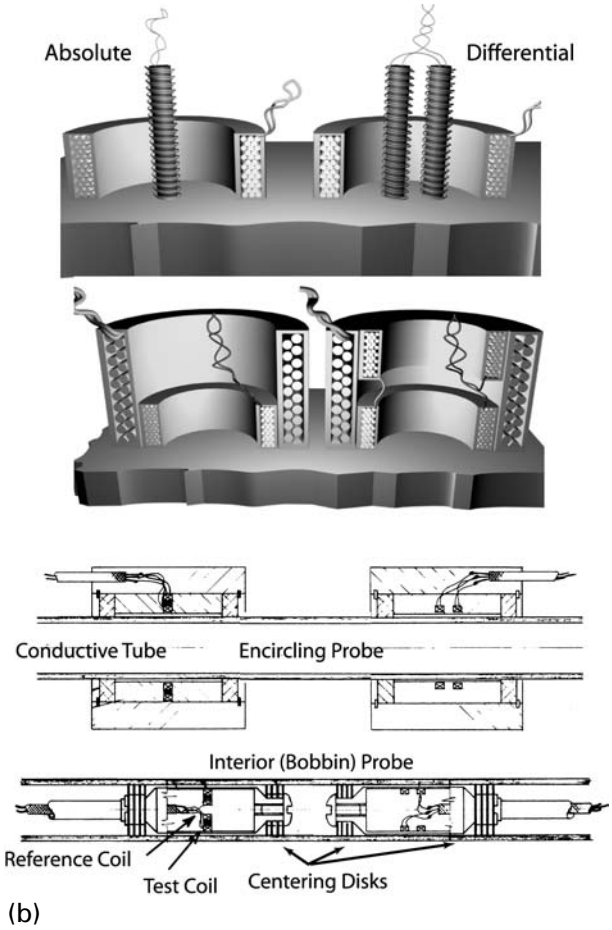


FIGURE 5.23 (*continued*) (b) More realistic probe geometries.

through the corroded section, where the wall thickness is slowly varying. Figure 5.24 graphs the probes' responses versus position along the tube; as you can see, an absolute probe responds to the change in wall thickness as well as the cracks.

On the other hand, assume a pair of identical pickup coils is configured differentially. Both coils see approximately the same wall thickness, and therefore produce responses of equal magnitude. But because the coils are wound in opposite directions, the signals have opposite polarity (i.e., they are 180° out of phase), so they cancel each other and indicate no change as the material is scanned. (The small "noise" indicates surface roughness.) The advantage of the differential configuration is that the probes are insensitive to signals common to

Copyright © Marcel Dekker, Inc. All rights reserved.

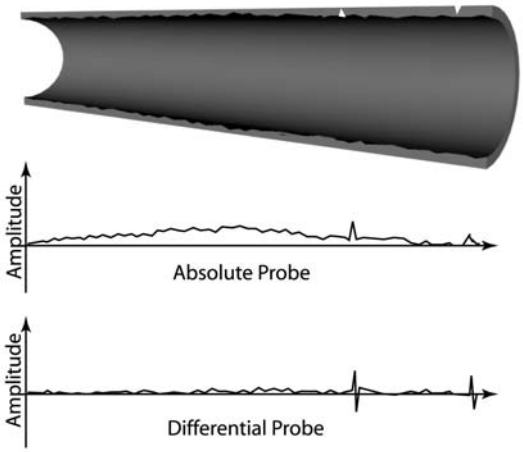


FIGURE 5.24 Eddy current probe responses for probes configured in absolute and differential modes. Scan of a tube with wall reduction due to corrosion and OD and ID cracks.

both, slowly varying geometric and material properties such as thickness, conductivity, magnetic permeability, temperature, liftoff, etc. Differential probes have high *common mode rejection*. They are therefore sensitive to sudden changes such as cracks, voids, and edges—in part because the signal is not masked by responses from slowly varying changes. Absolute probes, on the other hand, are sensitive not only to sudden changes (such as discontinuities), but also to slowly varying geometry and material properties. Figure 5.24 shows the response of absolute and differential probe to an ID and an OD crack.

Common Configurations of EC Probes

As with most NDE techniques, the specific geometry of the EC probe is dictated by the application. There are three general categories (Figure 5.23):

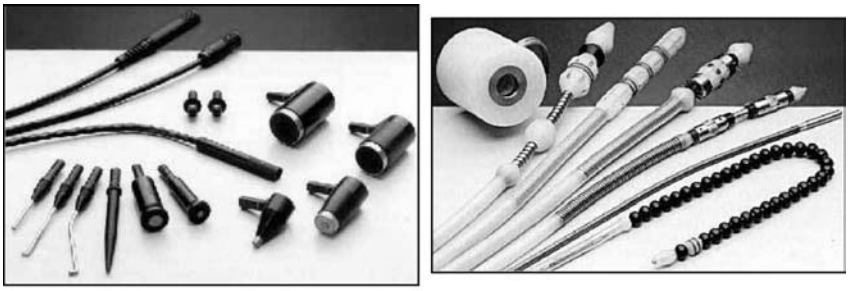
Encircling: exterior inspection of tubing and rods
Bobbin (interior): interior inspection of tubing
Probe (surface): surface inspection (single-sided or through-transmission) with the probe coil's axis normal to the inspection surface

Figure 5.25 shows a variety of surface probe coil arrangements and housing configurations designed for specific inspection applications. Figure 5.26 is a photograph of EC probes designed for inspection of flat and curved surfaces (Fig 5.26a), and for the interior and exterior of tubing (Fig 5.26b). Generally, these probes may be configured for absolute or differential operation. Additionally,

Copyright © Marcel Dekker, Inc. All rights reserved.



FIGURE 5.25 Probe geometries for specific applications.



(a) (b)

FIGURE 5.26 Commercial probes for specific applications.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

probes can be arranged in arrays or on a flexible substrate such as Kapton to follow the contour of the test piece.

EC Reference Coils

Although EC probes can be used without a reference, by far most common applications require one. Without a reference, the probe simply measures the total impedance of the coil and specimen combination, and it is difficult to determine any additional meaning from the signal. Generally, a reference signal is subtracted from the sensed signal. Any deviation from the reference signal indicates a change in variables (such as conductivity or thickness) or the presence of discontinuities. The many specific methods of obtaining a reference fall into three categories:

Internal reference coil
External reference coil
Self-reference

An *internal reference* coil is an additional coil mounted directly into the probe housing. It is connected in series to the pickup coil, and mounted perpendicular to and as far away from it as possible (Figure 5.27a). This configuration reduces the reference coil's sensitivity to the test specimen. When the probe is in air (free space), the pickup coil's signal and the reference coil's signal are equal. If they are counterwound, the signals cancel each other. When

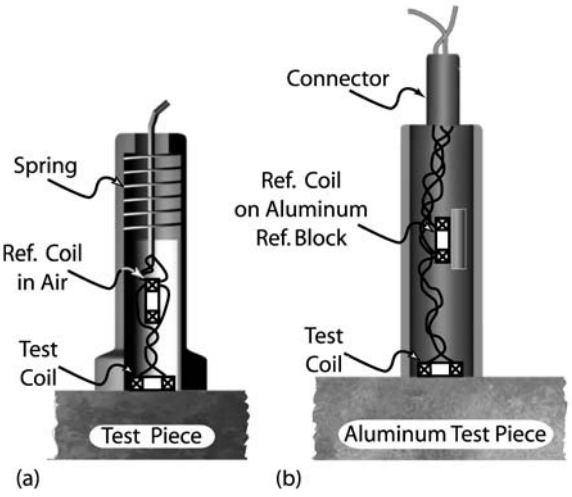


FIGURE 5.27 Eddy current probes with (a) internal reference coil and (b) internal reference coil mounted on a block of reference material.

Copyright © Marcel Dekker, Inc. All rights reserved.

the probe is placed on the test specimen, the probe output is the difference between the response from air and the response from the sample. We can make this configuration even more sensitive by allowing the reference coil to see a defect-free sample of the same material that the pickup coil sees (Figure 5.27b). Simply, a small piece of material is inserted next to the reference coil.

A more practical way to improve sensitivity is to use two matched probes containing only the pickup coil (and an excitation coil, if we use double probes). One probe acts as the sensor, and the other as an *external reference coil*. When both probes are in air, the signals cancel. During testing, however, the second (reference) probe can readily be placed on any material—such as standard conductivity blocks—resulting in greater flexibility and sensitivity of measurements.

Self-reference is a modern improvement on the external reference coil. The method uses only one probe; it is placed on a reference block, and the response signal is stored electronically. The test signal is then compared to the stored reference signal. Since finding two identical probes is difficult (even if you buy them as a matched set), this method offers the advantage of an "excitation probe" and "reference probe" that are absolutely identical.

Guiding and Shielding Magnetic Fields

Another important variation of EC probe design is the use of high-conductivity or high-magnetic-permeability materials to shield or guide the magnetic fields. In industrial environments, stray electromagnetic fields from high-frequency equipment (such as induction furnaces) can influence the probe response; the EC probe, remember, is an inductive antenna. Similarly, the probe itself may produce a stray field which interferes with neighboring sensors (although this is not a common problem). To resolve both of these issues, metals or metal oxides with high conductivity and/or high magnetic permeability are commonly used as shrouds to shield stray fields, or as guides to confine them. The most common materials for this purpose are mu-metal, ferrite, and steel.

To increase the magnetic field coupling between the probe and specimen, the coils are commonly wound on ferrite cores.* Ferrites have high magnetic permeability ($1000 < \mu_r < 10{,}000$) and low resistive losses. Adding them to an EC probe provides a low-reluctance path to and from the test specimen. Guiding the magnetic field along the ferrite increases the amount of magnetic field delivered to the surface of the test specimen from the excitation coil as well as the amount of reflected field that reaches the pickup coil. A ferrite *cup-core* increases probe-to-specimen coupling as well as electromagnetic shielding (Figure 5.28) (14).

*Although other materials can be used to increase the coupling between the probe and specimen, ferrites are by far the most common.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

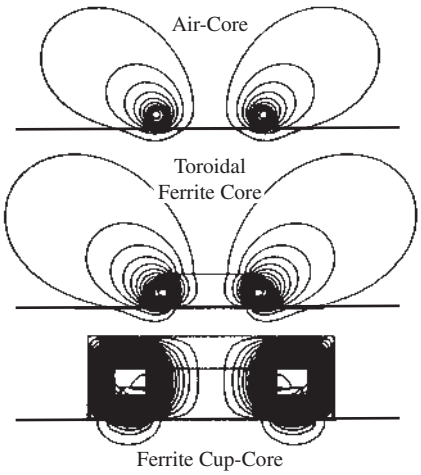


FIGURE 5.28 Eddy current probes showing the effect of an air-core, toroidal ferrite core and a ferrite cup-core or pot to guide the magnetic field. The cup-core also shields the sensor from external electromagnetic noise. (N Ida. Introduction to Electromagnetic Testing. In: RC McMaster, P McIntire, and ML Mester, eds. Nondestructive testing Handbook, Vol. 4, Electromagnetic Testing. American Society for Nondestructive Testing, 1986, p 77.)

In addition to increasing the probe-to-specimen coupling, ferrites can direct the magnetic field in specific geometric directions. Consider how we might optimize an EC probe to determine the direction of a crack in a conductive material. For best detection, the eddy currents should flow perpendicular to the discontinuity—so the magnetic field should flow uniformly parallel to the length of the crack. One solution uses a modified toroidal ferrite as the probe core (Figure 5.29) (15,16). In this example, the excitation coil doubles as the pickup coil. The gap in the toroid allows the magnetic field to interact with the specimen; otherwise, the field would remain entirely within the toroid. The pole pieces of the toroid are shaped to create a large, uniformly directed magnetic field. As this *uniform field eddy current* (UFEC) probe is rotated over a linear crack, the signal is maximum when the probe's magnetic field is parallel to the crack, and is approximately equal to zero when the field is perpendicular to the crack.

Coil Parameters

EC probe coils vary greatly in size, number of turns, aspect ratio (height to diameter), ferrite or air core, etc. All of these parameters affect the coil's inductance, and thus the probe's performance. This section discusses only a few of these variables.

Copyright © Marcel Dekker, Inc. All rights reserved.

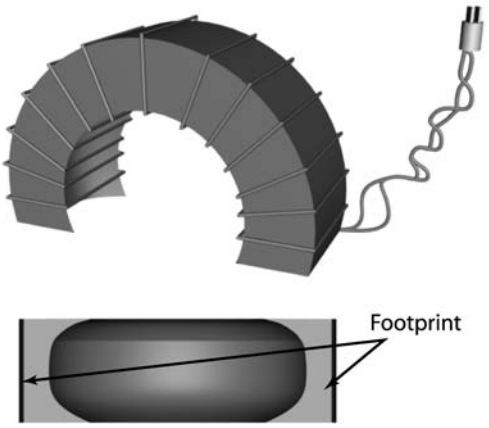


FIGURE 5.29 Uniform field eddy current probe. Coil is wound on a ferrite toroid with a shaped gap.

The larger the inductance, the larger the interrogating magnetic field. Recall that for a single-loop coil, Faraday's law states

$$V_{\text{emf}} = -\frac{d\phi}{dt} \tag{5.72}$$

the induced voltage is proportional to the change in the total magnetic flux ($\phi = BA_{\text{loop area}}$) within the loop. Magnetic flux is increased by adding multiple turns to the coil:

$$V_{\text{emf}} = -N\frac{d\phi}{dt} \tag{5.73}$$

where N is the number of turns. This equation can be directly related to the inductance and the current flowing in the coil by $\phi = L_{\text{single turn}}I$,

$$V_{\text{emf}} = -NL_{\text{single turn}}\frac{dI}{dt} = -L\frac{dI}{dt} \tag{5.74}$$

where L is the total coil inductance. In addition to the number of turns and the current, the probe diameter (area) also influences the probe inductance: the larger the probe diameter, the greater the inductance, and thus the larger the magnetic field. Why not just have a very large number of turns and a large-diameter probe? We must balance the increase in field from an increase in turns and/or probe diameter with the high spatial resolution of a small-diameter probe. Although large-diameter probes have greater sensitivity because of increased inductance, the probe's response is the average response over the entire probe area. A smaller probe averages over a smaller area, and therefore has a greater spatial resolution.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

For double-coil probes, the excitation coil is wound with a low-gauge (large-diameter) wire with just a few turns. The pickup coils are typically wound with a large number of turns (50–250) of very fine wire (34 to 48 AWG, 0.16–0.028 mm diameter). In single-coil probes, the wire diameter is typically higher, and thus the number of turns is smaller. Typically the number of turns is dictated by the space available for a probe of a given diameter.

Another restriction on coil parameters is the coil impedance itself. Any parameter that affects the coil inductance will, by definition, affect the coil impedance:

$$Z = |\vec{Z}| = \sqrt{R^2 + \omega^2 L^2}$$

This impedance must match the impedance of the associated EC instrumentation. Typical impedances are in the range of 20–200 ohm, with 50–100 ohm the most common. Note that this implies that a probe's useful frequency range is also restricted: ωL is the reactive impedance.

For additional details on the influence of such EC probe parameters as wire size, coil diameter-to-length ratio, and ferrite cores see work by Grubinskas (17), Capobianco (18), and Dodds (19).

5.3.2 Measurement Equipment

The range of EC measurement systems in industry varies greatly. Simple and inexpensive single-purpose equipment is very common. These instruments, which are designed for such applications as determining sheet or coating thickness and conductivity, are durable and easy to use. Instrumentation becomes more complex with the increase in flexibility and sensitivity of the measurement, as well as with the amount of data reduction required. Very sensitive fundamental laboratory equipment is available for any type of simple or advanced EC measurements.

As you know, an EC sensor is nothing more than an electromagnetic transformer with variable coupling. The variable coupling occurs in two forms: (a) different samples will produce different coupling, and (b) the small variations in sample coupling caused by variations in the sample or test parameters that affect coupling, such as discontinuities, conductivity, permeability, or liftoff. This implies that for a constant AC drive voltage at the primary (excitation) coil, the output at the secondary (pickup) coil can be considered as the sum of two signals: (a) a *base signal*—a large signal that represents the presence of the test sample in proximity to the sensor, and (b) an *interest signal*—small fluctuations about the base signal that represent variations in sample or test parameters, discontinuities, liftoff, etc. Our measurement system, therefore, must be capable of discerning a small variation in a large amplitude signal—as little as a 0.01% change of the base signal; this ability to measure very small deviations is fundamental to EC

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

measurement instrumentation. It is also important (though not required) to discriminate the phase difference between one signal and another—that is, to determine how much the phase of the output signal changes in the presence of some material or test parameter variation. Other common features of EC measurement instrumentation are:

Variable drive oscillator amplitude and frequency
Display in quadrature
Ability to rotate the display axis for suppression of response of changes in undesirable parameters
Storage of reference signals (amplitude and phase)
Multifrequency excitation
Generation of a variety of different excitation waveforms, e.g., continuous or pulsed

The specialized case of remote field EC systems is detailed in Section 5.4.2.

Instrumentation Solutions: Dynamic Range Versus Precision

A measurement instrument must accommodate the range of base signals produced by different types of samples or test conditions, e.g., samples with greatly different conductivity or permeabilities; the large base signal output itself varies greatly. The instrument must also be capable of detecting the small interest signal (often less than 1% of the base signal) with high precision. In other words, we need both *high dynamic range* and *high precision*. We also know, however, that a measurement system cannot have both high dynamic range and high precision, but must sacrifice one for the other.

Two common systems address this measurement problem: the electronic bridge circuit and the filter circuit. Both of these systems electronically balance the impedance at the value associated with the base signal. Because the system is balanced about the large impedance of the test setup, the measurement system needs to detect only the small perturbations of the interest signal associated with small deviations in a parameter (discontinuities, changes in thickness, conductivity, etc). Only the bridge circuit is explained here.

Figure 5.30a shows a basic electronic bridge circuit, commonly called a *Wheatstone bridge*. The bridge consists of four legs: two reference legs with fixed impedances (Z_1 and Z_2), one balance leg with variable (or adjustable) impedance Z_{variable} and the sensor leg with the impedance seen by the EC probe, $Z_{\text{EC probe}}$. Equal drive voltage, V_{ab}, is applied to both the parallel circuits of (a) the fixed impedances, Z_1 in series with Z_2 and (b) the variable impedance, Z_{variable}, in series with the sensor impedance, $Z_{\text{EC probe}}$. A precision voltmeter monitors V_{cd}. The impedance of the EC probe can be determined by adjusting the variable

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

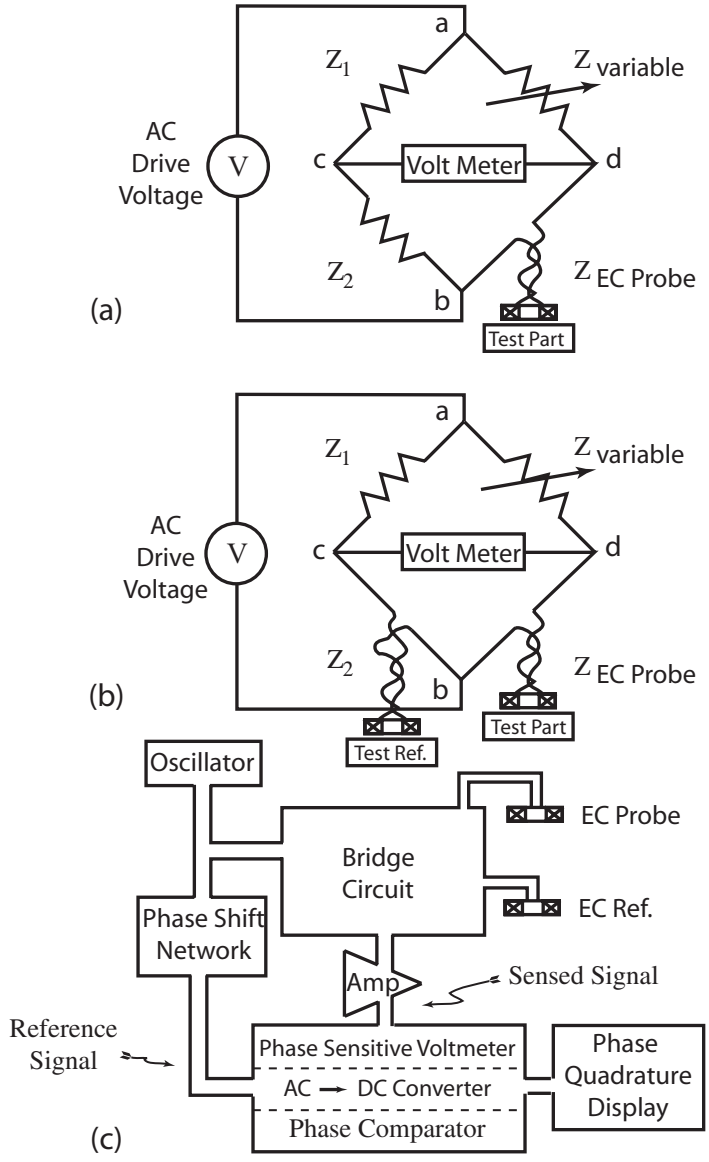


FIGURE 5.30 Wheatstone bridge used to increase the sensitivity of the EC measurement: (a) a basic balancing bridge, (b) bridge circuit with an external reference probe, and (c) bridge with phase reference and phase shifting electronics.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

impedance Z_{variable} until $V_{cd} = 0$. The unknown probe impedance (with or without the specimen) is then calculated by the voltage divider equation:

$$Z_{\text{EC probe}} = \frac{Z_2}{Z_1} Z_{\text{variable}} \tag{5.75}$$

Now that the system is balanced, any small changes in the reflected impedance in the EC probe is measured directly on the precision voltmeter (V_{cd}). Note that two separate measurements are made: a balance measurement through the adjustable Z_{variable} (base signal), and a voltage measurement of small perturbations in the test sample or setup (interest signal). This simple circuit provides the fundamental measurement for an EC system: the *magnitude* of the impedance of the EC sensor. Fig 5.30b shows how an EC system with an external reference probe can be accommodated.

Simple modifications to the Wheatstone bridge allows and EC probe to measure both phase and magnitude. A phase comparator is added to the circuit to compare the change in phase (lead or lag) between the original drive voltage and the phase of the output signal (Figure 5.30c).

Phase Sensitivity: Phase and Quadrature

The ability to measure and manipulate the phase of an EC signal is fundamental to many EC measurements, such as impedance plane evaluation. The instrumentation must (a) measure the phase difference between the sensed signal and a reference (or another sensed signal), and (b) be able to shift (or rotate) this phase difference. Phase measurement was discussed in the previous section and depicted in Figure 5.30c as a phase-sensitive voltmeter. To rotate the phase, we usually add a phase-shifting network in the reference line; this adds some variable phase lead or lag to the reference signal before its comparison to the sensed signal (Figure 5.30c).

What do phase measurement and phase rotation do for EC measurement? Knowing the phase (and the magnitude of the sensed signal) allows us to use the impedance diagram (specific use will be presented in Section 5.4). Phase rotation, on the other hand, is not so much a requirement as a convenience: it improves visual interpretation of the signal. In any impedance plane measure, the signal response to a parameter (cracks, conductivity, etc.) must be distinguished from the other parameters sensed by the probe. For example, liftoff is almost always an unwanted parameter. Figure 5.31 illustrates the utility of phase rotation in the response for a surface probe as it is scanned over a discontinuity. The dotted box indicates the oscilloscope display; the dotted lines indicate the probe response. The probe begins at point 1 on the sample with a corresponding coil voltage (impedance), $\vec{V}_1(\vec{Z}_1^n)$, indicated by the vector from the origin to point 1 on the conductivity curve and assumed with an initial liftoff of zero. (The superscript n denotes normalized impedance.) As the scan progresses, the probe senses the

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

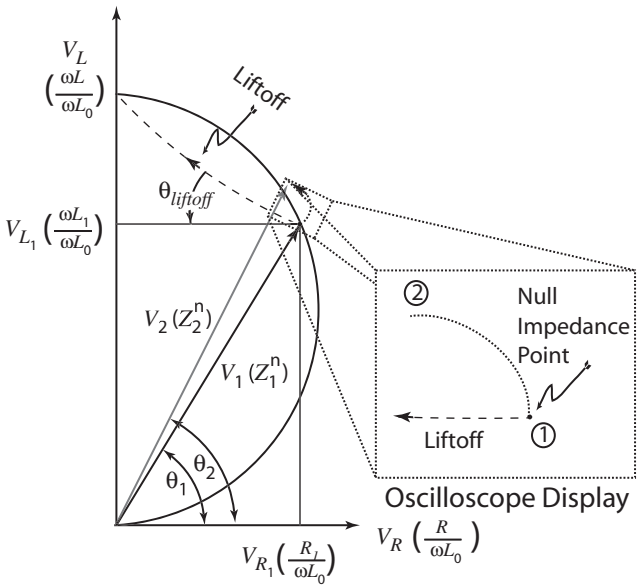


FIGURE 5.31 Sensor response voltage represented in quadrature. The dashed box represents the display on an oscilloscope of the scan in the inset.

discontinuity, as well as some liftoff from a little wobble (caused by the operator) in the probe. Now the response signal contains information for the parameter of interest (the discontinuity) and the unwanted parameter (the liftoff). Typically the operator rotates the signals until the liftoff is along the *x*-axis of the oscilloscope—a rotation of θ_{liftoff}. This rotation is illustrated by the oscilloscope display, rotated until the scope's *x*-axis is parallel to the liftoff trajectory. Now any signal that deviates from the *x*-axis is assumed to be associated with the discontinuity, and can be amplified independently of the liftoff. In practice, the liftoff signal is measured on a defect-free test sample and the signal rotated as part of the EC system setup (prior to actual measurements). Often the signal of interest must be extracted from additional signals; other techniques are required to reduce the influence of these unwanted parameter responses. The rotation method is simply a quick way of separating the liftoff, which is almost always present in any scan and usually not of interest. Note that this method can separate only two parameters.

What is quadrature? For a display in quadrature, an EC instrument measures an amplitude and the phase difference (expressed as an angle) between the interest and reference signals. These values, a magnitude and an angle, are expressed as DC values. (Note the AC to DC converter in Figure 5.30c.) On the impedance plot in Figure 5.31, the sensor response at point 1 had a magnitude V_1

Copyright © Marcel Dekker, Inc. All rights reserved.

with a phase of ϕ_1. The quadrature simply represents this vector, $\vec{V}_1$, in Cartesian coordinates with the (real) resistive component along the x-axis and the (imaginary) inductive reactive component along the y-axis. Although the x and y directions on the oscilloscope display the resistive and reactive voltages from the sensor response, the axis also represents the normalized resistive impedance $R/\omega L_0$ and the inductive reactance impedance, $\omega L/\omega L_0$; of course, the scales would be different. Note that an EC impedance plane oscilloscope does not plot time; therefore, no discontinuity position information is obtained from an impedance plot.

The details of measuring and manipulating the magnitude and phase for a specific system may be far more complex than indicated here, but the concept generally holds.

EC Response Display

Information displays range from a single digital readout to multiwindowed computer presentations. Simple displays are common for specialized equipment, such as paint thickness meters or some crack detectors. More advanced equipment can display magnitude and/or phase vs. time (position) in conjunction with the impedance plane plot. The time plots are displayed on strip charts and the cursor is used to coordinate the position on the strip chart with the point on the impedance plane. Assuming a constant scan rate, the operator can determine the position of a discontinuity from the strip charts, and the discontinuity details from the impedance plane. The computer monitor display in Figure 5.32 shows a full-featured display, including scan statistic and part identification, strip charts that can be used for multiple probes or for amplitude and phase, impedance plane, and a amplified version of the strip charts. Other displays yield actual impedance values as a function of frequency.

Resonant Circuits

Resonant probes are commonly used for EC crack detector systems and other applications. The instrumentation consists of an R–L–C circuit designed to resonate about a specific frequency. Figure 5.33 shows a simplified circuit for a resonance system where the probe is represented by the inductor. This resonant probe circuit represents the sensor leg in a bridge-balancing circuit. Both the capacitor and the resistor may be variable. Typically, the capacitor is chosen such that the inductive reactance ωL is approximately equal to the capacitive reactance $1/\omega C$. At resonance, the total impedance is

$$Z = \frac{R(\omega C)^2}{R^2 + (\omega L - \omega C)^2}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

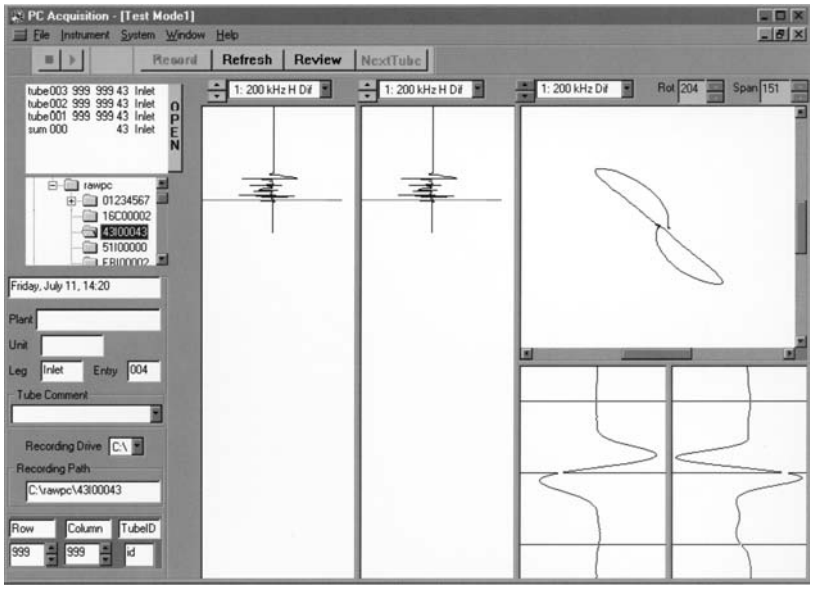


FIGURE 5.32 Computer based EC instrument display. (Courtesy Zetec.)

This impedance is minimized (response amplitude maximized) when the frequency at which the power of the inductance and capacitance are equal:

$$f = \frac{1}{2\pi\sqrt{LC}}\left(\frac{Q^2}{1+Q^2}\right)$$

where $Q = \omega L/R$, the quality factor, relates the total energy stored to the total energy dissipated per period.

In applications such as crack detection, the frequency of operation is chosen as a compromise between maximizing the slope of the resonance curve and minimizing the effect of liftoff. The maximum slope approximates the

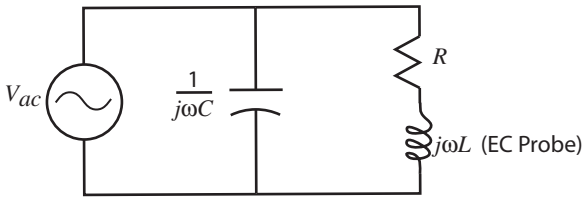


FIGURE 5.33 Simplified eddy current resonant circuit.

Copyright © Marcel Dekker, Inc. All rights reserved.

maximum sensitivity to a crack, i.e., the maximum amplitude change for a given change in probe impedance. More sophisticated systems monitor the shift of the rapid change in phase associated with resonance. (The phase of an ideal resonant system rapidly shifts from 180° to −180° as the frequency passes through resonance.)

5.3.3 Actual EC Equipment

EC measurement equipment is available in a wide variety of formats, from single- or multipurpose handheld units, to custom-designed high-speed scanning systems designed to inspect specific parts (often integrated into manufacturing or assembly lines), to highly versatile impedance analyzers (typically used in research and development). The cost of these systems range from a few hundred dollars (U.S.) for simple handheld units to tens of thousands of dollars for manufacturing-inspection systems.

Figure 5.34 shows two battery-operated handheld units: a single-purpose EC unit designed to measure conductivity thickness, and a versatile multipurpose unit designed to measure discontinuities (cracks, pitting, corrosion, etc.), conductivity, and nonconductive coating thickness. The second device has a

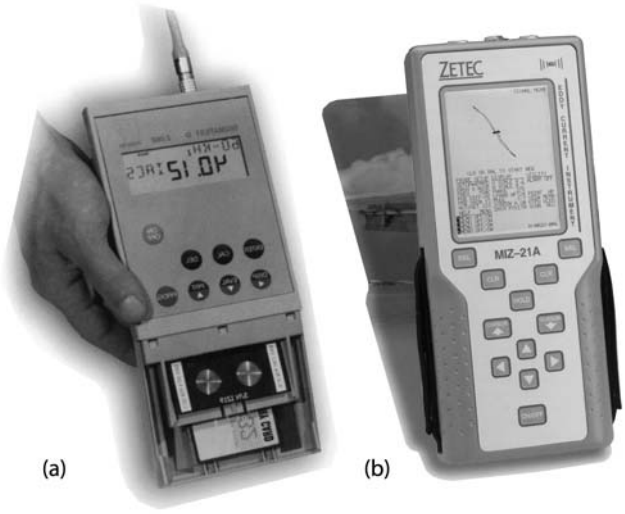


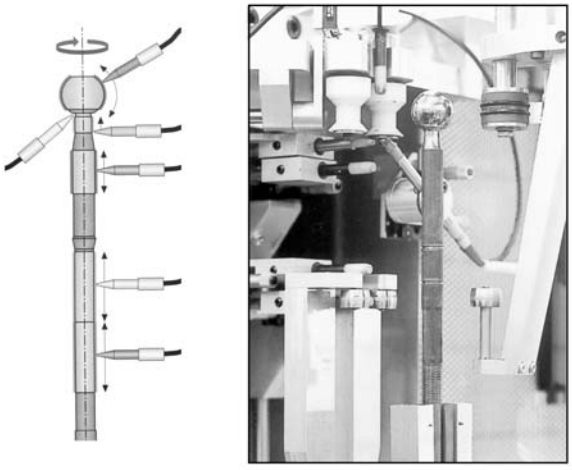
FIGURE 5.34 (a) Handheld EC unit designed to measure conductivity. (Courtesy Förster Instruments.) (b) A versatile handheld EC unit for measuring conductivity, non-conductive coating thickness, and crack detection with dual frequency signal discrimination capability. (Courtesy Zetec.)

Copyright © Marcel Dekker, Inc. All rights reserved.

variable frequency range from 50 Hz–8 MHz. The discontinuity response can be displayed on the impedance plane, as well as on a strip chart. The conductivity can be displayed in either % IACS (with the conductivity of copper equal to 100%) or MegaSiemens/m (a Siemen is equal to inverse Ohms). The unit's dual-frequency mode allows the operator to reduce the effects of unwanted signals, such as those from liftoff or temperature.

EC systems can also offer high-speed inspection of mass-produced parts for process and quality control. Typically, these systems tend to be very complex with multiple EC probes. The complexity is usually associated with part or probe manipulation, not EC measurement. Figure 5.35 shows a fully automated EC crack-detection system for automotive steering push rods. Six simultaneous-scanning EC probes inspect 100% of all critical sections of the push rod for cracks. This robotic quality-control EC system has a throughput of approximately 600 parts per hour, and automatically separates the parts according to an accept-or-reject criterion. The system is designed to handle a variety of push rod geometries.

The most versatile EC instrument is the network/impedance analyzer (Figure 5.36). This basic laboratory instrument offers extremely high sensitivity, frequency sweep (sub-Hertz–100 MHz), multifrequency, direct measurement of real and imaginary impedance as well as phase and amplitude, versatility in programming, compensation, etc. The three major drawbacks of the instrument



(a) (b)

FIGURE 5.35 Automated eddy current inspection system designed to test automotive steering push rods. (Courtesy Förster Instruments.)

Copyright © Marcel Dekker, Inc. All rights reserved.



FIGURE 5.36 Eddy current measurements on graphite-epoxy tubing using a network/impedance analyzer. (Courtesy Hewlett-Packard.)

are that (a) it is not industrially hardened and can be damaged easily, (b) significant operator experience is required, and (c) cost is high.

5.4 INSPECTION PRINCIPLES (20, 21)

5.4.1 Impedance Plane Measurement

EC impedance plane measurement offers the most flexibility in discerning variations in parameters such as conductivity, liftoff, permeability, discontinuities, and coating thickness. The impedance plane is practical because we can plot the EC probe impedance vector, $\vec{Z}_{\text{EC probe}}$, which has a magnitude and phase, in terms of its resistance and reactance. As we have seen, EC instrumentation plots small portions of this impedance plane about some set point with the x- and y-directed voltages corresponding to the resistance and reactance of the probe. In Figure 5.31, assume that a probe sees a material with thickness d_1 and has a response $\vec{V}_1$ with an associated normalized impedance $\vec{Z}_1^n$, point 1. (The superscript indicates normalized impedance.) On the impedance plane, this point is located by its magnitude, Z_1^n, and phase angle, ϕ_1, or by the normalized real resistive and imaginary reactive components, where $\vec{Z}_1^n = (R_1/\omega L_0) + j(\omega L_1/\omega L_0)$, as we have done before.

Copyright © Marcel Dekker, Inc. All rights reserved.

As the probe scans the sample, assume a sudden thickness change to d_2 that reflects a new probe impedance $\vec{Z}_2$ and voltage $\vec{V}_2$. On the impedance plane, the impedance moves along the thickness response trajectory from point 1 to point 2. EC measurement equipment balances (or *nulls*) the impedance about the large base signal and displays the changes about this (user-defined) *null point*; in the current example, we might choose the *null impedance* or balance point at $\vec{Z}_1^n$. As we scan across the material with thickness d_1, the measurement display indicates zero (see point 1 on the oscilloscope display in Figure 5.31). As the probe encounters the change in thickness, the display reflects this change by moving to point 2; it then returns to zero (point 1) as the probe passes out of the region with thickness d_2 and returns to thickness d_1.

We must distinguish between (a) the response of a probe as it sees a change in some parameter during a measurement and (b) the impedance plane map of the locus of probe responses as a parameter is changed through its entire range of variability (e.g., the comma curve mapping the loci for $0 > \sigma > \infty$). During a measurement, a probe response begins at some bias point, and must trace a continuous curve to the impedance representing the final (maximum) impedance for a given change in some parameter. We will call this measurement curve the *scan response trajectory* (or *response trajectory*), and we call the loci of the maximum points of the response trajectories for the entire range of the parameter the *impedance loci*. (This nomenclature is not standardized.)

The following sections show the scan response trajectories for variations in a number of different parameters. The trajectories are shown on the full impedance plane, as well as the amplified and rotated response from an EC instrument during a measurement. The plots on the full impedance planes, while not to scale, help us understand the relationships among the various parameters. Additionally, we will present a method to optimize the balance point such that the response maximizes the signal from the parameter of interest and minimizes the response from other parameters.

Cracks in Magnetic and Nonmagnetic Materials

Although there are many variations of a crack, each with associated nuances on the impedance plane, this section discusses only two categories: surface-breaking and subsurface cracks.

Surface-Breaking Cracks. Figure 5.37 show the response trajectories as the probe is scanned across surface-breaking cracks in both nonmagnetic (aluminum, Figure 5.37a) and magnetic (steel, Figure 5.37b) materials. These cracks are simulated with depths ranging from 5 mm to a depth much greater than the skindepth; at this depth the EC probe field does not see the bottom of the crack. We see that the scan response trajectories for all of the surface-breaking cracks lie between the liftoff and conductivity loci, so we could consider the

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

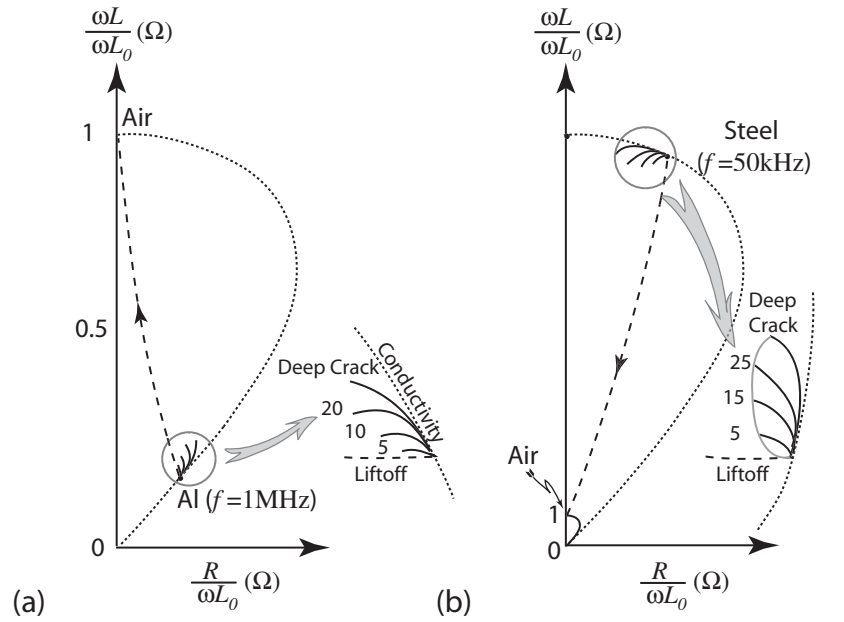


FIGURE 5.37 Flaw response trajectories for surface breaking cracks of varying depths: (a) aluminum (nonmagnetic) and (b) steel (magnetic). Impedance locus for surface breaking cracks on steel is represented with a light line connecting the tips of the flaw response trajectories.

presence of the crack to change the impedance as a combination of increased liftoff and decreased conductivity. As expected, the magnitude of change in impedance between "no flaw" and "flaw" increases with increasing crack depth. The crack response trajectories are also presented with a phase rotation, placing the liftoff along the x-axis on the display. Thus a vertical (y) axis amplification would magnify the crack signals without increasing the liftoff signal.

The impedance locus of surface-breaking-crack response for increasing depth is simply the curve connecting the response peaks for increasing crack depth, shown with a light line in Figure 5.37b. The trajectory begins on the σ curve tangential to the liftoff and curves to meet the response of a deep crack.

Subsurface Cracks (22, 23). This discussion presents the impedance loci of subsurface cracks as a function of two separate parameters: the height of the crack and the depth of the crack below the surface. (A surface-breaking crack has zero depth.) Figure 5.38a depicts the probe response to varying both depth below the surface d and the crack height s on aluminum with $f = 1$ KHz. The solid

Copyright © Marcel Dekker, Inc. All rights reserved.

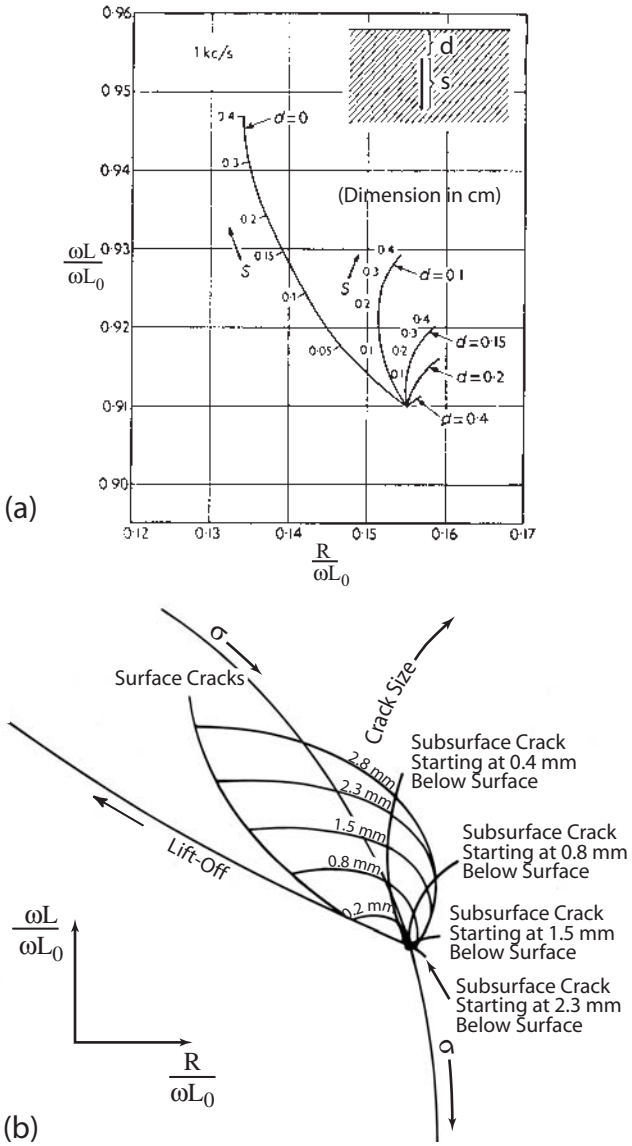


FIGURE 5.38 (a) Flaw response trajectories for four crack depths, d, below the surface as the height, s, of the crack is varied on aluminum with $f = 1$ kHz (dimensions in cm) and (b) the impedance loci for fixed crack heights as the depth of the crack is increased from $d = 0$ (surface breaking) to $d > 2\delta$ on graphite-epoxy composite (dimensions in mm).

Copyright © Marcel Dekker, Inc. All rights reserved.

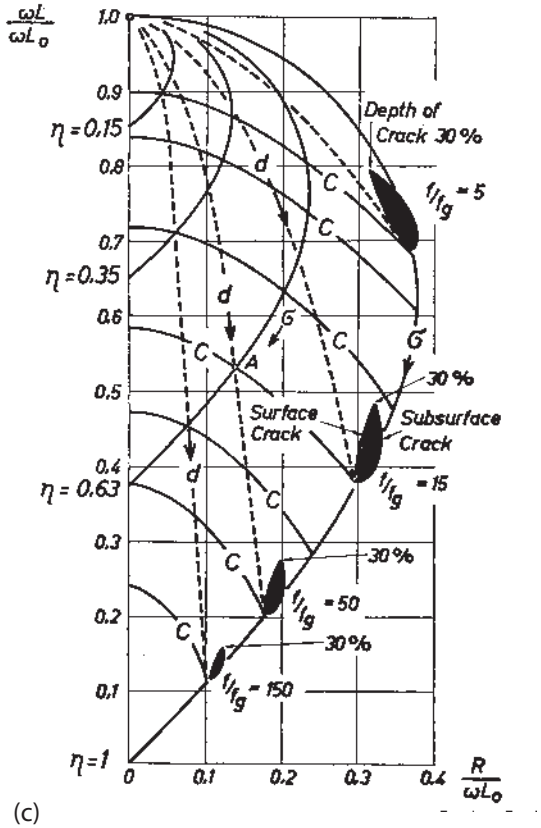


FIGURE 5.38 (*continued*) (c) Regions of crack response for different excitation frequencies or conductivities. (CV Dodds. The Use of Computer-modeling for Eddy-current Testing. In: RS Sharpe, ed. Research Techniques in Nondestructive Testing. London, Academic Press, 1977, p 469.)

curves represent constant depth below the surface as the crack height is varied. The comma (conductivity) curve has been removed for clarity. On the other hand, if we plot the impedance locus for a family of fixed size cracks as its depth increases, the locus curve begins on the maximum point on the response trajectory of a surface-breaking crack and ends on the conductivity curve where the crack is no longer detectable (Figure 5.38b). Of course, cracks don't move, and this impedance locus would never be a real probe response curve. Figure 5.38c shows the response regions for subsurface cracks (shaded areas) for various frequencies of operation (or material conductivity) on the impedance plane. The curves for magnetic material have been omitted for brevity.

Copyright © Marcel Dekker, Inc. All rights reserved.

Some Crack Realities. The shape of the response curve as a probe scans over a crack depends on many features of the crack, e.g., depth, length, width, whether the crack is planar, whether the crack walls touch. Here we examine the influence of the shape of the crack, assuming the length is much greater than the probe footprint. Provided that the crack walls do not make electrical contact, the maximum impedance magnitude is determined only by depth as measured normal to the surface, and does not depend on the actual linear length of the crack as it progresses into the depth (24). Figure 5.39 illustrates the probe response for a scan of three planar cracks with variations on the crack plane only; all of the cracks penetrate the same distance below the surface. When the crack is normal to the surface, the scan response trajectories for the probe's approach to and retreat from the crack trace identical paths. As the crack tilts from the surface normal, the approach and retreat are simply different, and hence their trajectories are different. On approach the probe "sees" the crack grow from a deep subsurface crack toward a surface-breaking crack. On the retreat, the probe still "sees" the

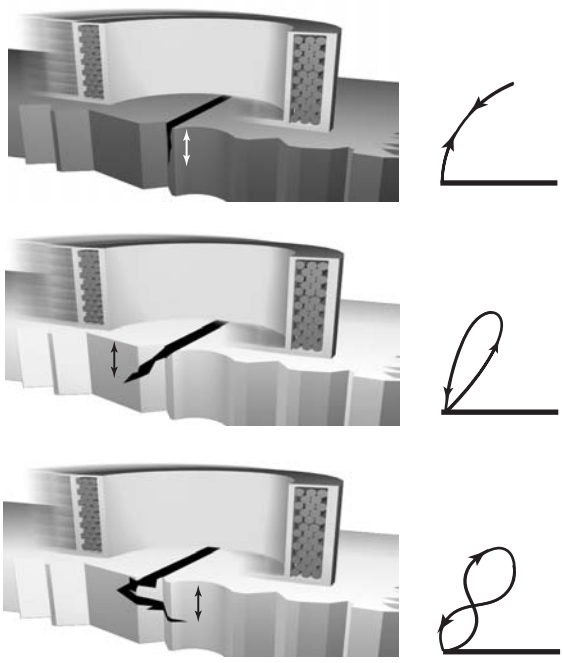


FIGURE 5.39 Impedance response trajectories as an EC probe scans flaws of equal depth but different linear lengths and geometries: (a) flaw perpendicular to surface, (b) flaw tilted, and (c) kinked flaw.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

crack growing towards a surface-breaking crack and then retreats from the crack altogether. Because the probe must return to the conductivity curve, the total impedance trajectory forms a loop (Figure 5.39b). Figure 5.39c illustrates the probe's response to a scan over a crack with a kink.

Consider: How does the probe's trajectory differ (on approach and on retreat) for a surface-breaking crack verses a subsurface crack?

The preceding discussion is an idealization of most real cracks; the crack length along the surface was assumed infinite (or at least much greater than the probe diameter) and the crack walls were assumed not to touch. (The effect of the crack length to probe diameter is presented in a subsequent section.) To varying degrees, the walls of most real cracks do touch, implying that the crack does not represent a complete change from a conductor to a nonconductor (as an ideal crack does) but a lesser change somewhere in between. For example, consider the fatigue load on the axle of a railroad locomotive. The crack will often be so tightly closed during inspection that its walls not only touch in some places, but also are in intimate electrical contact—a *kissing bond*. Thus the probe response underestimates the extent of the crack.

Effect of Probe Diameter on Sensitivity to Cracks.* The primary factors in a probe's ability to detect a given discontinuity are the liftoff, skin depth, and coil parameters. As liftoff increases, less of the magnetic field reaches the surface of the test part, thus reducing sensitivity. Skin depth describes the decay rate of the induced eddy currents relative to the eddy currents at the surface. Both of these factors effectively address how much of the magnetic field from the drive coil reaches the discontinuity. But neither addresses how much magnetic field is produced at the drive coil—the larger the initial magnetic field, the more sensitive the measurement. The excitation voltage (amplitude and frequency) and coil parameters control the magnitude of the magnetic field in the excitation coil; liftoff and skin depth control how much of it reaches the discontinuity.

The inductance L describes a coil's ability to produce (or detect) a magnetic field. The coil inductance is most influenced by the number of turns N and the mean coil diameter $\bar{D}$:

$$L \propto N^2 \bar{D}^2 \tag{5.76}$$

By increasing either N or $\bar{D}$ we increase the initial magnetic field at the drive coil, and thus the probe's overall sensitivity. Because both parameters have similar

*Although the discussion on probe diameter and sensitivity focuses on crack detection, the concept is general and applies to any defect.

Copyright © Marcel Dekker, Inc. All rights reserved.

influence on the probe inductance, this section discusses only the effect of the coil diameter on probe sensitivity.

Figure 5.40 graphs the effect of a probe's diameter on its sensitivity to a surface-breaking and a subsurface crack. The diameter's influence on the probe sensitivity for a surface crack is plotted against the liftoff for four probe diameters (Figure 5.40a). (Note that the probe response is normalized to its response at zero liftoff and plotted on a logarithmic scale.) For a surface crack, we assume that the

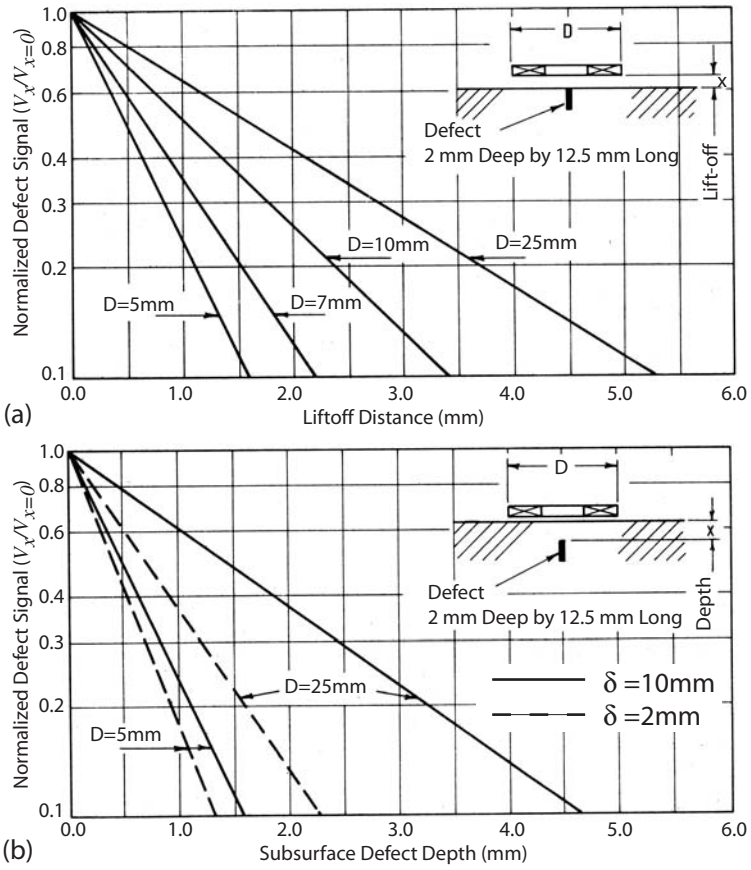


FIGURE 5.40 Probe sensitivity for different diameter probes to (a) a surface breaking flaw as a function of liftoff and (b) a subsurface flaw verses flaw depth. (VS Cecco, G Van Drunen, and FL Sharp. Eddy Current Manual, Vol. 1. Ontario: Chalk River Nuclear Laboratories,1981, p 64.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

skin depth has little influence on the response signal. For the first 1 mm increase in liftoff, we observe a decrease in signal of about 75% for the 5 mm-diameter probe, and only 28% for the 25 mm-diameter probe.

When sensing a subsurface defect, we need to account for effects of both liftoff and skin depth. Figure 5.40b graphs the response for two probe diameters as a function of depth. To separate the effects of liftoff from skin depth, two different skin depths are plotted, $\delta = 10$ mm (solid line) and $\delta = 2$ mm (dashed line). (The skin depth can be controlled by simply changing the operating frequency.) When the skin depth is large, the probe's response to an increase in depth of the crack below the surface is similar to that of increased liftoff from a surface crack (compare the solid lines in Figures 5.40a and b). In contrast, when the skin depth is small, $\delta = 2$ mm, the probe is influenced both by its distance from the top of the crack (liftoff) and by the absorption of eddy currents due to skin depth effect (compare the dashed to the solid lines in Figure 5.40b). In all cases, the larger the probe diameter, the higher the sensitivity to the presence of the crack.

The probe's diameter affects its sensitivity as much as, if not more than, the combined effects of liftoff and skin depth. That being the case, why not simply make large-diameter probes? The reason is that the increase in sensitivity achieved from an increase in diameter must be weighed against the high spatial resolution offered by the smaller-diameter probe. Recall that the total signal is integrated over the entire probe area.

Effect of Crack Length on Probe's Sensitivity. If the probe's response to a discontinuity is the average signal over the entire probe face (footprint), then it stands to reason that the size of the discontinuity compared to the probe footprint will influence the response signal. Figure 5.41 plots the probe response for two probe diameters ($\bar{D} = 1.3$ and 7 mm) and two skin depths ($\delta = 0.36$ and 1.16 mm). The normalized probe response is plotted versus crack length, with the crack width and depth constant for surface-breaking cracks. As the plot clearly shows, the larger the crack length, the larger the response signal. This trend holds until the crack length appears infinite to the EC probe, i.e., until the crack length extends beyond the interrogating magnetic field.

While it seems reasonable that a smaller probe would be more sensitive to a crack of smaller length, why does the sensitivity increase with increased frequency (assuming a constant probe diameter and crack size)? The lateral extent of the magnetic field is not limited to the physical footprint of the EC probe. Recall that the magnetic field must loop back upon itself. An effective probe diameter, $\bar{D}_{\text{eff}}$, can be approximated by

$$\bar{D}_{\text{eff}} = \bar{D} + 4\delta \tag{5.77}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

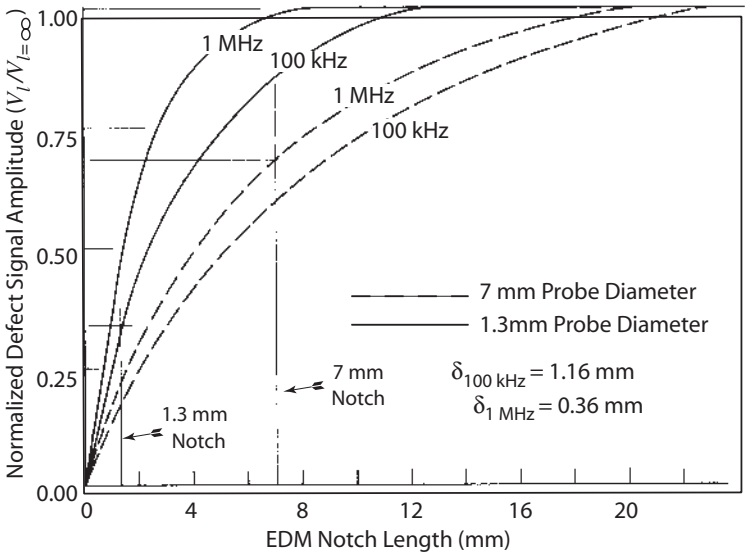


FIGURE 5.41 Probe sensitivity verses flaw length for two probe diameters and two operating frequencies. (VS Cecco, G Van Drunen, and FL Sharp. Eddy Current Manual, Vol. 1. Ontario: Chalk River Nuclear Laboratories,1981, p 67.)

Thus, the higher the frequency, the more spatially concentrated the field. For very low frequencies, Cecco et al. recommend the use of ferrite cup-core shielding probes to spatially concentrate the field (25).

As a rule, the probe diameter should be equal to or less than the length of the discontinuity.

Coatings (four types)

Determining coating thickness is an important use of EC sensing. Four distinct variations can be determined:

1. Conductive coating on a nonconductive substrate
2. Nonconductive coating on a conductive substrate
3. Low-conductivity coating on a high-conductivity substrate
4. High-conductivity coating on a low-conductivity substrate

Case 1 is equivalent to the impedance variations for a thin specimen (Figure 5.19); case 2 is a liftoff problem (Figure 5.17). This section discusses the new impedance plane variations of cases 3 and 4.

Plating a low-conductivity material onto a high-conductivity substrate will decrease the effective conductivity seen by the EC sensor. Figure 5.42a shows, for aluminum plated onto a copper substrate, the locus of the change in impedance as the low-conductivity coating increases in thickness. At zero coating thickness, the

Copyright © Marcel Dekker, Inc. All rights reserved.

impedance is that of copper; as the coating thickness reaches several skin depths, the impedance approaches the conductivity of aluminum.

Figure 5.42a also shows the direction of impedance change for copper plated onto an aluminum substrate as the plating thickness increases. Similarly (Figure 5.42b), the impedance locus for a magnetic material (nickel) plated onto a high-conductivity nonmagnetic substrate (copper) begins at the impedance of copper for zero coating thickness and approaches the impedance of nickel as the coating thickness increases, along with the apparent magnetic properties as seen by the probe.

Coatings involving a magnetic and a nonmagnetic material produce relatively large impedance changes compared to materials with like magnetic properties.

Spacing Between Two Conductors

A gap, or space, between two conductors can be an intentional design feature, e.g., an electrolytic capacitor where a dielectric film separates the two conductive

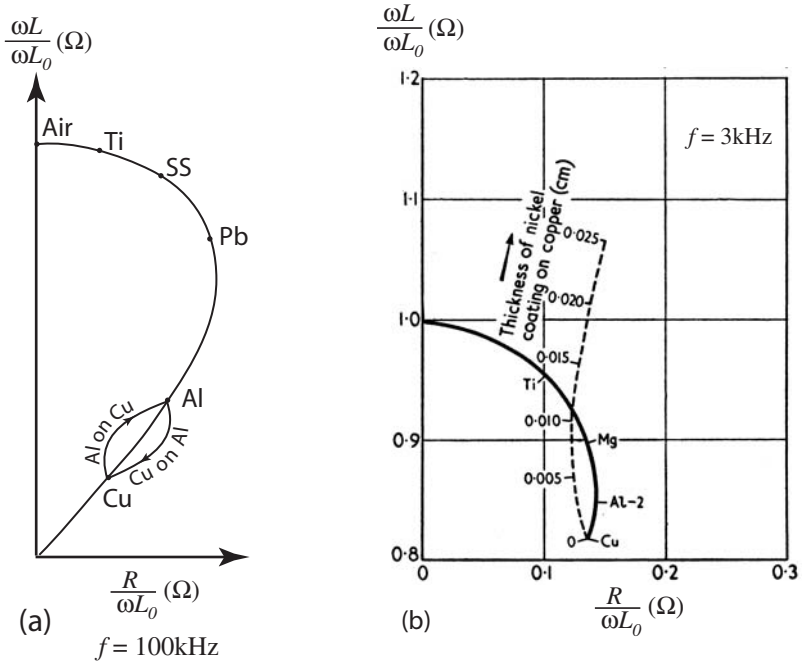


FIGURE 5.42 Impedance loci for conducting coatings starting with zero coating thickness to a thickness greater than 5 δ: (a) aluminum coated onto copper and copper coated onto aluminum, and (b) aluminum coated onto steel.

Copyright © Marcel Dekker, Inc. All rights reserved.

coatings; or it can represent a discontinuity in the structure, e.g., corrosion in a lap joint on a aluminum aircraft wing. For EC testing the thickness of at least one conductor must be less than several skindepths; otherwise, the probe field cannot see the gap. The details of the impedance loci—the peak values of a family of probe response trajectories—for varying gap thickness depend on:

Whether the two conductors are of the same material or different materials
Whether the combined thickness of the two conductors is greater or less than 5δ.

Let us assume the combined thickness of the two materials, excluding the gap, is greater than 5δ: an electromagnetically thick material. To predict the impedance trajectory of increasing gap size, we return to the engineering concept of examining the extremes: zero and infinite gap thickness. The extremes suggest where the impedance locus would begin and where it should end. If the gap equals zero, there are two possibilities: either the two materials are identical, or they are different. If they are identical, we simply have one block and the impedance falls on the conductivity comma curve. Different materials imply a coating, and the impedance falls on the coating trajectory according to its thickness. If the gap becomes large, the eddy currents penetrate through the thin material—the probe is on this side—and do not reach the second material. Therefore, the EC probe "sees" the impedance of a thin sheet.

What does the trajectory for increasing gap look like between these two extremes? As the gap increases from zero, the thick material's influence on the probe impedance diminishes, while the influence of the thin material remains constant. This suggests that initially, as the gap increases from zero, the response trajectories are similar to liftoff and not, for example, to a thinning sheet. As the gap increases, we know the impedance locus must cross the conductivity curve and then end with the impedance of a thin sheet. The final impedance locus is shown in Figure 5.43.

The exact impedance locus for a given set of materials, including the gap material, for increasing gap thickness would require empirical data or at least a quality model. Still, a reasonable curve can be deduced with knowledge of the fundamentals of the eddy current's interaction with materials.

Eddy Current Measurement Optimization

To optimize measurement of a specific parameter, we need to reduce the effects of the other parameters. Previously, we mentioned a simple phase rotation to reduce the effect of liftoff (second part of Section 5.3.2). In this section, we present two additional optimization techniques: offset null balance and multifrequency.

Optimizing the Frequency Through Offset Null Balance. For impedance plane measurements, we have information about both the magnitude and

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

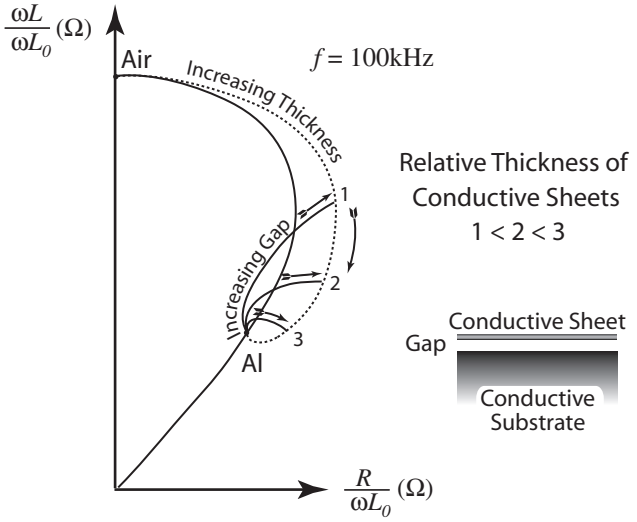


FIGURE 5.43 Impedance loci for a gap between electromagnetically thin and thick material.

the phase of the signal. On the display we see only a small amplified region of the signal, representing the difference between our reference vector $\vec{Z}_1^n$ and the response trajectory $\vec{Z}_2^n$ (Figure 5.31). These vectors begin at the origin $\vec{Z}^n = 0 + j0$. By altering the location of the "origin" of these vectors, we can selectively reduce the effect of unwanted variables while enhancing the interest variable. This shifting of the "origin" is called *offset null balance*. Equipment that allows offset null balance offers control of offset values through adjustable biasing resistance and reactance.

How does offset null balance work? Assume we wish to measure changes in conductivity, but probe wobble (liftoff) is corrupting the response signal. Previously, we simply rotated the signal such that the liftoff response was along the x-axis. Although we could amplify the signal in the y-direction, this method does not reduce the effect of the liftoff. Instead, we offset the origin to a new location and choose a frequency such that the magnitude of the response vector $\vec{Z}^n$ does not change significantly with a change in liftoff, and changes greatly with a change in conductivity. Figure 5.44 shows such a new "origin" for $f = 30$ kHz (1) and 1 MHz (2). Compare, in the figure, the magnitude change in $\vec{Z}^n$ for a variation in liftoff to a variation in conductivity. (The short dotted line represents the amplitude change on the meter for a change in conductivity. Also, $\vec{Z}^n$ is only marked on one response vector.) Note that the frequency of operation dictates the angular difference between the liftoff and conductivity curves. Note

Copyright © Marcel Dekker, Inc. All rights reserved.

the difference in null position for $f = 30$ kHz versus 1 MHz. For maximum sensitivity to the interest variable and minimum sensitivity to the unwanted variable, this angle should be as close to 90° as possible, and the offset null "origin" should lie radially to the locus of the unwanted variable. Figure 5.44 also depicts the optimum offset null "origin" and frequency combination for high sensitivity to liftoff and low sensitivity to conductivity (3).

As is usual with engineering systems, optimizing a single approach to detection comes at a cost. In this case, high frequencies tend to produce the maximum angular difference between the liftoff and conductive curves. This difference implies that the EC probe will have low penetration (small skindepth). Similarly, for crack detection discrimination from liftoff is maximized at high frequencies. Therefore, unless the cracks are surface-breaking and very shallow, the advantages of the offset null balance will have to be weighed against the reduced depth sensitivity.

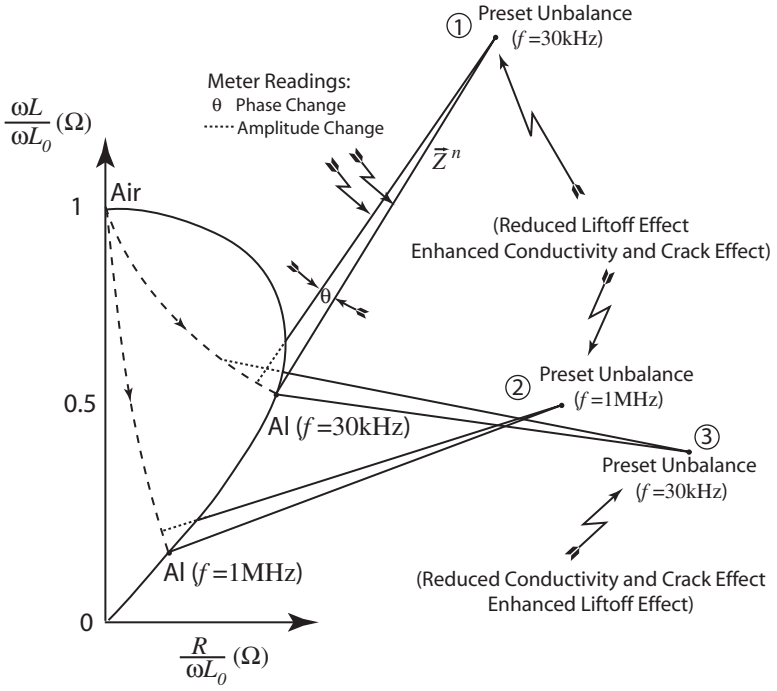


FIGURE 5.44 Offset null balance points to enhance cracks and conductivity variations while reducing liftoff effects for $f = 30$ kHz (1) and 1 MHz (2). An offset null balance point to enhance liftoff variations and reduce sensitivity to cracks and conductivity variations is also shown for $f = 30$ kHz (3).

Copyright © Marcel Dekker, Inc. All rights reserved.

Multifrequency-Multiparameter EC Method. EC sensors are very sensitive to many different parameters, all related to either a change in conductivity or magnetic permeability. We have seen how to differentiate between two parameters by choosing a frequency such that the angle between the two parameters is large, ideally 90° ($f_{90°}$). Then we chose an offset null origin that would enhance one parameter at a sacrifice of the other. Unfortunately, this simple method handles the variation of only two simultaneous parameters; in fact, we are often interested in a parameter that is superimposed on multiple unwanted parameters—liftoff, changes in electrical properties, cracks, geometry, etc. Multifrequency method offers a technique to extract a single parameter masked by its superposition of responses from many different sources. The following explanation has been adapted from Ref. 26.

Multifrequency EC method works much like solving simultaneous equations. Individual parameter responses are frequency-dependent. By measuring the response at multiple frequencies, we can treat each frequency response like a single algebraic equation. Thus we algebraically manipulate (add, subtract, multiply, and phase-rotate) the different frequency responses until the desired parameter is accentuated and the unwanted parameters are minimized. Figure 5.45a shows the response for several different defects in a tube at three different frequencies. In this calibration test, the signals represent the EC probe scanning along the calibration tube, where the probe encounters only one discontinuity at a time. (Remember that the impedance plane response is not a function of time, but of the accumulation of responses as the probe is scanned.) Thus each curve shows the probe response to an internal (ID) and external (OD) discontinuity, a through-wall hole (100%), a dent, a baffle plate, and a change in magnetic permeability (magnetite). At $f_1 = 20$ kHz, the EC probe responds very well to a change in magnetite and the presence of the baffle plate, but the probe can barely detect the presence of the ID and OD discontinuity or the 100% hole. As the frequency of excitation is increased, both the magnitude and the phase of the responses from all of the parameters change greatly. At $f_2 = 100$ kHz the probe has nearly equal magnitude response to all the discontinuities except the dent. At $f_3 = 500$ kHz only the ID, 100% hole, and dent are detected. This variation in sensitivity to features with frequency is to some degree predictable by the skin depth effect:

- At low frequencies (large skin depth), the highly penetrating eddy currents readily detect the baffle plate and the magnetite. But the nonlocalized eddy currents are not as sensitive to the ID, OD, and 100% discontinuities. There is little phase difference between eddy currents at the interior and exterior surfaces to differentiate these discontinuities.
- At intermediate frequencies, the eddy currents are more localized with significant phase difference between interior and exterior surfaces readily detect all discontinuities with clear phase discrimination.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

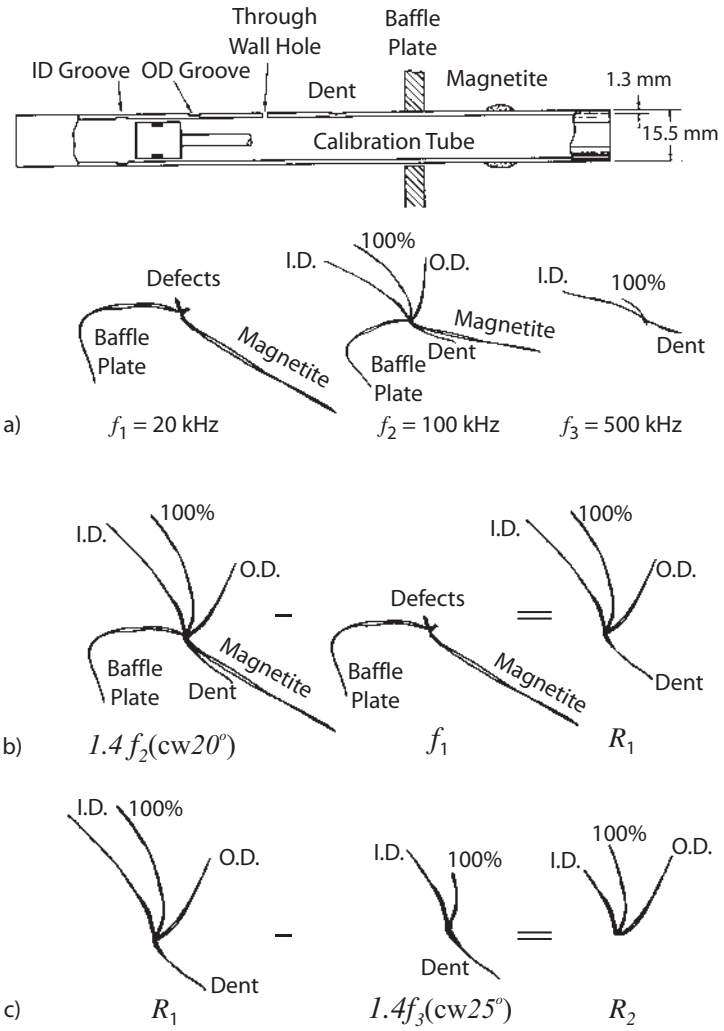


FIGURE 5.45 Multifrequency EC testing. (a) Scan results on a calibration tube with spatially separated flaws at $f = 20$, 200, and 500 kHz. (b) Results of algebraic manipulation of the data at $f = 20$ and 100 kHz. (cw 20° indicates a clockwise rotation of 20°.) (c) Results of algebraic manipulation of a scan of the dent only ($f = 20$ and 100 kHz).

Copyright © Marcel Dekker, Inc. All rights reserved.

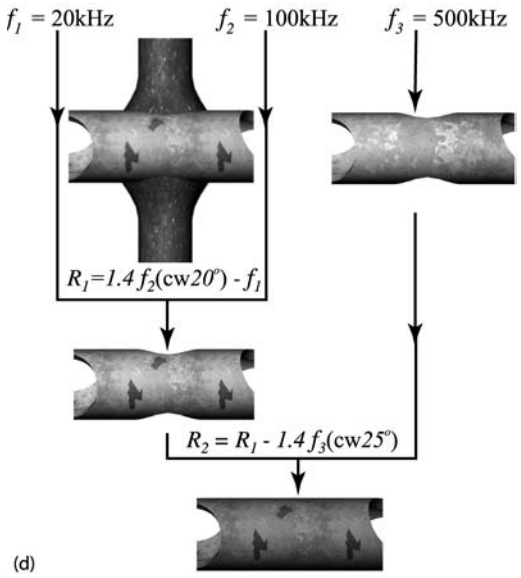


FIGURE 5.45 (*continued*) (d) Schematic of multifrequency EC testing on a heat exchanger tubing to remove the unwanted signal response from the baffle plate, magnetite deposits, and tube denting and the enhancement of the desired response from corrosion or cracking. (Adapted from VS Cecco, G Van Drunen, and FL Sharp. Eddy Current Manual, Vol. 1. Ontario: Chalk River Nuclear Laboratories,1981, p 156.)

At high frequencies, the small skin depth limits probe sensitivity to internal defects.

Given these three frequency responses, how can we manipulate the data algebraically to remove the effect of unwanted parameters (e.g., those from the baffle plate, magnetite, and a dent) so that we may readily detect only the parameter of interest (e.g., the internal or external wall thickness discontinuities)? First, we remove the effects of the baffle plate and the magnetite. We note that the probe response at both f_1 and f_2 shows sensitivity to both of the unwanted parameters. We multiply the response at f_2 by 1.4 and rotate the phase of the signal CW by about 20°, calling the resultant signal R_0. When we subtract the response at f_1 from R_0 (call this R_1), the resultant signal removes the effects of the unwanted signals and enhanced the interest signals (Figure 5.45b). Similarly, the effect of a dent can be removed, as shown in Figure 5.45c, where the probe scanned the dent at f_2 and f_3 only. Thus, given a calibrated specimen with clear, spatially separated discontinuities, we can determine (a) which probe frequencies

Copyright © Marcel Dekker, Inc. All rights reserved.

have high sensitivity to specific parameters and (b) which algebraic manipulations will remove unwanted signals and enhance the response from the desired parameters. The effects of these manipulations are characterized in Figure 5.45d.

We now apply this multifrequency method to an actual inspection, of a heat exchanger tube. Assume we want to inspect the tube for pitting or cracks at the location of a baffle plate (Figure 5.46a). Magnetite deposits have accumulated around the baffle plate, and during manufacturing the tube has been slightly dented at this location. Using the calibrated data from above, we know the appropriate frequencies and algebraic manipulations to remove the effects of the baffle plate, magnetite, and dent. Figure 5.46b shows the defective tube scan at $f = 20$, 100, and 500 kHz. Note since all of the "flaws" are spatial superimposed so is the signal response. Figure 5.46c shows the algebraic manipulation of both calibration results and the scan of the defective tube. The final signal R_2 shows pitting about 50% of the wall thickness.

In the multifrequency EC method, the excitation superimposes all the desired frequencies in one signal. Upon reception, the signal is filtered to extract the individual frequency data. Multifrequency EC method is a powerful tool to extract individual parameter effects otherwise masked by the response from unwanted parameters.

5.4.2 Remote Field Eddy Current (Optional) (27,28,29,30)

Remote field eddy current (RFEC)—commonly referred to in the industry as Remote Field Testing (RFT)—is a specialized method for inspecting tubes. It is primarily applied to ferrous material, and is used extensively in the inspection of (a) the carbon steel tubing used in steam boilers for power generators, (b) small-diameter heat-exchanger tubing, and (c) oil-well casing, for which it was originally developed by MacClean (31) and Schmidt (32). In conventional EC methods, the response of the induced eddy currents—the eddy currents that oppose the excitation (primary) field—to defects and material properties provides the information. Conventional EC inspection is limited by the depth of penetration, only a few skin depths. On the other hand, unlike conventional EC inspection, RFT monopolizes on the existence of a secondary field that completely penetrates the wall thickness, thus allowing through-thickness interrogation. RFT requires a specialized probe configuration and methods of visualization and analysis. Because of the unique features and methods of the RFT, this section details the theory and the visualization and interpretation methods.

RFT Basics

RFT relies on unique characteristics of the magnetic field caused by the specific geometric relationship between the excitation coil and the inspection material (the tube). The basic RFT setup (Figure 5.47) consists of a large internal excitation

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

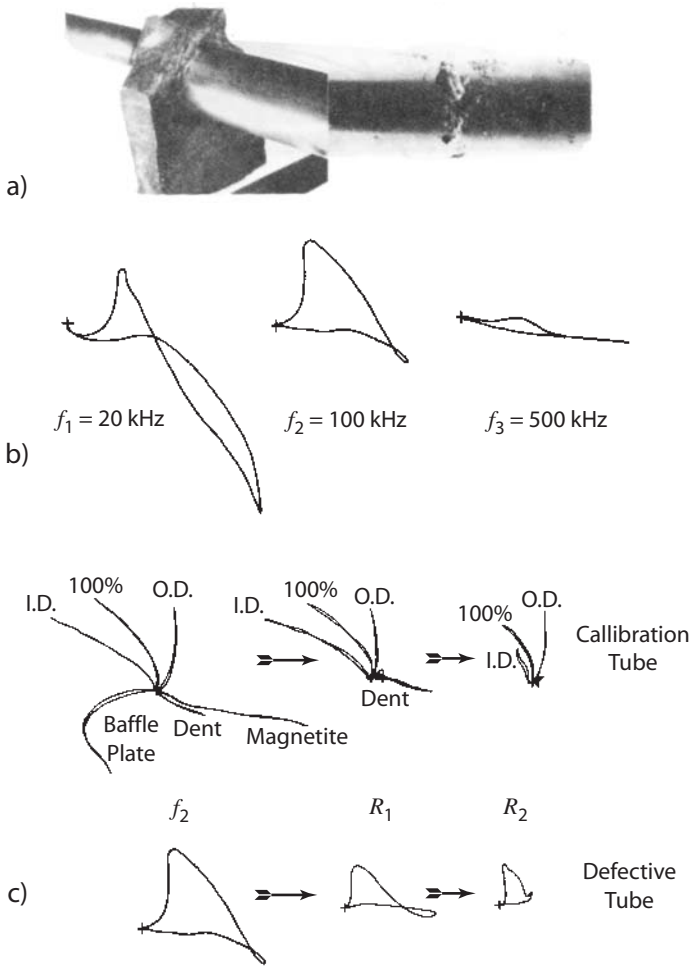


FIGURE 5.46 Multifrequency EC testing on an actual heat exchanger tube. (a) Photo heat exchanger tube at the location of the baffle plate. Note magnetite has built up to the point that the spacing between the plate and the tube has been eliminated. (b) Probe response of the tube scan at $f = 20$, 200, and 500 kHz. (c) Results of algebraic manipulation of the data. Pitting now clearly indicated. (Adapted from VS Cecco, G Van Drunen, and FL Sharp. Eddy Current Manual, Vol. 1. Ontario: Chalk River Nuclear Laboratories, 1981, p 161.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

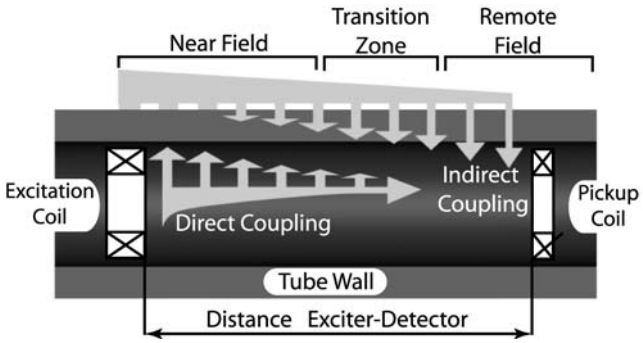


FIGURE 5.47 Schematic of RFT system depicting the direct and the indirect magnetic coupling field path. Conventional EC involves the direct field (near field) and RFT responds to the indirect field (far field).

coil and a pickup coil inserted into a tube. Upon excitation, the primary magnetic field induces eddy currents in the tube wall proportional to the strength of the excitation field. These induced eddy currents produce a magnetic field that opposes the primary field, thereby reducing the direct field impinging on the inside tube wall and thus changing the impedance of the primary coil. Impedance plane analysis interrogates the local material. This interaction in the vicinity of the excitation coil is equivalent to conventional EC inspection using a bobbin probe, but due to the geometry a magnetic field will also appear on the exterior (OD) of the tube. This field is guided axially along the tube. As the field propagates, it rediffuses back into the interior (ID) of the tube. Thus we have two fields inside the tube: the direct field from the excitation coil and the indirect field that has twice penetrated the thickness of the tube. RFT monopolizes on the indirect field to interrogate the entire thickness of the tube.

Schmidt (33) measured the magnitude and phase of the total (direct plus indirect) magnetic field on the interior of the pipe (Figure 5.48). The trace labeled "Amplitude–Inner Wall" represents the horizontal magnetic field component as the detector coil retreats axially from the excitation coil. Note the expected rapid decay in the vicinity of the coil (near field). At a distance of ~1.7 IDs, however, a sudden phase shift occurs, and the amplitude increases to a local maximum at ~2.5 IDs; it then decays at a distinctly slower rate (remote field). The distance at which the phase shift occurs marks the end of the near field and the beginning of the remote field. Typically, practitioners define a transition zone as the proximal region to the rapid phase change. Additionally, Schmidt measured the magnetic field exterior to the tube. Note that the difference of the OD and ID field strength is about four orders of magnitude in the near field, and only about one order of magnitude in the remote field. Most importantly, there is no dip in the OD field,

Copyright © Marcel Dekker, Inc. All rights reserved.

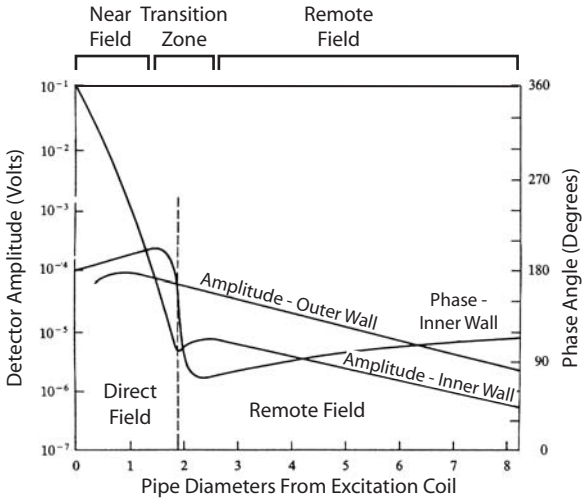


FIGURE 5.48 Plot of the magnetic field magnitude and the phase on the tube interior as the detector retreats from the excitor coil. The phase inversion separates the near and the far fields. The magnetic field magnitude measured on the tube exterior is also plotted. (Adapted from TR Schmidt. The Casing Inspection Tool: An Instrument for the In-Situ Detection of External Casing Corrosion in Oil Wells. Corrosion 17:81, July 1961.)

and the slopes of the two fields are about equal in the remote field, indicating a direct relationship between them in this region.

Typically the detector is placed two to three tube diameters (ID) from the excitation coil. The detector can be absolute or differential, and is orientated as an axial probe or a surface probe, i.e., parallel or perpendicular to the surface of the tube. The detector does not require a saturating field to work on ferromagnetic materials. Typical excitation frequencies are between 10–300 Hz. The response signal is on the order of microvolts. Although small, this signal consists only of the response from the indirect field. In conventional EC inspection, the response signal must be extracted from the excitation signal (recall the balance bridge instrument in the first part of Section 5.3.2). The RFT probe has low sensitivity to fill-factor changes and probe wobble. At frequencies of 10–30 Hz, the probe can penetrate up to 1 in. of steel. Most importantly, the probe responds to ID and OD discontinuities with nearly equal sensitivity.

RFT Response Visualization and Interpretation

In RFT, we measure a magnetic field that has twice penetrated the thickness of the tube wall and traversed the axial distance between the excitation and pickup coils.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

The traversal between the coils is nominally along the outer tube wall, and incurs low attenuation and phase lag. On the other hand, the skin depth effect predicts large attenuation and phase lag as the magnetic field diffuses through the walls. Therefore, any variation in the wall (thickness caused by corrosion, erosion, or fretting or material properties such as magnetite deposits) will change both the amplitude and the phase of the received signal.

For example, assume the RFT probe inspects a tube whose walls are thinning due to corrosion. As the probe scans from a region of full wall-thickness to the thinned area, the signal amplitude increases and the phase lag accordingly decreases. A phase vector polar plot of this response is the basis for the RFT visual scheme. Note that this method is a direct measurement of the field—in contrast to impedance plane analysis, where the reflected or transmitted field changes the impedance of the measurement coil. Before detailing the RFT response plot, let us review the skin depth relationship.

The skin depth equations previously developed to predict the amplitude and phase changes assumed a planar, semi-infinite solid. Nonetheless these equations are commonly used with reasonable accuracy (34). The fundamental relationship for the magnetic field density decay and phase lag as it diffuses through a medium is

$$B = B_0 e^{-x\sqrt{\pi f \sigma \mu}} e^{-j(x\sqrt{\pi f \sigma \mu} - \omega t)}$$

or

$$B = B_0 e^{-x\sqrt{\pi f \sigma \mu}} \sin(\omega t - x\sqrt{\pi f \sigma \mu})) \tag{5.78}$$

assuming harmonic (sinusoidal) excitation. Recall that B is the field density at a distance x below the surface, and B_0 is the field density at the surface. Note that the amplitude decays according to the first exponential term, and the phase lags according to the second term. Thus the skin depth (distance for the signal decay to e^{-1} and the phase to lag by one radian) is

$$\delta = \frac{1}{\sqrt{\pi f \sigma \mu}} \tag{5.79}$$

and the concomitant phase lag is

$$\phi = \frac{d}{\sqrt{\pi f \sigma \mu}} \tag{5.80}$$

where d is distance. Thus the signal strength can be controlled by frequency. Typically, thick walls are inspected with lower frequencies than thin walls. But phase measurement is very important in RFT analysis. As the frequency is lowered to increase the signal magnitude, the phase lag decreases. As usual with engineering problems, a balance between the two competing effects must be achieved.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

In support of the majority of the phase lag occurring in the wall transit, some people think of the EC diffusion as possessing wavelike characteristics—a critically "crushed" wave. In this scenario, we can define a phase velocity as

$$v = \frac{\partial x}{\partial t} = \frac{\partial(\omega t - x\sqrt{\pi f \sigma \mu})/\partial t}{\partial(\omega t - x\sqrt{\pi f \sigma \mu})/\partial x} = \sqrt{\frac{f}{\pi \sigma \mu}} \tag{5.81}$$

For a steel tube the "wave" propagates at about 1 m/s, while the guided magnetic field on the exterior surface travels at about 10 m/s.

RFT Vector Polar Plot: The Attenuation Spiral. The attenuation spiral provides a simple method to monitor the signal attenuation and phase lag for a through-transmission EC signal from an absolute probe (Figure 5.49). Assuming zero wall thickness, the signal would have maximum amplitude and zero phase lag. This phase vector is oriented along the *y*-axis on the attenuation spiral. As the wall thickness increases, the magnitude decreases as the phase lag increases, thus tracing a spiral. Therefore, at a given wall thickness there is a specific vector with magnitude and phase lag related through the skin depth equations. Note that for RFT, the sum of the through-transmission responses at the excitation and the pickup coil plus the axial traversal between the coils comprise the total detected signal.

Generally, we are interested in only a small portion of the attenuation spiral. An inspection of a specific pipe would have a full wall response vector as a

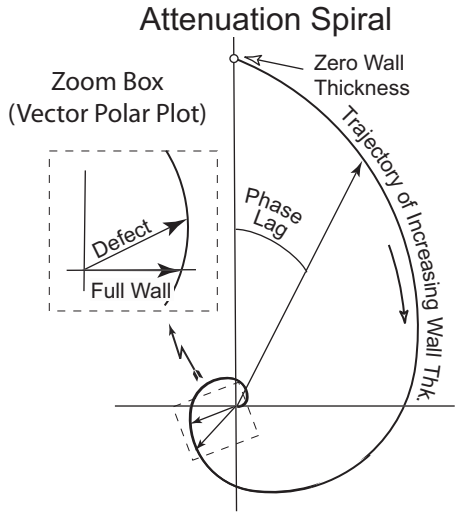


FIGURE 5.49 RFT probe response plotted on the *Attenuation Spiral* as the wall thickness increases from zero thickness to infinite thickness.

Copyright © Marcel Dekker, Inc. All rights reserved.

reference signal. Then perturbations from this vector would indicate discontinuities such as wall thinning due to fretting or corrosion. To improve signal visualization, the signal is amplified and rotated such that the full wall vector provides a clear base signal along the *x*-axis (inset Figure 5.49). *Zoom box* refers to this amplification and phase rotation, and *vector polar plot* denotes the resultant graph.

As a tube is scanned, the tip of the vector traces the changes as the probe traverses a discontinuity. The display shows the entire accumulated signals during the scan.

In conjunction with the amplitude and phase monitoring, one generally wants to know the position of the discontinuities. Therefore, most RFT systems incorporate a simple strip chart in addition to the vector polar plot. The strip chart measures probe position verses amplitude or phase. Two charts allow both. Computer storage allows signal play back for further analysis. Cursors on the strip charts and the vector polar plot align the signals with distance. Figure 5.50 shows the response for a scan on a tube with a discontinuity on an amplitude strip chart and a vector polar plot.

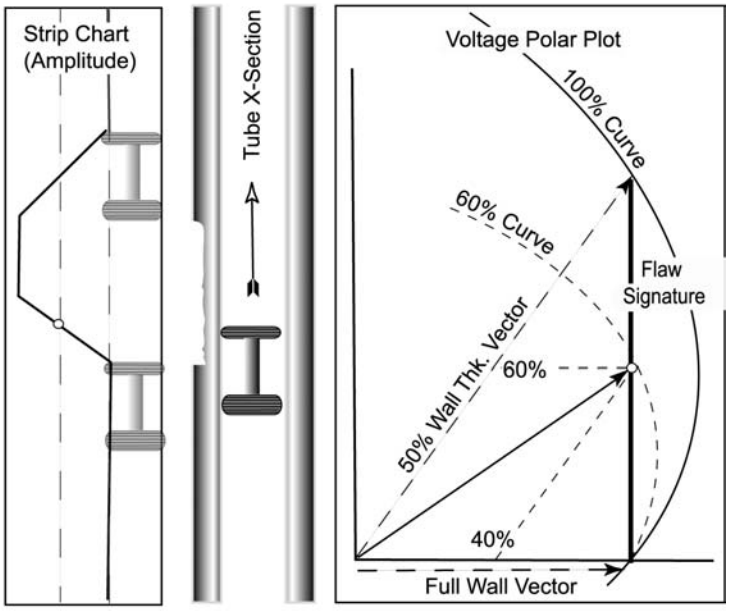


FIGURE 5.50 RFT response to a flaw of 50% wall thinning. The strip chart is used to record position of the flaw. Cursors (open circles) mark correspondence between plots.

Copyright © Marcel Dekker, Inc. All rights reserved.

Discontinuity Signatures. A combination of the strip charts and the vector polar plot monitors three RFT parameters: (a) probe position, (b) the change in magnitude, and (c) phase lag in response to the double through-transmission of the tube. The discontinuity signature specifics are detailed not only by the discontinuity itself, but also by the RFT probe configuration (absolute, differential, radial, or circumferential detectors, as well as excitation coil geometry). This section presents a representative discontinuity signature using an absolute RFT probe.

Wall reduction by corrosion and fretting is of significant concern in many tubing applications. The unique through-transmission and rapid-inspection properties of RFT allow us to assess the integrity of wall thickness. Figure 5.50 depicts a combined amplitude strip chart and a vector polar plot; the polar plot has been normalized and rotated such that the full wall thickness lies along the *x*-axis. The discontinuity (wall thinning) is long compared to the probe length, as shown in the diagram with the pipe cross-section and the probe. Therefore, the scanning probe encounters the following sequence:

> full wall→transition into the thinned region → approximately constant thinned wall → transition from the thinned region → full wall

The probe response vector to full wall thickness begins at the origin and ends on the attenuation curve. Likewise, the response vector to an approximately constant thinned wall begins at the origin and ends on the attenuation spiral, with increased amplitude and decreased phase lag over the full wall signal. During the transition, however, the probe "sees" a combination of the two signals. Recall that the RFT probe has two through-transmission trips, one at the exciter and one at the detector. Assume the probe lies 60% through the abrupt transition from full wall to 50% wall thickness. The response would be the vector sum of 60% of the reduced wall vector and 40% of the full wall thickness vector, as indicated. The dashed 60% attenuation spiral traces the response for varying wall thickness when the probe lies 60% through the abrupt transition (25%, 50%, and 75% attenuation spirals are often drawn on the voltage plane). Thus the vector tip trace for the scan begins at the full wall signal, moves through the 60% position to the reduced wall signal, and then returns along the same path. Note that the strip chart signal is skewed relative to the discontinuity position.

There are several important points to remember while reading these charts. We know that relative to the full wall thickness, the reduction in wall thickness reduces the phase lag. Therefore the angle between the full and reduced wall vectors (the phase lag) is directly proportional to the wall reduction relative to the full wall dimension. Also, the strip chart response length includes the discontinuity length *plus* the probe length. If the discontinuity length had been less than the probe length, the trace would not have extended to the 50% reduced wall vector. Also, a nonsymmetric discontinuity would trace a loop, as we have seen

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

for conventional impedance plane EC. A differential probe scanning a long discontinuity would produce a strip chart signal with asymmetrical "blips" in the transition region only, and no discontinuity signal in-between; recall that the differential probe has strong common mode rejection. The vector plot for a differential probe has a trace in the first and third quadrants passing through the origin.

5.4.3 EC Calibration

Nearly all EC testing requires a comparison of the response signal to some calibrated reference signal. The reference signals are obtained from a well-characterized material containing discontinuities or material properties similar to those to be measured. These calibrated reference pieces are called *artifact standards*. Typical EC artifact standard pieces are manufactured containing artificial discontinuities, usually EDM (electrical discharge machine) notches, flat bottom drilled holes, or a series of blocks containing different coating thicknesses or conductivities. Artifact standards exist for surface, bolt hole, and tubing inspection.

One problem with artifact standards as a calibration method is that the method relies on the quality and consistency of the physical piece; another is that the discontinuity standards are not actual cracks or void but poor imitations. For example, EDM cracks have a constant width over the entire crack—unlike a real crack, which is tapered if not closed at some points along the crack walls.

EC calibration uses artifact standards in a number of ways. In the multi-frequency EC method, calibration using artifact standards determines which frequencies respond to a given type of discontinuity. For discontinuity detection, response shapes on model flaws (artifact standards) are used to help identify the type and character of an actual discontinuity. Probes with either internal or external references use artifact standards as a comparison signal to help identify the conductivity or magnetic permeability of a given material. In tube inspection, a tube with dimensions and material properties similar to the test material is manufactured with model discontinuities, and is scanned both to obtain a reference signal and to quantify discontinuity type and size. Finally, artifact standards are used to calibrate EC instrumentation.

5.5 APPLICATIONS

5.5.1 Remote Field Testing (35–38)*

Remote field testing (RFT) is an electromagnetic method for NDE of metal tubes. It is widely used to examine carbon steel tubes, such as those found in heat

*Application submitted by David Mackintosh, Russell NDE Systems, Inc., Edmonton, Alberta, Canada.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

exchangers and boilers. Since its commercial growth in the late 1980s, RFT has led to large increases in efficiency in many heavy industries. For example, before RFT was available, leaky boilers often had to be discarded, because there was no economically viable way to determine the nature and extent of damage. By locating and characterizing damage, RFT allows boilers, heat exchangers, and other such equipment to be kept operating safely and near peak efficiency, thus reducing the likelihood of unplanned shutdowns (Figure 5.51). RFT also reduces the risk of leaks that could lead to explosions.

RFT has many advantages over other NDE methods used in tube inspection. It can interrogate full-wall-thickness even in steel tubing, whereas conventional eddy current suffers from low sensitivity caused by the low penetration (skin depth effect), and thus has almost no sensitivity to external tube flaws. Although ultrasonic techniques such as the *Internal Rotary Inspection System* (IRIS) accurately measure wall thickness, they are very slow (probe pull speed is typically 1 in./sec, compared to 1 ft/sec with RFT) and require extensive tube cleaning. RFT requires only minimal tube preparation.

RFT presents some unique challenges. In an RFT probe, the electromagnetic field travels some distance on the outside of the tube. Hence any conductive or ferromagnetic object on or near the tube creates an indication that may obscure, distort, or be mistaken for a flaw indication. Another source of uncertainty is the magnetic permeability of a given grade of carbon steel, which can vary considerably between tubes. To compensate for changes in



FIGURE 5.51 Demonstration of RFT on a tube bundle extracted from a heat exchanger. The technician on the left is pulling the RFT probe, while the technician on the right operates the instrument, storing examination data by tube name. Only the tubes on the outside of the bundle can be examined visually; hence the need for RFT. The thick circular plate where the ends of the tubes are visible is called the tubesheet. The thinner plates further to the right are the tube support plates. (Courtesy Russell NDE Systems.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

permeability, the RFT operating frequency may be varied during an examination, or the RFT examination may be followed up with an ultrasonic examination to verify or adjust results at specific locations.

RFT probes are designed to accommodate the harsh environment of tubing inspection. The tubes typically have significant, very abrasive corrosion. As shown in Figure 5.51, the pull distances can be very long and the probe speeds quite high. The probes must also negotiate bends in the tubing. Figure 5.52 shows a variety of commercial RFT probes.

Damage in Heat Exchangers

In a heat exchanger, one fluid inside the tubes exchanges heat with another fluid on the outside of the tubes. Tubes are often made of carbon steel that meets design requirements in such areas as thermal conductivity, corrosion resistance, availability, and cost. Typical damage in heat exchangers is shown in Figure 5.53, and many of the indicated flaws appear in the sample data in Figures 5.54a and 5.54b. Metal loss by erosion is relatively easy to detect and characterize because of its large volume and smooth contours. Metal loss near a tube's support plates is more difficult to detect and characterize, because the plates impede the field traveling on the outside of the tube.

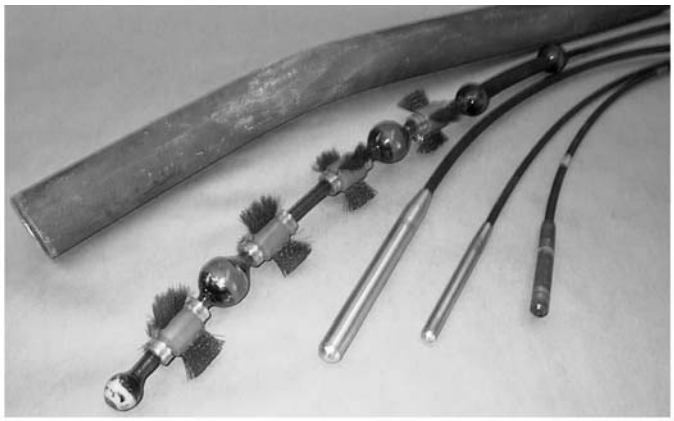


FIGURE 5.52 Typical RFT probes. The flexible probe (with centering brushes, left) is suitable for examining boiler tubes with bends, such as the sample tube in the figure. The two wear-resistant rigid probes with stainless steel casing (center) could be used to examine heat exchanger tubes. The probe without a casing (right) reveals the coil layout. (Courtesy Russell NDE Systems.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

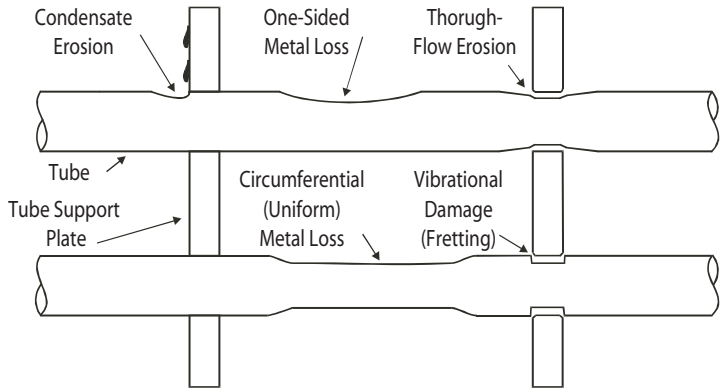


FIGURE 5.53 Common flaws in heat exchangers (simplified layout with exaggerated flaw dimensions).

Damage in Boilers

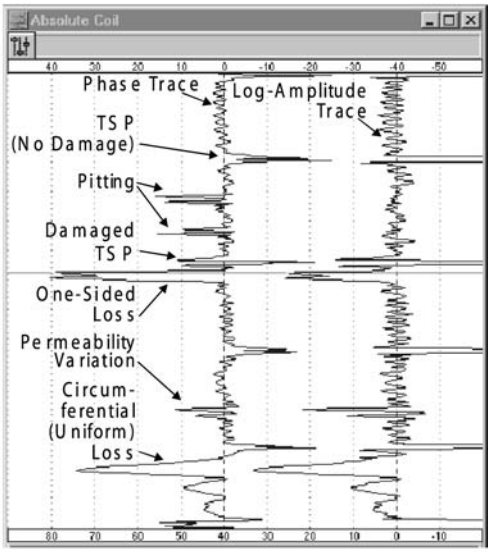
The function of a boiler is to turn water into steam. Sizes of boilers found in industry range from room-sized package boilers to boilers in power plants, several stories high. In coal-fired boilers, damage often takes the form of fly-ash erosion, caused by combustion products impinging on the tubes. One-sided metal loss due to erosion is commonly found opposite soot-blowers (lances that blow steam to remove soot deposits). Erosion takes place most often where gases are hot and moving fast: in high-velocity soot-blower lanes, in the top third of boilers, and near superheaters (plattens). Data displays such as those in Figures 5.54a and 5.54b allow the type of metal loss to be determined. Corrosion is also common in areas where sediment such as ash or byproducts of combustion can settle and combine with moisture to become acidic—e.g., on tubes at the mud drum in a generator bank. Near-drum examination is a challenge for RFT and requires a special probe. In boiler examinations, it is essential to be familiar with the mechanical layout in order to match boiler features with indications in the RFT data. Steel attachments on the outside of the tube create indications, as will other tubes in close proximity.

Field Methods

Before an RFT job, establish a tube numbering system and set the extent of the examination (whether coverage is to be complete or partial, whether a flexible probe is needed for bends, etc.). Obtain as much history as possible on the component to be examined. Ensure that the tubes are clean enough to allow passage of the probe, and where possible remove heavy deposits of magnetic

Copyright © Marcel Dekker, Inc. All rights reserved.

(a)



(b)

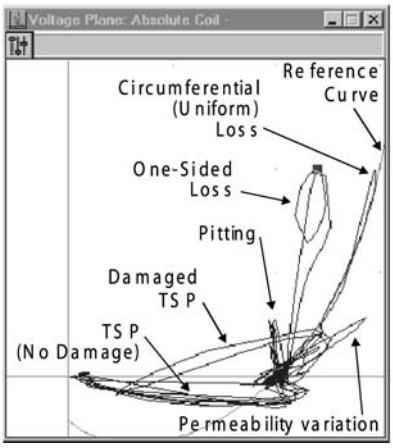


FIGURE 5.54 (a) Sample RFT data from an absolute coil probe in carbon steel tubing; strip chart view. Phase and log-amplitude traces are useful because both quantities are linear indicators of the depth of uniform wall loss. Except for the permeability variation, all leftward deflections of traces shown here are metal loss. The ''noise'' is due to periodic permeability variations in the tube wall. (Compare with Figure 5.54b.) (b) Sample RFT data from Figure 5.54a: Voltage Plane view. The Voltage Plane displays phase and amplitude data in polar format to help identify flaws by their indications. (TSP = tube support plate)

Copyright © Marcel Dekker, Inc. All rights reserved.

material. Set the RFT operating frequency using machined flaws in a reference tube (one technique is to set the frequency to obtain a certain angular spread between flaw indications on the voltage plane). Once the instrument settings are optimized, demonstrate sensitivity to flaws in the reference tube, pulling the probe at examination speeds. Acquire data from the tube bundle and evaluate the flaw indications. Often the client will require that all tubes with metal loss above a certain depth be marked for plugging. For complete details on field procedures, see the ASTM document *E2096-00 Standard Practice for In Situ Examination of Ferromagnetic Heat-Exchanger Tubes Using Remote Field Testing*.

5.5.2 Aerospace Applications*

NDE plays a major role in the aerospace industry. It has been said that if you designed an airplane that would never crack, it would be too heavy to fly. Aircraft airframes are made of materials that will bend and flex rather than break. NDE methods have become sophisticated enough that much of the airframe and wings can be inspected without disassembly. NDE allows inspectors to find flaws in aerospace components before they reach critical engineering thresholds that might lead to catastrophic failures.

EC testing can be used on a wide range of airframe components for detection and/or sizing of cracks or corrosion. The typical test units used in the industry today may work in either the single-frequency or the multiple-frequency mode. Advances in electronics have allowed these test units to become quite compact; even the smaller units (hand-held, 3.5-pound, battery-powered) can be used with a wide range of probe designs, test frequencies, and peripheral devices. Most on-aircraft inspections use these smaller portable units. Although EC testing is not always called for in the inspection codes, it can provide a more sensitive, timely, and therefore cost-effective inspection than other NDE processes for surface or near-surface discontinuities.

We know from earlier discussions that there are practical limits for how deep we can go into materials and still perform an effective EC test. We normally use something in the range of 0.2 in. as a typical value for "effective depth of penetration" ($\delta_{\text{eff}} = 3\delta$). If the test is to detect near-surface flaws, we would probably expect to be operating in the 100–500 kHz range. For discontinuities deeper in materials, we must use lower frequencies (100 Hz–1 kHz) as well as specially designed hybrid (excitation/pickup) probes. Some of the most critical aerospace components use very-high-strength alloys, which must be inspected using a "zero tolerance" approach. Even the smallest surface discontinuity, either from the manufacturing processes or from normal (or abnormal) operational

*Application submitted by Jim Cox, Training Director, Zetec, Inc., Issaquah, Washington.

Copyright © Marcel Dekker, Inc. All rights reserved.

conditions, could result in a rapidly propagating flaw. In these types of surface exams it is not unusual to use inspection frequencies well up into the megahertz range (1–10 MHz).

Engine Components

The rotary propulsion systems found on modern aircraft use alloys of higher strength than the typical aluminum alloys found in the bulk of the airframe. Jet engine components have to survive in a very harsh environment; they may have operating temperatures between 400–800°C and rotation speeds of 8000–40,000 revolutions per minute. The areas that require very careful inspection are not readily accessible when the engine is mounted on the wing or fuselage, so inspections can only be fully accomplished when the engine is disassembled.

Engines are removed and maintenance performed on an established schedule. This process is governed by the expected life cycle of the internal rotating components, how many full-power cycles that engine has gone through, or possible malfunctions that require repairs. The normal aircraft maintenance cycle falls between 2000 and 5000 flight hours, depending on design. The allowable operating cycles vary from one rotating component to another, even within the same engine. For this reason, inspection and maintenance schedules can be very complicated and must be closely tracked and carefully documented.

Some of the more critical engine components are the fan disk (and blades) as well as the compressor disk (and blades). The leading and trailing edges of the blades are exposed to potential damage from foreign objects while the airplane is taxiing or in flight. If a blade breaks off, the loose piece is ingested through the rest of the engine, breaking off more blades as it goes. The rotating parts become unbalanced, creating vibration that will severely damage the engine—and may cause it to break away from the wing.

EC testing is the most sensitive NDE method for finding small cracks in the blades. Probes are designed to fit the exact contour of the blade edge. Sometimes called J-probes, their coils are mounted at the optimum point in the probe body for detecting small cracks radiating out from these edges. If the probe is allowed to lose contact with the inspection surface, then the varying liftoff and edge effect signals can create a screen response that could be misinterpreted as a flaw.

Another area of interest is the fillet region at the root of the blade. Usually this region is inpsected by a pencil probe scanning back and forth across the fillet. The disks and blades are usually made of titanium or other high-strength metals with very low electrical conductivity. For this reason, the frequency used for the inspection is typically 2 MHz or higher.

The dovetail slots in the disk that the blades mount to also require careful inspection. Cracks may develop in these slots due to normal operational stresses. If these cracks are not detected when they are small (0.035 in. long and 0.015 in. deep), the disk may break apart, typically damaging the rest of the engine and

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

possibly also damaging the passenger compartment and other aircraft systems, as has happened with fuselage-mounted engines. Probe fixtures are manufactured to fit the exact contour of the slot. Some probes have as many as four coils mounted such that all critical stress points are inspected in one pass. Figure 5.55 shows an instrument that can display the signals from the four coils on the screen simultaneously. In most situations the inspector reviews this information in real time and decides if there are any indications of concern. One manufacturer requires that the instrument be capable of recording all of the EC data for further review and evaluation.

A standard flaw is used to calibrate the system. The typical screen display is adjusted so that all signals related to nonflaws (geometrical conditions) are oriented on the horizontal axis (x-axis) of the screen. The known flaws in the calibration standard are oriented to respond in the vertical direction (y-axis) when they are detected. Maximum detection potential can be demonstrated when there is a relatively large displayed phase angle difference between the known flaw responses in the standard and the geometry responses ($\theta_{ideal} = 90^\circ$). The operator, based on the test system's response to these standards, must be able to clearly distinguish between flawed and nonflawed conditions. The system's sensitivity to flaws is verified not only before inspection, but also perhaps at regular intervals during the test.



FIGURE 5.55 Eddy current inspection of a turbine disk. Note this particular system displays the response of a specialized probe with four eddy current sensors. (Courtesy Zetec, Inc.)

Copyright © Marcel Dekker, Inc. All rights reserved.

PROBLEMS

1. In order to induce secondary currents, the material in proximity to the primary field **must** be:

   (a) a conductor
   (b) an insulator
   (c) a ferromagnetic material
   (d) all of the above

2. Draw the orientation of an eddy current coil and all associated magnetic fields and currents to best detect a linear crack on a flat plate.
3. If an alternating current were applied to a resistive circuit, the phase difference between the voltage and current is

   (a) $+90^\circ$
   (b) -90°
   (c) $+180^\circ$
   (d) -180°
   (e) 0
   (f) not enough information

4. If an alternating current were applied to an inductive circuit, the phase difference between the voltage and current is

   (a) $+90^\circ$
   (b) -90°
   (c) $+180^\circ$
   (d) -180°
   (e) 0
   (f) not enough information

5. For Inconel at a probe operating frequency of 100 kHz (a) calculate the standard depth of penetration (1 δ) and (b) the associated phase lag.
6. If you want to interrogate Al-6061 for a suspected flaw beginning 1 mm below the surface, what operating frequency would you use to achieve $1\ \delta = 1$ mm?
7. Calculate the standard depth of penetration (1 δ) for carbon steel at a test frequency of 10 Hz. (a) without a saturation field and (b) with a saturation field. (Assume $\mu_r = 400$.)
8. Which material would have the largest standard depth of penetration:

   (a) 50% IAC
   (b) 87.2% IAC
   (c) 0.135% IAC
   (d) frequency must be known to answer the question

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

9. Show that $\delta = \sqrt{1/\pi\mu_r\mu_0\sigma f}$ can be written as $\delta = 50\rho/\mu_r f$ using the commonly used units of $\rho(\mu\Omega\,\text{cm})$, $f(\text{kHz})$, $\sigma(\text{mm})$, and $\mu_r(\text{dimensionless})$.
10. An EC probe has a resistance of $10\,\Omega$ and an inductance of $110\,\mu\text{H}$. What operating frequency is required to produce an impedance of (a) $50\,\Omega$ and (b) $200\,\Omega$.
11. For the previous problem what is the inductive and resistive impedance for an operating frequency of 60 kHz?
12. A decrease in material conductivity is equivalent to

    (a) an increase in magnetic permeability.
    (b) an increase in frequency.
    (c) an increase in resistivity.
    (d) a decrease in magnetic permeability.
    (e) an decrease in frequency.
    (f) a decrease in resistivity.

13. If the result of a measurement indicated a reactive impedance of $30\,\Omega$ and a resistive impedance of $15\,\Omega$, (a) what is the magnitude and phase of the total impedance? (b) The phase angle is a measure between what and what?
14. Liftoff signals can be utilized in

    (a) measuring permeability changes.
    (b) measuring conductivity changes.
    (c) measuring the thickness of nonconductive coatings.
    (d) determining proper test frequency.

15. Determine the effective diameter of an eddy current probe with an average diameter of 8 mm when operating at (a) $f = 50$ kHz and (b) $f = 500$ kHz when testing Al-7075.
16. State the advantages and disadvantages of (a) an absolute and (b) a differential probe.
17. For a bolt hole crack detection would an absolute or a differential probe be appropriate? Support your answer with an argument in words and state any assumptions.
18. Show that the limiting frequency can be written as $f_g = 5.07\rho/\mu_r D_0$ where the variables are given in common units used in EC: $\rho(\mu\Omega\,\text{cm})$, $D_0(\text{mm})$, and $f_g(\text{kHz})$.
19. Assume that for an inspection of a solid rod for longitudinal surface cracks, the best sensitivity/flaw discrimination is determined to be at the knee of the impedance curve. Calculate using Figure 5.20 the required frequency for a 10 mm diameter rod of (a) 316 SS and (b) copper.
20. Draw the orientation of an eddy current coil and all associated magnetic fields and currents to best detect a crack on a tube. Assume (a) a transverse crack and (b) a longitudinal crack.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

21. A Wheatstone balancing bridge has four legs. Its principal use in EC testing allows the operator to

    (a) sense smaller flaws.
    (b) inspect multiple materials at once.
    (c) testing using a reference block.
    (d) none of the above.

22. Assume that you wish to inspect the uniformity of a Mylar® film that has been metalized on both sides. As this material is to be used in a high voltage capacitor, the Mylar® thickness must not be less that a fixed amount. You suggest to your boss that eddy current might have potential. Argue and draw the impedance trajectory for a multilayer thin film system.
23. Although it has been suggested and tested, why is a magnetic saturation field undesirable for RFT inspection of carbon steel tubing?
24. From the typical impedance plane or vector polar (RFT) diagrams, can you tell the location of a flaw? Explain your answer.

GLOSSARY (39,40,41)

Absolute probe: Probe having a single sensing coil; it may or may not have a separate drive coil.

Alternating current: Current flow changing in amplitude and direction with time.

Anomaly: Unexpected, unclassified eddy current signal; a false defect indication.

Bobbin probe or coil: Eddy current probe designed to inspect the interior of a tube; also called an *interior* or *internal probe*, as it is inserted into the tube.

Bridge circuit: Electrical circuit consisting of four impedance arms designed to balance two of the arms against the other two arms; often called a *Wheatstone Bridge*.

Calibration standard: Test standard used to estimate defect size and setup/calibrate eddy current instrumentation.

Capacitive reactance: Reactance in a circuit that opposes any changes in voltage.

Capacitor: Electronic device that resists changes in voltage; stores energy in the form of an electric field.

Characteristic (or limiting) frequency: Combination of variables (conductivity, magnetic permeability, and length; e.g., tube diameter) that is normalized by frequency to create the *characteristic parameter* in frequency.

Characteristic frequency: Normalization of test parameters to a dimensionless *characteristic parameter* in frequency (frequency/frequency), allowing the

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

test coil operating point to be specified in terms of a single quantity rather than four independent variables (see *characteristic parameter*).

Characteristic length: Normalization of test parameters to a dimensionless *characteristic parameter* in length (length/length), allowing the test coil operating point to be specified in terms of a single quantity rather than as four independent variables (see *characteristic parameter*).

Characteristic parameter: Normalization of test parameters to a dimensionless parameter (such as frequency/frequency or length/length), allowing the test coil operating point to be specified in terms of a single quantity rather than four independent variables.

Circumferential coil: See *encircling probe* and *internal probe*.

Coil fill factor: For an encircling coil, the ratio of the diameter of the tube to the internal diameter of the coil; for an internal coaxial coil, the ratio of the diameter of the coil to the internal diameter of the tube.

Conductivity: Measure of the ability of a material to conduct electrical current.

Conductor: Material capable of carrying electrical current.

Current: Flow of electrons in a material.

Depth of penetration (standard): Depth at which the eddy current density has decreased to e^{-1} or 36.8% of the current density at the surface facing the eddy current probe (see *skin depth*).

Defect: Flaw or discontinuity that reduces a material's integrity, load-carrying capacity, or aesthetics; may or may not involve a loss of material.

Differential probe: Probe having two sensing coils located side-by-side. The probe may or may not have a separate drive coil.

Direct current: Current flow that is constant in amplitude and direction with time.

Discontinuity: A change in the material as sensed by an eddy current probe. A discontinuity may or may not be considered a flaw.

Eddy current: Closed loop alternating current flow induced in a conductor by a time-varying magnetic field.

Eddy current method: Electromagnetic NDT method based on the process of inducing electrical currents into a conductive material by magnetic induction and observing the interaction between the currents and the material; in France it is known as the *Foucault currents* method.

Edge effect: Signal obtained when a surface probe approaches a sample's edge.

Effective depth of penetration: Depth at which eddy current density drops off to 5% of the surface density; equivalent to 3δ.

Effective permeability: Magnetic permeability of a complex component, assuming the field strength in the magnetizing coil to be constant throughout the coil and sample.

Electromotive force (EMF): Force, measured in volts, supplied by a source of electric current that drives electric charges around a circuit.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Electromagnetic induction: Mechanism whereby a time-varying current induces a secondary current or EMF in a nearby circuit; the coupling mechanism is the magnetic field produced by the time-varying current.

Electromagnetic coupling: Electromagnetic interaction between two or more circuits; in eddy current examination, the specimen to be tested becomes a circuit (see *magnetic coupling; mutual inductance*).

Encircling probe: Probe that completely surrounds the test material; can be absolute or differential. Also referred to as a *feed-through coil*.

Feed-through coil: See *encircling probe*.

Ferrite: Ferromagnetic oxide material used for cores for eddy current probes to enhance the magnetic flux density, guide the magnetic field, or shield the probe from stray magnetic fields; the same material used in high-frequency transformers.

Flaw: See defect.

Ferromagnetic: Material with a relative magnetic permeability greater than 1.0.

Fill-factor: The fraction of the eddy current coil filled by the test sample; typically used in conjunction with an encircling or bobbin probe. It is a measure of coupling between the coil and test object.

Foucault currents method: Name by which the *Eddy current method* is known in France and parts of Canada.

Frequency: Number of cycles of alternating field (current, voltage, electric, or magnetic) per second.

Frequency (angular): Frequency expressed in radians per second; $\omega = 2\pi f$.

Hysteresis: Magnetization curve.

IACS: International Annealed Copper Standard. Conductivity as a percentage of oxygen free high purity (OFHC) copper. First developed by D. E. Hughes in 1880.

Inductance: Opposition of a circuit to the flow of an alternating current; the complex quotient of voltage divided by current.

Impedance plane diagram: Graphical presentation of the impedance and impedance changes of an eddy current probe as a function of changing frequency, conductivity, coil size, etc..

Impedance method: Eddy current test method which monitors the change in probe impedance, both magnitude and phase.

Inductive reactance: Reactance in a circuit that opposes any changes in current flow.

Inductor: Electronic device that resists changes in current flow, storing energy in the form of a magnetic field. Typically formed by a coil of conductive wire.

Internal probe (coil): Probe for testing tubes (or holes) from the inside. The coil(s) is/are circumferentially wound on a bobbin.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Liftoff: Spacing between the test coil and the specimen; indicates the degree of magnetic coupling between the specimen and the probe.

Liftoff effect: Effect of a change in magnetic coupling between the specimen and the test coil as the distance between them is varied.

Lissajous figure: Locus of the displacement of a point that undergoes the effect of two (or more) periodic motions. Generally, the two periodic motions are at right angles to each other and of the same frequency, giving rise to elliptical formations.

Magnetic coupling: See electromagnetic coupling; mutual inductance.

Magnetic field strength H (magnetic intensity): Measure of the magnetic field from the flow of electric currents; measured in A/m.

Magnetic flux density B (magnetic induction): Magnetization of a ferromagnetic material. $B = \mu_r \mu_0 H$.

Magnetic permeability: Ease with which a magnetic field can be induced in a solid.

Magnetic saturation: Limit of magnetization of a specimen, where an increase in the magnetizing force produces no increase in the magnetic flux density in the specimen.

Mutual inductance: Process by which current flowing in one circuit induces an electromotive force in a neighboring circuit; the magnetic field produced by the drive circuit induces a voltage in the pickup circuit proportional to the amount of magnetic field captured.

Noise: Any undesired signal, such as electrical noise or a signal from a specimen's dimensional or property variations, that obscures the signal of interest.

Null (offset) balance: Impedance plane eddy current technique where the initial operating point is chosen such that the response to certain parameters is minimized while enhancing the response to the parameter of interest.

Ohm's law: Electromotive force across a circuit is equal to the current flowing through the circuit multiplied by the total impedance of the circuit.

Operating point: Point on the impedance diagram that specifies the normalized inductive reactance and resistance of a coil.

Oscillator: Electronic unit in an eddy current instrument that generates alternating probe excitation current.

Parameter: Material property or instrument variable.

Permeability (magnetic): Ratio between flux density, B, and magnetizing force, H; describes the intrinsic willingness of a material to conduct magnetic flux lines.

Phase lag: Lag in phase (or time) between the sinusoidal currents flowing at the surface and those below the surface; in general, the lag in time between any two alternating variables.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Phasor: Vector describing sinusoidal signals; it has both amplitude and phase.

Primary field: Magnetic field surrounding the primary or drive coil due to the current flowing through it.

Probe: Eddy current transducer.

Reactance (X): Part of the impedance of an electric circuit given by the inductance of the circuit and the angular frequency ω of the alternating current plus the impedance caused by any capacitance in the circuit, where the inductive reactance is $X_L = \omega L$ and the capacitive reactance is $X_C = 1/\omega C$. The total reactance, measured in ohms, depends on whether the inductive or capacitive reactances are in parallel or series.

Reference coil: Coil that enables bridge balancing in absolute probes; its impedance is close to the test coil impedance, but does not couple to the test material.

Relative permeability (μ_r): Magnetic property of a medium, given by the ratio of the self-inductance of a circuit surrounded by the medium to the self-inductance of the circuit in free space.

Resistance (R): Ability of a material to resist the flow of an electric current; measured in ohms.

Resistivity: Resistance per unit length of the material of unit cross-section. Reciprocal of conductivity ($\rho = 1/\sigma$).

Resonance: Circuit having an inductor and capacitor connected in series or parallel such that the inductive reactance equals capacitive reactance; presents the minimum impedance to current flow (also called a *tuned* circuit).

Saturation (magnetic): Condition where incremental or reversible magnetic permeability of a ferromagnetic material becomes 1.0.

Secondary field: Magnetic field produced by induced eddy currents.

Signal: Response in eddy current instrumentation to excitation; has both amplitude and phase.

Signal-to-Noise Ratio: Ratio between defect signal amplitude and that from nonrelevant indications; minimum acceptable ratio is 3:1.

Skin depth: Effect that causes alternating currents to preferentially flow near the surface of a conductor; or, the depth of penetration of the alternating current. At one skin depth from the surface the current has drop by 36.8% and the phase has changed by one radian (see *depth of penetration*).

Skin effect: Phenomenon where currents and magnetic fields are restricted to the surface of a test sample; increasing test frequency reduces penetration (see *skin depth*).

Surface probe: Eddy current probe designed for testing surfaces; the coil is typically wound in a pancake shape.

Test coil: Eddy current probe.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Voltage: Electric potential or driving force for current.
Voltmeter: Instrument used to measure voltage.
Vector: Quantity having both amplitude (magnitude) and direction. Velocity is a vector, and speed is a magnitude.
Wheatstone Bridge: See *bridge circuit*.

VARIABLES

N	Number of turns in the eddy current probe
$\vec{B}$	Magnetic flux density
$\vec{H}$	Magnetic field strength
μ	Magnetic permeability
μ_0	Magnetic permeability in free space
μ_r	Relative magnetic permeability
μ_{rev}	Reversible magnetic permeability
ϕ	Magnetic flux
$\vec{E}$	Electric field
$\vec{D}$	Electric flux density
ε	Electric permittivity
L	Inductance
L_S	Self inductance
L_M	Mutual inductance
Z	Complex impedance
R	Resistance
X	Reactance
C	Capacitance
V	Voltage
ρ	Resistivity
σ	Conductivity
$\vec{J}$	Current density
δ	Skin depth
θ	Phase angle referenced to the phase at the surface
f	Linear frequency
ω	Angular frequency
k	Spatial frequency (wave number)
ν	Phase velocity
$\bar{r}$	Mean probe radius
$\bar{D}$	Mean probe diameter
$\bar{a}$	Mean tube or rod diameter
η	Fill factor
f_g	Limiting frequency
D	Tube or rod diameter

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

d Tube wall thickness
Q Quality factor

REFERENCES

1. JC Ørsted. Experiments on the Effect of a Current of Electricity on the Magnetic Needle. Ann Phil 16:273–276, 1820.
2. Hochschild. Electromagnetic Methods of Testing Materials. In: EG Stanford, ed. Progress in Non-destructive Testing Vol. 1. London: Heywook & Company Ltd., 1958, pp 59–109.
3. E Hughes. Philosophical Magazine. Series 5, Vol. 8 1879. pp 50–57.
4. F Förster. Theoretische und experimentelle Grundlagen der Zerstörungstreien Werkstoffprüpfung mit Wirbelstromverfahren. Zeitschrift fr Metallkunde, 43:163–171, 1952.
5. F Förster and Breitfeld. Nondestructive Test by an Electrical Method. Aluminum 25:14–15, 1943.
6. HL Libby. Multiparameter Eddy Current Concepts. In: RS Sharpe, ed. Research Techniques in Nondestructive Testing. London: Academic Press, 1970, pp. 345-382.
7. HL Libby. U.S. Patent No. 3,229,198. 1966.
8. HL Libby. Basic Principles and Techniques of Eddy Current Testing. Florida: Robert E. Krieger Publishing Company, 1979.
9. RC McMaster. Introduction to Electromagnetic Testing. In: RC McMaster, P McIntire, and ML Mester, eds. Nondestructive testing Handbook, Vol. 4, Electromagnetic Testing. American Society for Nondestructive Testing, 1986, p 16.
10. R Hochschild. The Theory of Eddy Current Testing in One (Not-So-Easy) Lesson. J Nondestructive Testing 12:31–40, 1954.
11. HL Libby. Introduction to Electromagnetic Nondestructive Test Methods. 2nd ed. Florida: Robert E. Krieger Publishing Co., 1971, pp 25–26.
12. CV Dodd. Computer Modeling for Eddy Current Testing. In: RS Sharpe, ed. Research Techniques in Nondestructive Testing. London: Academic Press, 1977, p 449.
13. VS Cecco, G Van Drunen, and FL Sharp. Eddy Current Manual. Vol. 1. Ontario: Chalk River Nuclear Laboratories, 1981, p 105.
14. SN Vernon and TAO Gross. Effects of shielding on properties of eddy current probes with ferrite cup cores. Proceedings of QNDE 6A:713–719, 1987.
15. PJ Shull, TE Capobianco, and JC Moulder. Design and characterization of uniform field eddy current probes. Proceedings of QNDE 6A:695–703, 1987.
16. JC Moulder, PJ Shull, and TE Capobianco. Uniform field eddy current probes: experiments and inversion for realistic flaws. Proceedings of QNDE 6A:601–610, 1987.
17. RC Grubinskas. An evaluation of probe coils with ferrite cores for use in electromagnetic testing. Materials Evaluation 24:557–563, 1966.
18. T Capobianco, JD Splett, and HK Iyer. Eddy Current Probe Sensitivity as a Function of Coil Constructions Parameters. Research in NDE 2:3:169–186.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

19. CV Dodds. The Use of Computer-modeling for Eddy-current Testing. In: RS Sharpe, ed. Research Techniques in Nondestructive Testing. London, Academic Press, 1977, pp 429–479.
20. B Vigness, JE Dinger, and R Gunn. Eddy current type flaw detectors for non-magnetic metals, J Appl. Phys., 13:377–383, 1942.
21. VS Cecco, G Van Drunen, and FL Sharp. Eddy Current Manual, Vol. 1. Ontario: Chalk River Nuclear Laboratories, 1981, pp 155–163.
22. RC McMaster. Nondestructive testing handbook. New York: The Ronald Press Company, 1959, pp 40–18–40–19.
23. G Birnbaum, G Free. Eddy-current characterization of materials and structures. Philadephia: American Society for Testing and Materials, 1981, pp 140–153.
24. J Blitz. Electrical and Magnetic Methods of Non-destructive Testing. 2nd ed. London: Chapman & Hall, 1997, pp 160–161.
25. VS Cecco, C Van Drunen, and FL Sharp. Eddy Current Test Manual. Chalk River, Ontario: Chalk River Nuclear Laboratories, 1983, p 66.
26. VS Cecco, C Van Drunen, and FL Sharp. Eddy Current Test Manual. Chalk River, Ontario: Chalk River Nuclear Laboratories, 1983, pp 155–163.
27. TR Schmidt. The Remote Field Eddy Current Inspection Technique. Materials Eval. 42:225–230 1984.
28. W Lord, YS Sun, SS Udpa, S Nath. Physics of Remote Field Eddy Current Effect. In: DO Thompson and DE Chimenti, eds. Review of Progress in Quantitative NDE. Vol. 7A, pp 165–172.
29. TR Schmidt, and DL Atherton. Understanding Remote Field Logs (Part 1). Materials Eval. 41:1276–1285, 1998.
30. TR Schmidt, and DL Atherton. Understanding Remote Field Logs (Part 2). Materials Eval. 42:1356–1362, 1998.
31. ER MacClean. U.S. Patent 2,573,799, November 6, 1951.
32. TR Schmidt. The Casing Inspection Tool: An Instrument for the In-Situ Detection of External Casing Corrosion in Oil Wells. Corrosion. 17:7, pp 81–85, 1961.
33. TR Schmidt. The Remote Field Eddy Current Inspection Technique. Materials Eval. 42:225–230, 1984.
34. TR Schmidt, DL Atherton, S Sullivan. Use of One-Dimensional Skin-Effect Equation for Predicting Remote-Field Characteristics, Including Wall-Thickness Verses Frequency Requirements. Materials Evaluation 47:76–79, 1989.
35. Anon. E2096-00 Standard Practice for In Situ Examination of Ferromagnetic Heat-Exchanger Tubes Using Remote Field Testing. In: Annual Book of ASTM Standards, 03.03. Philadelphia: American Society for Testing and Materials, 1990.
36. DD Mackintosh, L Cagle, D Russell. A New ASTM Standard for Remote Field Testing. Materials Evaluation 58:1213–1215, 2000.
37. Anon. Manual for Investigation and Correction of Feedwater Heater Failures, prepared for EPRI by Heat Exchanger Systems, Inc., and Encor-America, Inc. Special report, September 1991.
38. DD Mackintosh, DL Atherton, TR Schmidt, DE Russell. Remote Field Eddy Current for Examination of Ferromagnetic Tubes. Materials Evaluation, 54:652–657, 1996.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

39. Anon. E2096-31 Nondestructive Testing Definitions. In: Annual Book of ASTM Standards, 03.03:3. Philadelphia: American Society for Testing and Materials, 1999.
40. Anon. Nondestructive Testing Eddy Current. Jim Cox, ed. Harrisburg, NC: PH Diversified, 1997, pp B1–B8.
41. VS Cecco, C Van Drunen, and FL Sharp. Eddy Current Test Manual. Chalk River, Ontario: Chalk River Nuclear Laboratories, 1983, pp 187–193.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

6

Acoustic Emission

William H. Prosser
NASA Langley Research Center, Hampton, Virginia

6.1 INTRODUCTION

6.1.1 Technical Overview

Among the various NDE techniques, *acoustic emission* (AE) monitoring is arguably based on the simplest physical concepts, but is one of the most difficult techniques to practically implement. Nearly everyone has heard audible AE in the form of popping and cracking noises from materials under stress. Wood is an example of a material that produces such audible AE. As illustrated in the cartoon in Figure 6.1, a tree limb will emit cracking sounds when heavily stressed prior to breaking. One can estimate whether failure of the limb is imminent by listening to the amplitude and frequency of occurrence of these sounds.

A formal definition of *acoustic emission* is often given as "the release of transient elastic waves produced by a rapid redistribution of stress in a material." A key aspect of acoustic emission, in comparison to other NDE techniques, is that the signals are produced within the material itself. An external stimulus such as a change in load, pressure, strain, or temperature, is required to induce the source to emit acoustic waves. But the actual source is a change within the material that is usually permanent and irreversible, such as the growth of a crack. As such, AE monitoring is not strictly nondestructive, as some damage or change

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016



FIGURE 6.1 Analogy of AE of listening to audible cracking of wood.

in the material occurs. However, AE monitoring is included as an NDE technique because there is usually only a small, microscopic, amount of damage associated with a given AE event in comparison to the ultimate failure of the material or structure. In more common usage, the term acoustic emission refers to not only the release of sound waves as defined above, but also the acoustic waves themselves, as well as the NDE technique of monitoring these acoustic signals to assess damage.

Another analogy to AE is that of seismic monitoring, which is illustrated in Figure 6.2a. Seismic waves are elastic or acoustic waves that propagate through the earth. They are detected with a network of seismometers located around the world. Sources of seismic waves include the operation of fault mechanisms (earthquakes) as well as large explosions, such as underground nuclear testing. Analysis of seismic waves can provide the location and depth of a seismic source. In addition, earthquake and nuclear sources can be differentiated. Details of the structure of the earth, including the density depth profile and the depth of various layers such as the core and mantle have been determined from seismology. The similarities between seismology and AE monitoring, which is illustrated in Figure 6.2b, are so great that AE has also been called microseismology.

These analogies illustrate the simplicity of the basic concept of AE monitoring. However, there are many factors that make its practical implementation for NDE difficult. As discussed in the chapter on ultrasonics, acoustic wave propagation in solids is quite complicated. Multiple modes that propagate with different velocities, reflection, refraction, mode conversion, velocity dispersion, and attenuation, may all occur and affect the measured signal. This point can be illustrated with the use of a simulated AE source, a pencil lead fracture on the surface of a test structure as shown in Figure 6.3a. The pencil lead fracture, also

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

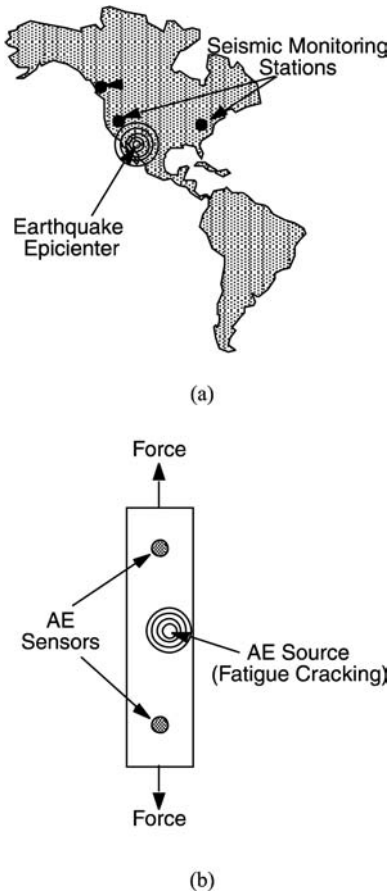
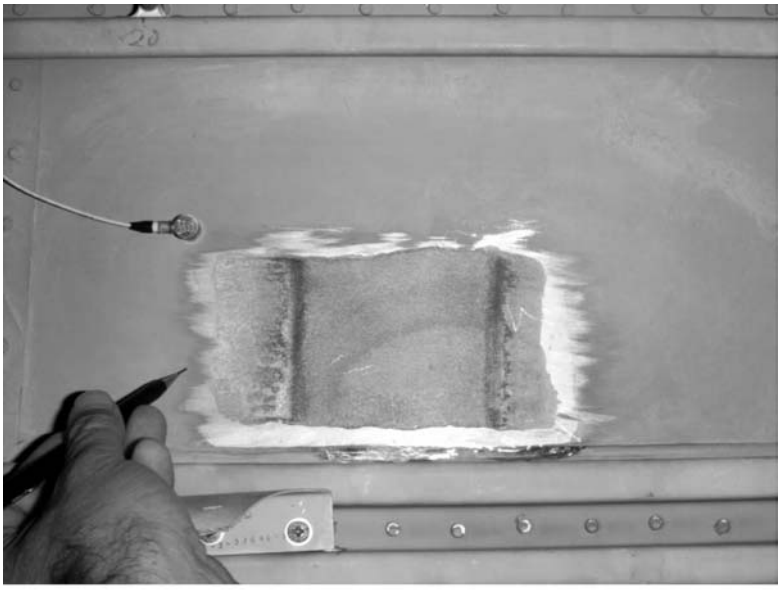


FIGURE 6.2 (a) Seismic analogy of AE and (b) illustration of AE test.

known as a Hsu-Nielsen source, is a standard method for simulating acoustic emission signals because it provides a fast rise time or step function like transient force that is very representative of real AE sources, and the resulting waveform is remarkably reproducible. The rather simple force function for a pencil lead fracture is shown on the top of Figure 6.3b. The resulting very complicated detected acoustic waveform after undergoing multiple reflections, mode conversions, etc. is shown on the bottom of Figure 6.3b.

In addition, AE testing is further complicated in comparison to ultrasonic testing by the fact that the characteristics of the "input" or emitted signal are relatively unknown. The emitted acoustic signal is dependent on the character-

Copyright © Marcel Dekker, Inc. All rights reserved.



(a)

FIGURE 6.3 (a) Photo of pencil lead break being performed on aircraft fuselage specimen along with AE sensor.

istics of the source, which is located inside the material. Different AE sources may produce vastly different AE waveforms. Even the same AE source mechanism may result in different received signals if the source is oriented differently with respect to the geometry of the medium or the propagation path to the detector.

In spite of these difficulties, AE monitoring has some unique advantages in comparison to other NDE techniques. First, by using an array of AE sensors, a global region or volume of material can be monitored. Time consuming and expensive point by point scanning is not required. Second, in many cases, AE can be used in situ to monitor a structure while it remains in service. The tremendous economic benefit and the potential for improved safety that can be derived from in situ, global, continuous monitoring provides a significant impetus to develop and apply AE techniques instead of more conventional NDE methods.

6.1.2 Historical Perspective

In its simplest form of listening to audible sounds from fractures, AE is probably the oldest NDE technique and has been around since the earliest history of man.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

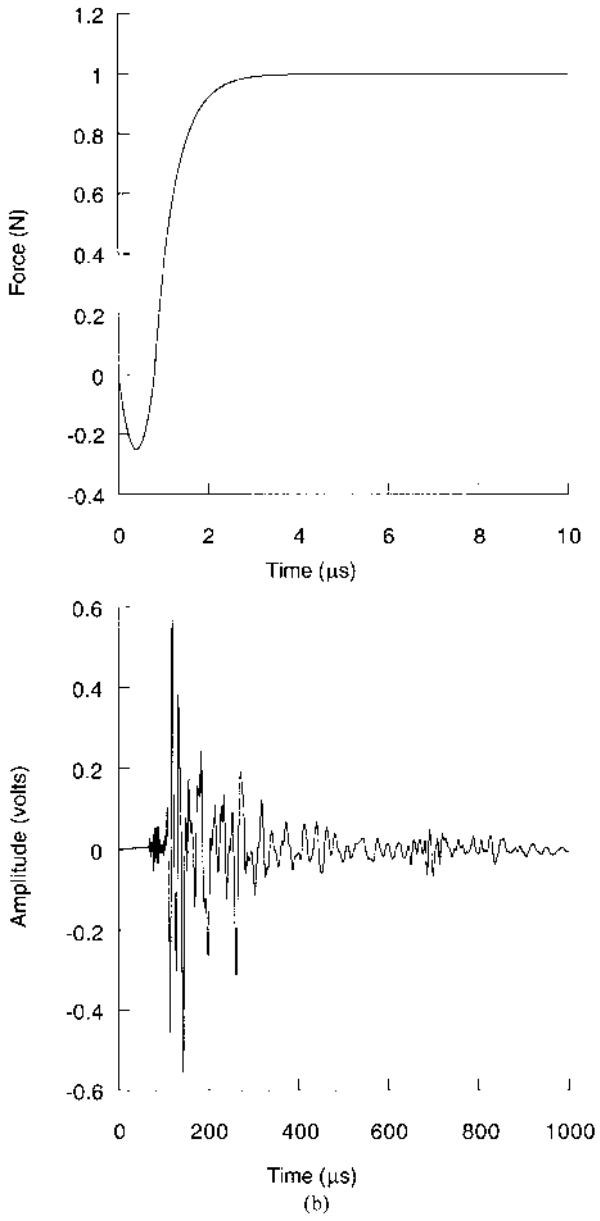


FIGURE 6.3 (b) Comparison of forcing function for pencil lead fracture simulated AE source (top) with complicated, received AE signal after propagation and detection (bottom).

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

As an example, the audible cracking sounds produced by clay vessels cooling in the kiln were used by early potters to assess the quality of the ceramic (1). Significant acoustic emissions were an indicator of poor quality. Later, audible acoustic emissions were noted in metals such as tin, which produces sounds, known as tin cry. when mechanically deformed. However, it was not until the 1930s to 1950s that electronic transducers were used to study AE. Early studies (2–4) monitored acoustic signals produced by deformation processes in metals. Kaiser (5) performed one of the most famous, early studies. His work demonstrated that AE is an irreversible phenomenon. Damage mechanisms such as cracks release acoustic waves as the load or stress is increased and they initiate and propagate. If the load is released, and then gradually reapplied, no additional AE will be generated until the load exceeds its *previous* maximum value. Above the previous maximum value, existing damage may propagate or new damage may initiate, both of which may produce AE. This irreversible phenomenon of AE has been designated the *Kaiser effect*. The Kaiser effect has been widely used in AE testing to determine information about previous maximum stress on a structure.

These early AE studies were followed by considerable research during the late 1950s through the 1970s. A published bibliography (6) of AE studies provides numerous references. Tatro and his students studied the basic physical process of the generation of acoustic emission. They also evaluated AE as a tool for studying deformation processes in materials. Their work is reviewed by Liptai et al. (7) and references therein. At the same time, there were numerous attempts to use AE to monitor the integrity of aerospace and nuclear structures. Some of these efforts are reviewed in a work edited by Miller and McIntire (1). The progress made in understanding AE phenomena and the practical applications developed during this time are amazing given the limited capabilities of instrumentation. Modern, high speed, signal digitizers were not available to record high frequency transient AE waveforms. Most measurements were made with RMS voltmeters or simple event counters.

Since the late 1970s, considerable advances have been made in waveform recording instrumentation. Transient recorders with multi-MHz digitization sampling frequencies necessary for high frequency AE signals were developed. The earliest of these high speed digitizers suffered from slow data transfer and storage rates allowing the acquisition of only a single or few events per second. In addition, the limited storage memory of computers prevented the acquisition of large numbers of events. Thus, these early transient recorders were impractical for typical AE applications. However, they were used in laboratory applications to study single AE events. A group at the National Bureau of Standards (now National institute of Standards and Technology) (8) carried out some of the earlier studies with this advanced waveform recording instrumentation. Their work provided considerable advances in the areas of AE sensor calibration, source

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

inversion, and theoretical modeling of AE. Models were adapted from seismology for the prediction and inversion of bulk mode AE signals.

More recent work has recognized the importance of studying acoustic waves in thin-walled structures in which AE signals propagate as guided modes rather than bulk modes. Many structures of practical testing interest are comprised of structural elements such as rods and thin walled plates, shells, and pipes. The dependence on how source orientation affects the generation of different guided modes in thin plates was demonstrated (9). Analysis of AE signal modal content thus provides a potential method for improved noise discrimination and source identification. Source location errors as a result of highly dispersive guided wave propagation were illustrated and an improved source location algorithm developed (10). Simultaneously, additional improvements in the capture and storage rate of digitizers have made full waveform capture and analysis possible for practical AE testing. In addition, increases in the speed and available memory of desktop computers have significantly improved analysis capabilities.

6.1.3 Potential of Technique

AE monitoring has been attempted for a wide variety of applications. Some of the more successful uses that have been developed include the evaluation of damage progression in laboratory materials studies, and the assessment of structural integrity of a wide variety of pressure vessels. Many types of pressure vessels used in petrochemical refineries and processing plants are routinely tested with AE. Trailer tank tubes containing multiple pressure vessels used for transportation of high pressure gases are also widely tested with AE. Some of the potential capabilities of AE that have led to these successful applications are discussed below.

Damage Activity Monitoring

The most basic AE test is to monitor acoustic "activity." For this simple approach, individual signals are not recorded or analyzed. Instead, an indication of the AE "activity" such as the number of signals, or the RMS voltage output from a sensor is recorded. The AE "activity" is then used as a qualitative indication that some damage or degradation has occurred. For this approach to be successful, it is critical that little or no background noise such as EMI or structural vibrations, exist. The Kaiser effect is an important analysis tool for AE "activity" monitoring. It provides a means for determining the previous maximum load or stress that has been applied to a structure of interest.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Damage Location

In addition to simply monitoring acoustic activity, an array of sensors can be used to locate damage. Differences in arrival times at the different sensor locations are combined with knowledge of the acoustic wave speed to triangulate the position of the source. Once the damage has been located, additional NDE techniques can be used to provide a further assessment of the severity of damage. This use of AE to locate damage can yield significant cost savings in comparison to fully scanning large structures with conventional NDE point sensors.

Damage Mechanism Identification

This is perhaps the most difficult AE analysis task, particularly when multiple damage mechanisms and/or noise sources are present. The difficulty in source identification arises from the significant changes that occur in AE signals as they propagate. In addition, the signals are further altered by sensors and measurement instrumentation. These changes due to propagation and measurement may mask characteristics that are related to the damage mechanism. However, source identification, when possible, significantly adds to the value of AE analysis. Source identification aids in noise discrimination and rejection, which improves the quality of AE data. Also, rather than simply detect and locate damage so that other NDE techniques can then be used, source identification makes it possible to begin to assess structural integrity based on the AE data alone. Numerous methods have been attempted to identify damage mechanisms from AE data and some of these will be discussed later. In general, however, detailed knowledge of the characteristics of the test structure including the geometry and material elastic properties, sensors and measurement system, and a good understanding of wave propagation effects are required to even attempt source identification. In addition, many carefully controlled laboratory experiments are also necessary to develop relationships between measured AE signals and specific damage mechanism.

Strength Predictions

The results from AE monitoring have also often been used in attempts to estimate the residual strength or life of a structure. However, there are many problems associated with such an analysis. First and foremost, there is no direct physical link between strength and acoustic emission, or for that matter, any other NDE phenomena. AE provides information about the existence and locations of defects within a structure. In theory, if such information were of sufficient accuracy and detail, it could be used as input in strength prediction models such as those based on fracture mechanics. However, AE provides little, if any, information about defect size and shape, which are critical parameters for strength prediction. Thus, in practice, most strength assessments from AE are based on empirical correla-

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

tions. Such correlations are developed from failure tests on a large number of nominally identical structures. A recent example of this approach is given in (11). One obvious limitation is the cost associated with generating a sufficient database. For very expensive components such as aerospace structures, building and testing to failure a large number of components is not feasible. Even for less expensive structures, the adequate simulation of numerous potential failure scenarios is nearly impossible. Yet another problem is that again AE signals can be significantly altered during propagation. Even small variations in the structural geometry, the material properties, and/or the relative locations of sensors and flaws, can invalidate previously developed AE/strength correlations.

6.2 FUNDAMENTALS

Measured AE waveforms are the result of the convolution of three effects; generation at the source, propagation, and measurement. These three fundamental aspects of AE are illustrated in Figure 6.4. This convolution can be expressed mathematically from theory developed for seismology (12). In this model (8), the AE sensor voltage ($V(t)$) is expressed as a function of the AE source mechanism, represented by a change in internal stress ($\Delta\sigma_{jk}$), a propagation term given by the spatial derivative of the Green's tensor ($G_{ij,k}$) and the transducer response (T). The Green's tensor is defined as the displacement components of an elastic medium of a given geometry at a specified receiver position, as a result of unit impulse forces along the coordinate axes at a given source position. Although the derivation of this theory is beyond the scope of this introductory text, the resulting equation is

$$V(t) = \int \int_{S_T} T(r, t - t')G_{ij,k'}(r, r_0', t' - t'')\Delta\sigma_{jk}(t'')drdt'' \tag{6.1}$$

where the source is positioned at r_0' within the interior of the material, the receiver is positioned at r on the surface of the material, and the transducer area is given by S_T. A number of assumptions are made in developing this model which include that the source is infinitesimal (i.e. small compared to the distance from the source to the receiver) and that the transducer does not disturb or load the surface. To successfully apply AE, a fundamental understanding of all three aspects, acoustic generation, propagation, and detection, and their respective effects on measured signals is helpful.

6.2.1 Sources

A number of source mechanisms have been postulated for acoustic emission. Some of these are listed in Table 6.1, along with some of the materials in which AE has been measured. The sizes of these sources range from the macroscopic

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

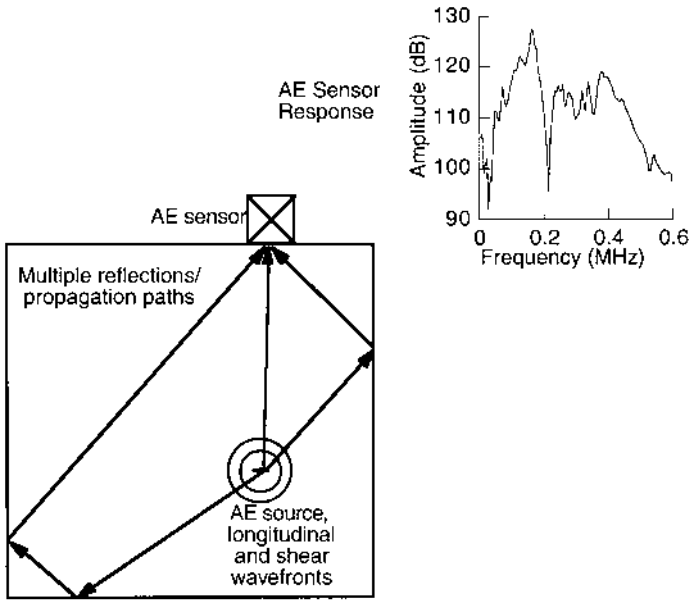


FIGURE 6.4 Illustration of how source, propagation, and sensor response affects AE signals.

such as earthquakes, to the microscopic as in dislocation movement. There are also other sources of acoustic waves that may not strictly meet the formal definition of acoustic emission given previously, but are of interest for structural monitoring. These include leakage of fluids and gases, projectile impact, and frictional noise from rotating bearings and machinery.

AE sources and the resulting signals are often characterized as being either discrete or continuous. Individual damage mechanisms such as an increment of crack growth or fracture produce discrete signals that are characterized by a clear beginning and end. Sources such as leakage may produce continuous signals that do not have a clear start or end. The signal is detected as long as the source remains active. In some cases, rapidly occurring multiple discrete sources can also produce continuous or near continuous AE. For most of the discussion in this chapter, discrete AE source mechanisms and signals will be considered.

In addition to AE produced by damage mechanisms and other sources of interest, there are also numerous sources of noise. Mechanical and structural vibrations from the operation of motors and hydraulic systems are an example. Rubbing or fretting of components in contact with the specimen being tested, or from mechanical fasteners, can also produce noise. In outdoor test settings, noise

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

TABLE 6.1 Materials in which AE Has Been Measured and Source Mechanisms Thought to Cause AE

Materials in which AE Has Been Measured
Metals
Ceramics
Polymers
Composites (including those with metal, ceramic, and polymer matrices and a wide variety of reinforcement materials)
Wood
Concrete
Rocks and geologic materials
Potential AE Sources
Microcrack sources such as intergranular cracking
Macrocrack sources such as fatigue crack growth
Slip and dislocation movement
Phase transformations
Fracture of inclusion particles
Fracture of reinforcement particles or fibers
Debonding of inclusions or reinforcements
Realignment of magnetic domains (Barkhausen effect)
Delamination in layered media
Rockbursts
Fault slip (Earthquakes)

can result from rain, hail, and windblown debris. Another noise source, although not acoustical in origination, is that of EMI. AE sensors, cables, amplifiers and instrumentation may all be susceptible to EMI. Noise creates a number of problems in AE testing and analysis such as making it more difficult to detect AE signals from sources of interest. In addition, measured noise signals that are not correctly identified and eliminated may lead to false indications of damage.

There are a number of important source characteristics that affect the acoustic waves that are generated. These include the time response of the change in internal stress, its magnitude and the area or volume over which it operates, and the orientation of the source. Unfortunately, however, such source characteristics are not well known or cataloged for different real AE source mechanisms. Accurate determination of source parameters is difficult, requiring complicated inversion and deconvolution to remove the effects of propagation and measurement. Additionally, even if source parameters could be easily measured, it is recognized that the characteristics of the same source mechanism can vary in

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

different materials, or under different loading conditions within the same material. However, there are some general relationships between source characteristics and the resulting acoustic waves that are helpful to understand and use in attempts at source identification.

Temporal Response

Acoustic emission is produced by a rapid change in the internal stress of a material. The time response over which this stress change occurs has a number of effects on the resulting signals. First, how quickly the stress changes, or the source rise time, affects the frequency content of AE waveforms. Neglecting changes in frequency content that are caused by sensor filtering, and by propagation effects such as attenuation, the bandwidth of AE signals is generally inversely proportional to the source rise time. Frequency contents range from the subaudible (20 Hz and below) as in earthquakes, to the audible (20 Hz–20 kHz) for cracking in wood, and into the MHz ultrasonic range for metals. In addition, the time response of the source has an effect on the measured amplitude. Rapidly operating damage mechanisms such as brittle crack propagation generate larger amplitude waveforms. Slower sources such as void formation and coalescence in ductile tearing of metals may not produce AE of sufficient amplitude to be measured. The time response of a few AE sources have been determined in laboratory studies of several materials including metals (13), glass (14), and concrete (15). In metals and glass, the rise time of crack sources is on the order of a microsecond or less. For comparison, the rise time of the pencil lead fracture, shown in Figure 6.3b, occurs on a similar time scale. In concrete, the measured rise time of AE sources is on the order of 100 microseconds, and for earthquake sources, rise times are much longer still.

Magnitude

Internal stress changes caused by damage mechanisms, and the associated internal displacements and changes in strain, release strain energy that was stored within the material. This strain energy is converted to other forms of energy including heat and acoustic waves. Assuming the same time response, the amplitude of AE waves are thought to be proportional to the magnitude of the strain energy released. However, there are several important things to note. First, the same source mechanism, such as crack growth, may release different amounts of strain energy in different materials and/or under different loading conditions. Also, as discussed above, if the time response is different, it may also have an effect on the signal amplitude. In addition, there are numerous other factors that have large impacts on measured signal amplitude such as attenuation during propagation. Thus, no simple correspondence between measured AE signal amplitudes and different source mechanisms should be expected as often has been proposed.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Orientation

As the tensor form of the source term in Eq. (6.1) suggests, AE sources have associated direction and orientation in addition to their time response and magnitude. This orientation of the source results in a radiation pattern in which the amplitude of the resulting AE wave is a function of the propagation direction. To this point, however, AE sources have been discussed as if only a single acoustic wave is emitted that is dependent on the time response, magnitude, and orientation of the source. In general, however, both longitudinal and shear modes are generated by a single AE source. The radiation patterns for both modes are dependent on the source orientation and are usually different.

For example, equations for the radiation patterns for longitudinal and shear mode can be derived for simple source models such as a monopole point force (12, 13). The force is assumed to be acting at the origin and along the z-axis with a time dependence of $F_Z(t)$, in an unbounded isotropic elastic medium with elastic properties λ and μ. For this case, the longitudinal wave displacement (u_l) as a function of the angle (θ) of propagation with respect to the z-axis is given by

$$u_l = [\cos(\theta)/4\pi(\lambda + 2\mu)r]F_z(t - r/v_l) \tag{6.2}$$

where v_l is the bulk longitudinal wave velocity. The radiation pattern for the shear wave displacement (u_s) is given by

$$u_s = [\sin(\theta)/4\pi\mu r]F_z(r - r/v_s) \tag{6.3}$$

where v_s is the bulk shear wave velocity. The resulting displacement radiation pattern plots for a fixed distance of propagation are shown in Figure 6.5 (13). For the longitudinal mode, the peak amplitude occurs at 0° and 180°, or along and opposite of the direction of the applied monopole force. For the shear wave, the amplitudes along these directions are zero, and its maximum occurs along propagation directions of 90° and 270°. For those more familiar with engineering elastic properties for isotropic materials, expressions relating the Lamé constants to the Young's modulus (E) and Poisson's ratio (v) are provided in Chapter 3 Eqs. (32, 33).

Example 1: For a constant distance of propagation (r) and force input (F_3), calculate the ratio of the peak longitudinal displacement (u_l) at 0 degrees to the peak shear displacement (u_s) at 90 degrees for aluminum which has elastic properties of $\lambda = 55.94$ GPa and $\mu = 25.95$ GPa.

$$\frac{u_l(\text{Peak})}{u_s(\text{Peak})} = \frac{[\cos 0^\circ/4\pi(\lambda + 2\mu)]}{[\sin 90^\circ/4\pi\mu]} = \frac{\mu}{(\lambda + 2\mu)}$$

$$\frac{u_l(\text{Peak})}{u_s(\text{Peak})} = \frac{(25.95 \times 10^9)}{((55.94 \times 10^9) + 2(25.95 \times 10^9))} = 0.24$$

The simple force monopole model is applicable to externally applied sources such as the pencil lead fracture AE simulation. However, real internal

Copyright © Marcel Dekker, Inc. All rights reserved.

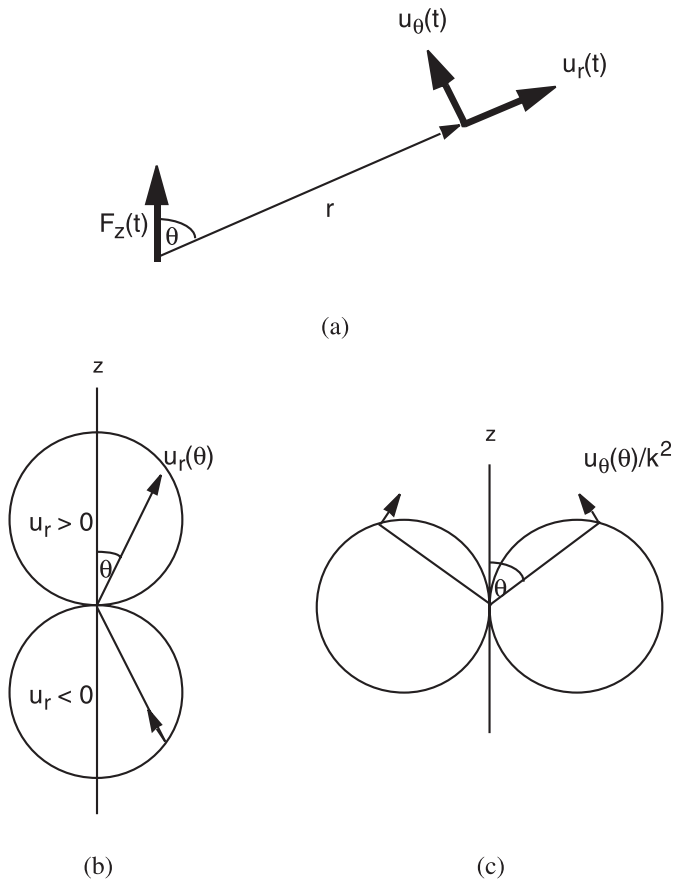


FIGURE 6.5 (a) Illustration of point force applied along z direction in an unbounded medium and the resulting radial and tangential displacements at an angle θ with respect to the z-axis, along with the angular dependence of the displacement amplitudes for (b) radial displacement caused by longitudinal wave and (c) tangential displacement caused by shear wave.

AE sources require combinations of force dipoles and double shear couples. Radiation patterns for such sources are even more complicated. Equations for radiation patterns from realistic sources have been derived. Experimental radiation pattern measurements from a crack source in glass were also made, which verified these equations (14). Radiation patterns were also measured for cracks in large concrete specimens demonstrating that crack orientation could be deter-

Copyright © Marcel Dekker, Inc. All rights reserved.

mined and that tensile and shear cracks could be differentiated from AE analysis (15). One complicating factor in attempting to use radiation pattern analysis for practical AE testing is the presence of reflections and mode conversions in real structures. The interference of these waves makes it more difficult to accurately measure the amplitudes of individual modes.

Source orientation effects have also been observed in thin plates and pipes (8, 16, 20). In these geometries, the bulk longitudinal and shear modes produced by the source undergo multiple reflections and mode conversions that interfere and result in propagating guided modes. For this more complicated case, analytical models relating source orientation to guided mode radiation patterns are not as readily available. Nevertheless, the dependence of the amplitudes of plate modes on source orientation for simulated AE sources was experimentally established. This work was extended to the more practical application of discriminating AE signals from crack growth from noise signals in tensile testing of thin metal and composite plates and coupons.

6.2.2 Wave Propagation

After acoustic waves are generated by a damage source, they must propagate through the structure to detectors located on the surface. There are numerous propagation effects that may alter the AE waves. These include multiple and dispersive propagation velocities, attenuation, reflection, mode conversion and guided mode propagation. These topics were discussed in detail in the chapter on ultrasonics. Thus, only a brief review of some of the more important effects on AE signals is again presented.

Propagation Velocity and Dispersion

Knowledge of the velocity of acoustic waves is very important for analysis of AE signals, especially so for location determination. In solids, the velocity of acoustic waves depends on the density and elastic properties of the material. For some propagation modes such as guided waves, the velocity may also be dependent on geometrical factors. Acoustic velocities may also be dispersive, or dependent on frequency, such as the case for guided modes or in viscoelastic materials. Additionally, the velocity is a function of propagation direction in anisotropic materials such as composites because of the dependence of elastic properties on direction. A brief review of velocity dispersion relationships for some of the more common AE modes of propagation is now provided.

Bulk Modes in Isotropic Materials. AE sources generate both longitudinal and shear acoustic waves, which will continue to propagate and be detected as AE signals in bulk or large specimens. In isotropic materials, the

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

velocity relations for the bulk longitudinal velocity (v_l) and the bulk shear velocity (v_s) are given in terms of the density and Lamé elastic constants by

$$v_l = \sqrt{\frac{\lambda + 2\mu}{\rho}} \tag{6.4}$$

and

$$v_s = \sqrt{\frac{\mu}{\rho}} \tag{6.5}$$

The longitudinal mode propagates with the fastest velocity and thus should be the first measured AE signal arrival if the source produced a sufficiently large longitudinal component. Neither mode is predicted to be dispersive, or have a velocity dependent on frequency. In addition to having different propagation velocities, the direction of particle displacement relative to the direction of propagation is different for these modes. For the longitudinal mode, the direction of particle displacement is the same as that of the direction of propagation. For the shear mode, the direction of particle displacement is transverse to the direction of propagation. Knowledge of particle displacement directions is of importance in understanding detected AE signals, as AE transducers may be sensitive to only certain directions of surface displacement.

Example 2: Calculate the longitudinal and shear velocities for aluminum, which has a density of 2700 $\mathrm{kg/m^3}$ and elastic properties of $\lambda = 55.94$ GPa and $\mu = 25.95$ GPa.

$$v_l = \sqrt{\frac{\lambda + 2\mu}{\rho}} = \sqrt{\frac{(55.94 + 2(25.95))\ \mathrm{GPa}}{2700\ \mathrm{kg/m^3}}}$$

$$= \sqrt{39.94 \times 10^6\ \frac{\mathrm{kg\,m/sec^2/m^2}}{\mathrm{kg/m^3}}} = 6320\ \mathrm{m/sec}$$

$$v_s = \sqrt{\frac{\mu}{\rho}} = \sqrt{\frac{25.95\ \mathrm{GPa}}{2700\ \mathrm{kg/m^3}}} = 3100\ \mathrm{m/sec}$$

Rayleigh (Surface) Waves in Isotropic Materials. Since AE transducers are located on the surface, the detected AE signal often also contains a Rayleigh or surface wave component. Calculation of the Rayleigh velocity (v_R) is more complicated than bulk mode velocities. In isotropic materials, v_R is calculated as the real solution to the following equation

$$\alpha^6 - 8\alpha^4 + (24 - 15\kappa^{-2})\alpha^2 + 16(\kappa^{-2} - 1) = 0 \tag{6.6}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

where

$$\alpha = \frac{v_R}{v_s} \quad \text{and} \quad \kappa = \frac{v_l}{v_s} \tag{6.7}$$

An approximate solution for the Rayleigh velocity is given by

$$v_R = v_s\left(\frac{0.87 + 1.12v}{1 + v}\right) \tag{6.8}$$

The particle displacement for Rayleigh waves is also more complicated. It is elliptical with components both along, and transverse to the direction of propagation. Rayleigh waves propagate along the surface, with an amplitude that decreases with depth from the surface.

Example 3: Calculate the approximate Rayleigh velocity for aluminum, which has a density of 2700 kg/m^3 and elastic properties of $\lambda = 55.94$ GPa and $\mu = 25.95$ GPa.

From Chapter 3 Eqs. (3.32, 3.33) the Poisson's ratio (v) can be expressed in terms of the Lamé elastic coefficients as

$$v = \frac{\lambda}{2(\lambda + \mu)}$$

Therefore,

$$v = \frac{55.94}{2(55.94 + 25.95)} = 0.3416$$

and

$$v_R = v_s\left(\frac{0.87 + 1.12(0.3416)}{1 + 0.3416}\right) = 0.9337v_s$$

$$v_R = 0.9337(3100 \text{ m/sec}) = 2894 \text{ m/sec}$$

For comparison, solving (9) exactly yields

$$v_R = 0.9209v_s = 2855 \text{ m/sec}$$

Lamb (Plate) Waves in Isotropic Materials. Many real structures of practical testing interest contain geometrical elements of plates, shells, pipes, beams, bars, or rods. In these geometries, bulk acoustic waves undergo multiple reflections and mode conversions at the surfaces. These multiple waves may superimpose to form propagating guided waves. The velocities of guided waves are dependent on not only the elastic properties, but also on geometrical parameters such as plate thickness or rod diameter. The velocities of guided modes are also often dispersive, or dependent on frequency. An additional complication of guided modes is that for a given geometry, there are multiple

Copyright © Marcel Dekker, Inc. All rights reserved.

modes that may propagate. In fact, theoretically there are an infinite number of modes. However, for a given bandwidth as contained in AE signals or captured by AE sensors, only a fixed number of modes are detected.

Similar to the procedure for bulk waves that was described in the ultrasonics chapter, the dispersion relations for guided modes are derived by substituting displacement wave solutions into the linear elastic equations of motion. However, the appropriate boundary conditions must be satisfied. The resulting guided mode dispersion equations are very complicated, typically requiring numerical solution techniques. An example is given by the Rayleigh–Lamb equations for guided modes in a plate that were presented previously in Chapter 3, Ultrasonics. A plot of the dispersion for the first three symmetric and antisymmetric Lamb modes in aluminum is shown in Figure 6.6. Note that in this figure the velocity is plotted against the product of the frequency and plate thickness as the dispersion is dependent on both.

As can be seen in Figure 6.6, for low values of the product of frequency and thickness, only the lowest order symmetric and antisymmetric modes may propagate. This is a case of much interest in practical AE testing as many structures can be approximated as thin plates. Examples include coupon specimens, many aerospace structures, and thin-walled pressure vessels and pipes.

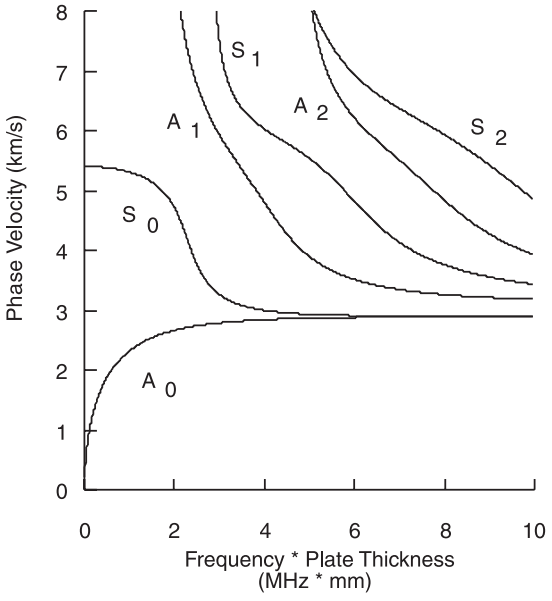


FIGURE 6.6 Dispersion for first three symmetric (S) and antisymmetric (A) Lamb modes.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Even in somewhat thicker plate materials in which higher order modes may be present, AE signals may be dominated by the two lowest order modes. Approximate plate theories have been developed that provide much more simple dispersion relations for these modes. For the lowest order symmetric mode, which is also called the extensional plate mode, approximate plate theory predicts that the velocity is not dispersive or dependent on plate thickness and is given by

$$v_e = \sqrt{\frac{E}{\rho(1-\nu^2)}} \tag{6.9}$$

For the lowest order antisymmetric mode, or flexural plate mode, the plate theory dispersion is

$$v_f = \sqrt{\omega\sqrt{\frac{D}{\rho h}}} \tag{6.10}$$

where h is the plate thickness and D is the plate bending stiffness given by

$$D = \frac{Eh^3}{12(1-\nu^2)} \tag{6.11}$$

Figure 6.7 shows a comparison of the exact Lamb mode dispersion with the plate theory predictions demonstrating good agreement for low values of thickness frequency product.

Example 4: (a) Calculate the plate theory extensional velocity for aluminum with density of 2700 kg/m^3, Poisson's ratio of 0.3416, and Young's modulus of 69.63 GPa.

$$v_e = \sqrt{\frac{69.63\ \text{GPa}}{2700\ \text{kg/m}^3(1-(0.3416)^2)}} = \sqrt{29.20\times10^6\ \text{m}^2/\text{sec}^2}$$
$$= 5404\ \text{m/sec}$$

(b) Calculate the plate theory flexural velocity for a 3.175 mm thick aluminum plate as a function of frequency and for specific frequency values of 10 kHz and 100 kHz.

$$v_f = \sqrt{\omega\sqrt{\frac{210.2\ \text{kg m}^2/\text{sec}^2}{2700\ \text{kg/m}^3\ 3.175\times10^{-3}\ \text{m}}}} = \sqrt{4.952\omega\ \text{m}^2/\text{sec}}$$

At 10 kHz, $\omega = 2\pi f = 6.283\times10^4\ \text{sec}^{-1}$

$v_f = \sqrt{4.952(6.283\times10^4)\ \text{m}^2/\text{sec}^2} = 557.8\ \text{m/sec}$

At 100 kHz, $\omega = 2\pi f = 6.283\times10^5\ \text{sec}^{-1}$

$v_f = \sqrt{4.952(6.283\times10^5)\,\text{m}^2/\text{sec}^2} = 1764\ \text{m/sec}$

Copyright © Marcel Dekker, Inc. All rights reserved.

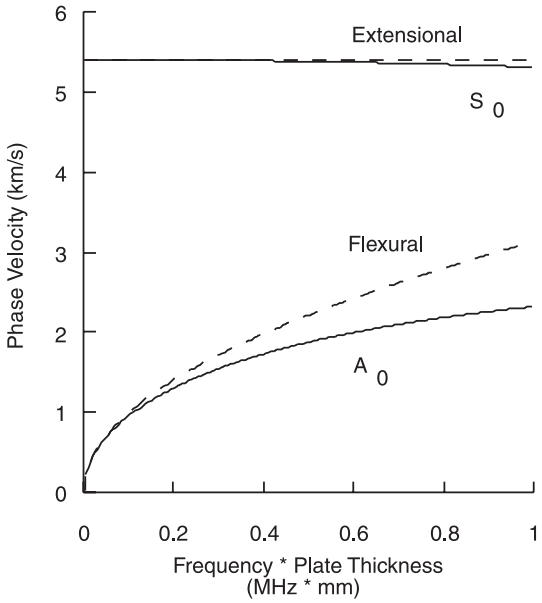


FIGURE 6.7 Comparison of Lamb mode dispersion (solid lines) and plate theory dispersion (dashed lines).

The particle displacements of guided modes are also more complicated. Similar to Rayleigh waves, the particle displacement for the extensional plate mode has components both along and perpendicular to the propagation direction. However, the largest component of particle displacement is along the propagation direction in the plane of the plate. For the flexural mode, the largest component of particle displacement is perpendicular to the propagation direction and normal to the plane of the plate.

Anisotropic Materials (Optional). Advanced composite materials, which are being increasingly used for structural applications, are one example of a class of anisotropic materials. The strength and elastic stiffness properties of anisotropic materials are a function of direction. This adds an additional complexity for AE and ultrasonic NDE methods in that the acoustic velocities also become dependent on direction.

To mathematically describe the directional dependence of stiffness in anisotropic materials, an increased number of independent elastic stiffness coefficients is required. The number of independent coefficients required depends on the symmetry of the material. Only two are necessary for isotropic materials,

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

which have symmetry about all axes and in all directions. Five are required for a unidirectionally reinforced fiber composite, which is assumed to be transversely isotropic. Such a material is isotropic in the plane perpendicular to the fiber direction in which the fibers are randomly distributed, but has different stiffness along and at angles to the fibers. For a composite laminate composed of multiple layers of unidirectional composite material, as many as nine independent elastic coefficients are necessary.

After the elastic coefficients and their interrelationships are determined from considerations of symmetry, equations can be derived for the acoustic velocities as a function of propagation direction. Such expressions are available for both bulk and guided acoustic modes. However, since many composite structures are comprised of thin laminates, only the example of guided plate wave propagation will be presented. Following Whitney (17) and Tsai and Hahn (18), a different set of elastic coefficients are used to describe the elastic behavior of thin laminated plates. The laminated plate stiffness coefficients relate the applied in-plane stresses (N_x, N_y, N_{xy}) and bending moment (M_x, M_y, M_{xy}) resultants to the midplane strains ($\varepsilon_x^0, \varepsilon_y^0, \varepsilon_{xy}^0$) and the plate curvatures ($\kappa_x, \kappa_y, \kappa_{xy}$). For a laminate with a thickness of h and orthogonal coordinate x- and y-axes in the plane of the plate and z perpendicular to the plane of the plate, the stress and bending moment resultants are defined by

$$(N_x, N_y, N_{xy}) = \int_{-h/2}^{h/2} (\sigma_x^{(k)}, \sigma_y^{(k)}, \sigma_{xy}^{(k)}) dz \tag{6.12}$$

$$(M_x, M_y, M_{xy}) = \int_{-h/2}^{h/2} (\sigma_x^{(k)}, \sigma_y^{(k)}, \sigma_{xy}^{(k)}) z dz \tag{6.13}$$

where the superscript k indicates the kth layer of the laminate. The midplane strains have the usual definition of strain but applied to the middle thickness of the laminate, while the curvatures are defined by

$$\kappa_x = -\frac{\partial^2 w}{\partial x^2} \qquad \kappa_y = -\frac{\partial^2 w}{\partial y^2} \qquad \kappa_{xy} = -2\frac{\partial^2 w}{\partial x\, \partial y} \tag{6.14}$$

where w is the displacement along the z direction.

To simplify the analysis, the case of a symmetric cross-ply, or "specially orthotropic" laminate is considered. Such a laminate has layers that are symmetrically stacked with respect to the midplane of the plate, and which contain fibers only along (0°) or perpendicular (90°) to the x-coordinate axis. For this simple laminate, there is no coupling between inplane and bending deformation and only

Copyright © Marcel Dekker, Inc. All rights reserved.

four inplane stiffnesses (A_{ij}) and four bending stiffnesses (D_{ij}) are necessary. The elastic behavior is expressed by

$$\begin{bmatrix} N_x \\ N_y \\ N_{xy} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & 0 \\ A_{12} & A_{22} & 0 \\ 0 & 0 & A_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_x^0 \\ \varepsilon_y^0 \\ \varepsilon_{xy}^0 \end{bmatrix}$$

and

$$\begin{bmatrix} M_x \\ M_y \\ M_{xy} \end{bmatrix} = \begin{bmatrix} D_{11} & D_{12} & 0 \\ D_{12} & D_{22} & 0 \\ 0 & 0 & D_{66} \end{bmatrix} \begin{bmatrix} \kappa_x \\ \kappa_y \\ \kappa_{xy} \end{bmatrix} \tag{6.15}$$

More complicated laminate stacking sequences require additional nonzero coefficients (A_{16}, A_{26}, D_{16}, and D_{26}).

To determine the extensional and flexural plate wave velocities as a function of propagation direction, equations for the transformation of the laminated stiffness coefficients for a rotation of coordinate axes are needed. A new set of orthogonal axes, x' and y' are defined such that they are still in the plane of the laminate, but with the x'-axis rotated by an angle θ from the original x-axis. The transformed inplane stiffnesses for the new coordinate axes are given by

$$\begin{aligned}
A'_{11} &= m^4A_{11} + n^4A_{22} + 2m^2n^2A_{12} + 4m^2n^2A_{66} \\
A'_{22} &= n^4A_{11} + m^4A_{22} + 2m^2n^2A_{12} + 4m^2n^2A_{66} \\
A'_{12} &= m^2n^2A_{11} + m^2n^2A_{22} + (m^4 + n^4)A_{12} - 4m^2n^2A_{66} \\
A'_{66} &= m^2n^2A_{11} + m^2n^2A_{22} - 2m^2n^2A_{12} + (m^2 - n^2)^2A_{66} \\
A'_{16} &= -m^3nA_{11} + mn^3A_{22} + (m^3n - mn^3)A_{12} + 2(m^3n - mn^3)A_{66} \\
A'_{26} &= -m^3nA_{11} + mn^3A_{22} + (mn^3 - m^3n)A_{12} + 2(mn^3 - m^3n)A_{66}
\end{aligned} \tag{6.16}$$

where $m = \cos\theta$ and $n = \sin\theta$. The bending stiffness coefficients are transformed in an identical fashion. It is noted that for the transformed coefficients, A'_{16}, A'_{26}, D'_{16}, and D'_{26} are now nonzero. The resulting dispersion relations for extensional and flexural plate waves are

$$v_e = \sqrt{\frac{A'_{11} + A'_{66} + \sqrt{(A'_{11} - A'_{66})^2 + 4A'^2_{16}}}{2\rho h}} \tag{6.17}$$

and

$$v_f = \sqrt{\omega\sqrt{\frac{D'_{11}}{\rho h}}} \tag{6.18}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Attenuation. Attenuation is the decrease in acoustic wave amplitude as a function of propagation distance. This has a number of detrimental effects in practical AE testing and analysis. Attenuation limits sensor spacing, which, in turn, limits the area that can be monitored by a fixed number of sensors. Attenuation may also affect source location accuracy. Decreases in signal amplitudes make it more difficult to accurately determine arrival times for sensors more distant from the source. Attenuation is also typically frequency dependent, which results in changes in not only the signal amplitude, but also the wave shape. This makes source identification analysis more difficult as well.

Attenuation, along with source differences, results in variations in AE signal amplitudes over orders of magnitude. Thus, a logarithmic amplitude scale is often used. The amplitude (A) of an AE signal is given in decibels (dB) by

$$A \text{ dB} = 20 \log \frac{V_{\text{sig}}}{V_{\text{ref}}} \tag{6.19}$$

where V_{sig} is the voltage of the measured signal and V_{ref} is an arbitrarily chosen reference voltage. A voltage of 1 μV at the sensor output is used in AE as the standard reference voltage. In addition to easily addressing the large range of AE signals, the logarithmic amplitude scale provides another advantage. The properties of the logarithmic function allow gain and attenuation values to be easily calculated by addition and subtraction. For example, a system with 40 dB of preamplifier gain and 20 dB of additional system gain provides a total gain of 60 dB.

The attenuation (α) is often reported in units of dB/unit length of propagation. There are a couple of things to note about attenuation parameters though. First, as will be discussed below, there are different attenuation mechanisms and not all of them result in a simple logarithmic amplitude loss. Also, although a single attenuation coefficient may be quoted for a given material, attenuation is dependent on a number of variables including the frequency of the wave, the particular wave mode, and for guided waves, geometric parameters such as plate thickness.

Example 5: (a) Assuming a preamplifier gain of 40 dB and a system gain of 20 dB, calculate the peak amplitude in dB (referenced to 1 μV at the sensor) for an AE signal whose peak voltage is 153 mV.

There are two ways in which this can be calculated. One is to calculate the peak amplitude of the signal in dB referenced to 1 μV at the sensor and then subtract the total gain as follows

$$A(\text{dB}) = \left(20 \log\left(\frac{153 \times 10^{-3}\ V}{1 \times 10^{-6}\ V}\right)\right) - 60 \text{ dB}$$
$$A = 43.7 \text{ dB}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

The second method is to calculate a new reference voltage after the total gain, and then calculate the amplitude of the signal referenced to this new value.

$$60 \text{ dB} = 20 \log\left(\frac{V_{\text{newref}}}{1 \times 10^{-6}\ V}\right)$$

$$V_{\text{newref}} = (1 \times 10^{-6}\ V)(10^3) = 1 \times 10^{-3}\ V$$

$$A = 20 \log\left(\frac{153 \times 10^{-3}\ V}{1 \times 10^{-3}\ V}\right) = 43.7 \text{ dB}$$

(b) For a material with an attenuation of 10 dB/m, calculate the distance of propagation for the peak voltage signal to be reduced by a factor of 2.

First calculate the logarithmic amplitude decrease of a factor of 2.

$$A_{\text{loss}} = 20 \log\left(\tfrac{1}{2}\right) = -6 \text{ dB}$$

The distance for a 6 dB loss is then given by

$$D = \frac{6 \text{ dB}}{10 \text{ dB/m}} = 0.6 \text{ m}$$

There are four attenuation mechanisms for acoustic waves. The first is that of absorption or internal friction, in which acoustic energy is converted to thermal energy within the material. This mechanism typically does result in a logarithmic loss with propagation distance. Absorption is also usually frequency dependent with higher losses occurring at higher frequencies. A second mechanism is that of geometric spreading. As the wave propagates outwards from the source, it disturbs an increasingly larger region of the material. For conservation of energy, the amplitude of the wave must correspondingly decrease. Thus, energy is not lost, but redistributed within the material. For bulk wave modes, the amplitude decreases in proportion to $1/r$ where r is the distance of propagation. For two-dimensional wave propagation such as in surface and plate waves, the amplitude decreases in proportion to $1/r^{1/2}$. As shown in Figure 6.8, geometric spreading is the dominant attenuation mechanism close to the source and causes significant amplitude decreases. In the far field, absorption may cause greater losses. In this figure, the amplitude versus distance is plotted for both absorption and geometric spreading. The attenuation due to either effect is the derivative, or instantaneous slope, of the respective curve at a given distance. The attenuation due to absorption in this figure is a constant 10 dB/m, while the attenuation (i.e., slope of the curve) due to geometric spreading is much higher close to the source, but less further from the source.

Example 6: Calculate the distance as a function of an absorption attenuation coefficient (α_A) such that attenuation due to geometric spreading (α_{gs}) and absorption are equal for bulk mode propagation.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

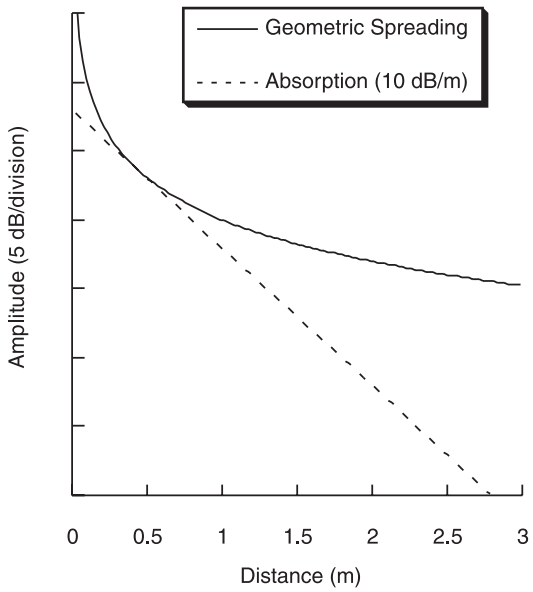


FIGURE 6.8 Comparison of amplitude as a function of distance for both geometric spreading in a plate (i.e. $1/r^{1/2}$ loss) and 10 dB/m absorption. The derivative or instantaneous slope of each line at any given distance is the attenuation for the respective loss mechanism.

For geometric spreading of bulk modes, the decrease in measured signal amplitude is proportional to $1/r$. Thus, the logarithmic amplitude is proportional to

$$A \text{ dB} \propto 20\log\left(\frac{\left(\frac{1}{r}\right)}{V_{\text{ref}}}\right) = ((20\log(1) - 20\log(r)) - 20\log(V_{\text{ref}})).$$

The attenuation from geometric spreading is the loss in amplitude with propagation distance, or the derivative of the previous expression

$$\begin{aligned}\alpha_{\text{gs}} &= -\frac{dA}{dr}(\text{dB/m}) = -\frac{d}{dr}((20\log(1) - 20\log(r)) - 20\log(V_{\text{ref}})) \\ &= -\frac{d}{dr}(-20\log(r)) = 20\frac{d}{dr}(\log(r)) \\ \alpha_{\text{gs}} &= 20\left(\frac{0.43429}{r}\right) = \frac{8.6858}{r}\end{aligned}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

If

$$\alpha_{gs} = \alpha_A$$

then

$$r = \frac{8.6858}{\alpha_A}$$

For distances closer to the source than this value of r, geometric spreading is dominant, while absorption is dominant for larger values of r.

A third attenuation mechanism is the dissipation of acoustic energy into adjacent media. Surface coatings and fluids contained within vessels can lead to losses of this type. Scattering of the signal from inhomogeneities such as grains in a metal or reinforcements in a composite can also cause this type of attenuation. Dispersion is a fourth attenuation mechanism. Frequency components superimposed at the source spread with propagation distance because of velocity variations. This results in a reduction in the peak amplitude. This type of attenuation can be quite large for highly dispersive modes such as guided waves. For example, the far field attenuation for the flexural plate mode in a composite plate has been measured to he greater than 80 dB/m (19).

Reflection, Refraction, and Mode Conversion

When an acoustic wave reaches a boundary between two media with different elastic properties, reflection, refraction, and mode conversion may occur. The reflection of acoustic waves is analogous to the reflection of electromagnetic waves such as light from a mirrored surface. The angle of reflection is equal to the angle of incidence. Refraction of acoustic waves is also similar to the refraction of light such as when propagating from air into water. The refracted wave has an angle of transmission that is different from the angle of incidence. The angles are related through the ratio of the acoustic velocities as given by Snell's law

$$\sin\theta_1 = \frac{v_1}{v_2}\sin\theta_2 \tag{6.20}$$

where v_1 and v_2 are the velocities of the acoustic waves in the two media and θ_1 and θ_2 are the angles of incidence and transmission. This applies for both incident and transmitted longitudinal waves, and incident and transmitted shear waves. Acoustic waves, however, differ from electromagnetic waves in that a mode converted wave component may be produced at a boundary. An incident longitudinal wave may produce not only reflected and refracted longitudinal waves, but also reflected and refracted shear waves with vertical polarization. Likewise, vertically polarized shear waves may produce reflected and refracted longitudinal waves in addition to the expected vertically polarized shear reflected and refracted waves. A horizontally polarized shear wave does not produce mode converted

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

longitudinal waves. Again, Snell's law relationships are used to calculate the relative angles of mode converted, reflected and refracted wave components. A much more detailed analysis of these phenomena including critical angle effects and the amplitude relationships between incident, reflected, refracted, and mode converted components was presented previously in the ultrasonics chapter and is not presented again here. Instead, the effects that reflection phenomena may have on AE signals and their analysis is discussed.

When performing an AE test in structures of finite geometry, the resulting detected signal will be a superposition of the directly propagating signal from the source, along with all reflections and mode conversions that propagate to the sensor. Depending on the size and complexity of the geometry, the superposition of these signals can produce a resultant waveform that is significantly different from that caused by the original source. This is illustrated in Figure 6.9, in which simulated AE signals caused by pencil lead fractures were detected after the same propagation distance in two specimens with different geometry. One specimen was a large plate of graphite epoxy composite material with dimensions of 38.1×50.8 cm. The other specimen was a small tensile test coupon of the same material of dimensions 2.54×30.48 cm. The signal in the small coupon contained numerous reflected signal components. As shown in this figure, AE signals containing reflections typically have larger peak amplitudes. Energy that would have otherwise radiated away is reflected and focused toward the receiving transducer. In addition to changes in the peak amplitude, the superposition of reflections and mode conversions will also change the overall shape of an AE signal. These changes in amplitude and wave shape from reflections and mode conversions complicate attempts to perform source analysis and/or noise discrimination. A given source mechanism may produce drastically different AE signals in the same specimen when located at different positions relative to specimen boundaries, or in specimens with different geometries.

6.2.3 Measurement

After generation and propagation, an AE signal must be detected, conditioned, and recorded for analysis. There are a wide variety of AE sensors and measurement systems available for this purpose. Some of the more commonly used detection and recording instrumentation will now be discussed.

Transducers

The first step in the AE measurement process is the conversion, or transduction, of the mechanical energy of the elastic wave into electrical energy. An ideal transducer would provide a voltage output that was directly proportional to the surface displacement, velocity, or acceleration at the measurement point. This ideal sensor would operate over the entire frequency spectrum, without introdu-

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

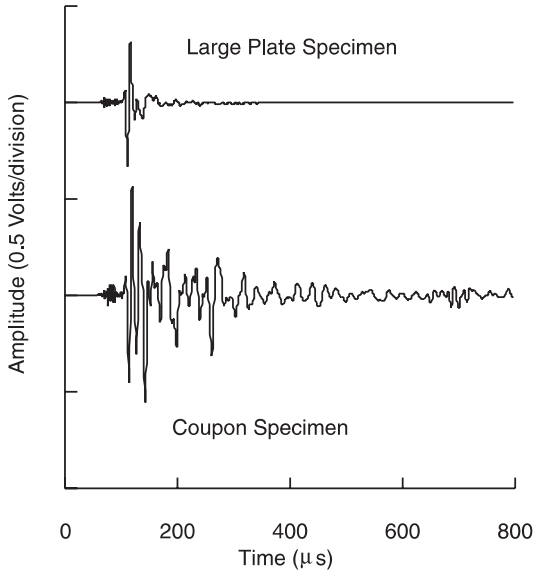


FIGURE 6.9 Comparison of simulated AE signals in a large composite plate specimen (38.1 × 50.8 cm) and in a small coupon specimen (2.54 × 30.48 cm).

cing any noise to the signal, and without disturbing the motion of the elastic wave. In reality, there is no such device. All real transducers have a limited bandwidth, over which the sensitivity may vary. Real sensors may also respond to different variables (i.e., displacement, velocity, or acceleration) or even a combination at different regions of the frequency spectrum. In addition, real transducers typically have a finite area over which the received signal is integrated. Since they usually contact the surface, real sensors may also disturb or alter the motion of the wave being detected. Furthermore, real sensors may introduce noise into the measured signal. Thus, selecting an AE transducer, like selecting any other physical measurement device, requires that compromises and tradeoffs be considered to determine the best suitable transducer for the measurement situation.

As discussed in the ultrasonics chapter, there are a variety of devices that can be used as acoustic wave transducers. These include strain gages, accelerometers, piezoelectric sensors, electromagnetic acoustic transducers (EMATs), and optical or fiber-optic interferometers. All of these devices have different characteristics such as sensitivity, signal/noise ratio, dynamic range, and frequency bandwidth. Several features of AE signals play a role in the selection of the type of device used. First, the detection of very small amplitude signals is

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

necessary. Surface displacements from AE signals rarely exceed 1 Å (10^{-10} m) and are typically even smaller than those detected in conventional ultrasonic testing. Thus, high sensitivity and signal/noise ratio is required in an AE sensor. In addition, a wide dynamic range is required as the amplitude of AE signals may vary by more than 100 dB. Also, because of the large, low frequency vibration background in many testing environments, AE monitoring is usually done in the ultrasonic frequency range (> 20 kHz). Piezoelectric sensors provide the best sensitivity, signal/noise, and dynamic range over this frequency range and are used almost exclusively in AE testing.

There are a wide variety of piezoelectric AE transducers available, several of which are shown in Figure 6.10. Also shown in this photo is a piezoelectric element, which is the basic sensing component from which these transducers are made. With respect to their frequency response, these can generally be broken down into two categories: resonant and high fidelity. Resonant sensors are designed with little damping of the natural resonant response of the piezoelectric element. This provides maximum sensitivity at or near this resonance frequency. Their response, however, usually falls off significantly at frequencies other than the resonance. Often, resonant sensors are called narrowband. However, this is not entirely correct as resonant sensors may have multiple resonances and have high sensitivity peaks over a wide frequency range. Figure 6.11a illustrates this characteristic showing the frequency response of a typical resonant AE sensor (20). This sensor has a primary resonance peak near 150 kHz, but also has multiple resonant peaks over a broad bandwidth. Sensors with a resonant frequency in the range of 150 kHz are among the most widely used in AE testing, and they are typically monitored over the 100–300 kHz bandwidth. This frequency bandwidth is considerably higher than that of noise associated with structural vibrations. However, it is low enough to minimize problems with

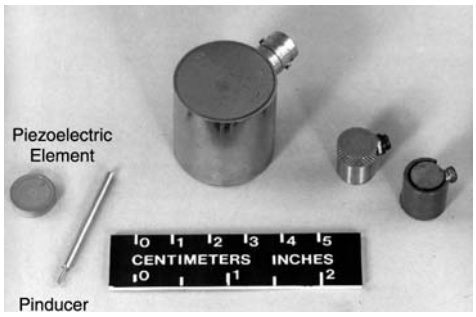


FIGURE 6.10 Photo of a variety of commercial AE transducers along with raw sensing piezoelectric element.

Copyright © Marcel Dekker, Inc. All rights reserved.

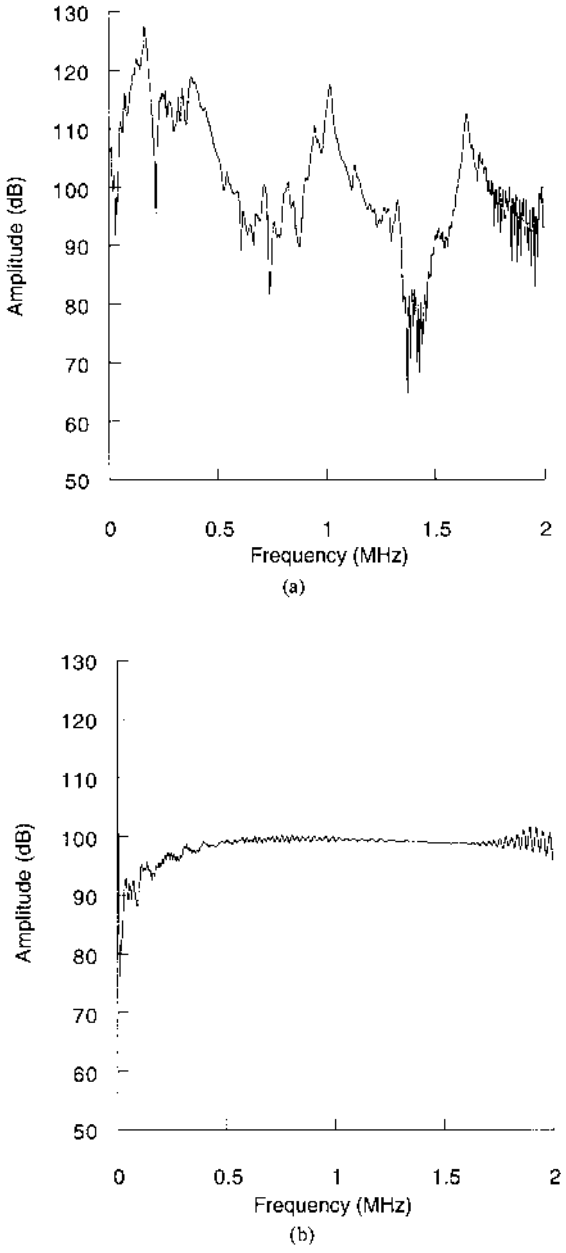


FIGURE 6.11 Transducer frequency response curves for (a) 150 kHz resonant sensor and (b) wide band, high fidelity AE sensor.

Copyright © Marcel Dekker, Inc. All rights reserved.

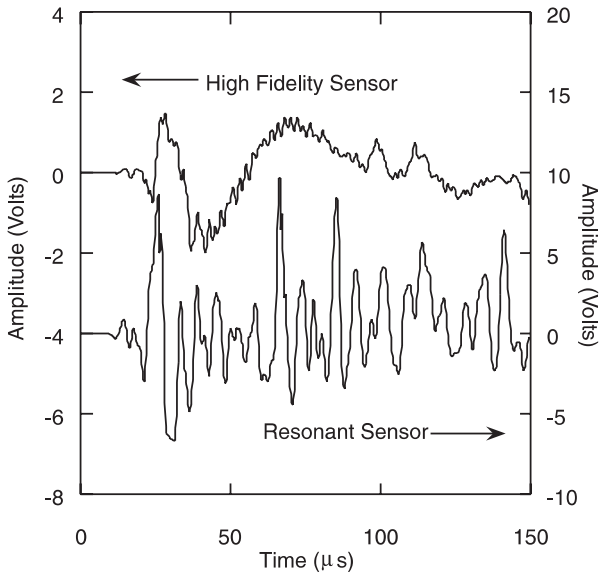


FIGURE 6.12 Comparison of high fidelity and resonant sensor outputs for simulated AE signal in a thin aluminum plate. Resonant sensor signal is highly distorted but much larger in amplitude.

electromagnetic interference as well as the increased attenuation that occurs at higher frequencies.

High fidelity sensors provide an output that is flat with frequency and thus more closely approximates the ideal sensor. High fidelity sensors have been manufactured by either heavily damping a resonant piezoelectric element, or by using a unique conical piezoelectric element with a large backing mass (21). The frequency response of a high fidelity AE sensor is shown in Figure 6.11b for comparison (20). Typically, high fidelity sensors have lower peak sensitivity than resonant sensors, but more accurately reproduce the actual wave displacement. Although resonant sensors provide the highest peak sensitivity, their complicated frequency response can significantly alter AE signals. This is illustrated in Figure 6.12 which compares simulated AE signals in an aluminum plate when detected by the high fidelity and resonant sensors whose frequency responses were shown in Figure 6.11. For some AE testing situations, the changes in the signal caused by a resonant sensor are unimportant, or may be outweighed by the need for increased sensitivity. However, for accurate waveform analysis to determine source mechanisms, high fidelity transducers are necessary.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Signal Conditioning

Even with the high sensitivity and signal/noise ratio of piezoelectric acoustic emission transducers, their voltage output is usually quite small. Typical values range from microvolt to the millivolt levels. Thus, signal conditioning at or near the sensor is necessary to allow transmission to and detection by the measurement system. For this purpose, low noise preamplifiers are usually placed at a cable distance of no more than 1.2 m from the sensors. Several examples of AE preamps are shown in the photo of Figure 6.13a, including a rugged version flight qualified for use on the X-33 prototype Reusable Launch Vehicle. Commonly used preamps provide gain values over the range of 20–60 dB. Sensors with preamps integrated into the sensor case, one of which is also shown in Figure 6.13a, are also available. Such integrated preamp sensors allow easier testing because of the fewer cable connections that are required. However, there are also several drawbacks. First, the electronic components of the preamp may not be as environmentally rugged as the sensor. The applicable temperature range or vibration specifications for an integrated sensor may be less than that of

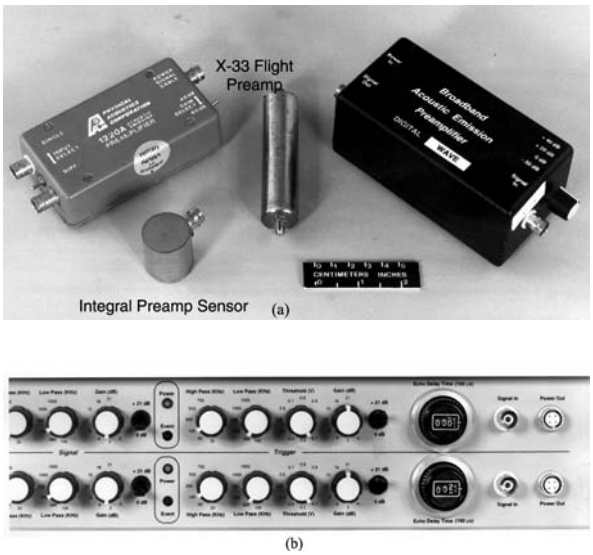


FIGURE 6.13 Photos of (a) AE preamplifiers including a sensor with an integral preamp and a rugged preamp flight qualified for the X-33 prototype Reusable Launch Vehicle and (b) a signal conditioning module used to provide additional gain and filtering for AE signals and to set acquisition trigger parameters.

Copyright © Marcel Dekker, Inc. All rights reserved.

the sensor alone. Also, if either the sensing element or integrated preamplifier fails in an integrated sensor, the whole device is lost as they are not usually economically feasible to repair. The last limitation is that a separate preamp is usually made with variable gain settings. To minimize size, preamps in an integrated sensor most commonly have fixed gain values.

To further condition the signal and improve signal/noise, analog filters are used at the preamps. For the widely used 150 kHz resonant sensor, the most common frequency range for bandpass filtering is 100–300 kHz. When such narrow frequency filtering is used with a resonant sensor, it does truly become a narrowband measurement system as the frequency content of the signal detected at other frequency resonant peaks is reduced or eliminated by the filtering. For high fidelity sensors, bandpass filtering of either 20 or 50 kHz to 1.2 MHz is often used. In addition to conditioning at the preamp, additional filtering and gain may also be applied at the measurement instrument. Figure 6.13b shows a photo of a two-channel signal conditioning module that provides additional filtering and gain and is also used to set acquisition trigger parameters.

Measurement Systems

Again considering the ideal case, an AE measurement system should allow continuous recording of signals from a number of transducers for the duration of a test, with a bandwidth and dynamic range exceeding that of the sensors, while adding no additional electronic noise. Again, as with sensors, no such ideal measurement system exists. In fact, it has been only since the development of *analog-to-digital* (A/D) converters with sufficient sampling rates in the 1980s that recording of large numbers of AE signals has been practical. And even with these A/D systems, storage memory and data bus transfer limitations prevent absolutely continuous monitoring. It is with this historical background that some of the more common measurement approaches are discussed.

"Activity" Monitoring. Because of measurement instrumentation limitations, many early AE systems only recorded that an acoustic event had occurred, along with perhaps an indication of its magnitude. A number of devices were developed to provide such measures of AE "activity." One of the most simple was to input the signal from the transducer into an AC voltage measurement device such as a RMS (root-mean-squared) voltmeter. The output from the RMS voltmeter provides evidence of the occurrence of AE and its magnitude, and can either be plotted during a test or recorded for analysis. Another simple "activity" monitoring approach was to use an electronic counter to count the number of events that exceeded a preset voltage threshold. A preset timeout is used to avoid counting multiple threshold crossings from the same event. Such simple measurement systems provide a record of the "activity" or level of AE, which is then correlated against other test parameters such as load, stress, strain, temperature,

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

and any observed damage development. These measurement approaches continue to be used today. They are especially useful when source identification and location are unimportant or not possible. "Activity" monitoring is often used for continuous or near continuous AE signals such as those from leak sources.

Feature Measurement. The next development in AE measurement systems was the capability to measure features or parameters of individual discreet AE signals. Features that typically have been measured include the arrival time, peak amplitude, duration, rise time, threshold crossings or counts, energy, and average frequency content. Some of these features are illustrated in Figure 6.14. As illustrated, signal detection and feature measurement are again based on a preset voltage threshold or trigger level. The arrival time is determined from the first signal excursion above the threshold voltage. The duration is the difference in time between the first and last threshold crossings. The rise time is the time from the first threshold crossing until the peak amplitude. In theory, the energy is an integrated measurement of the square of the voltage signal with respect to time. However, in most AE measurement systems, the measured energy value is only loosely based on this definition.

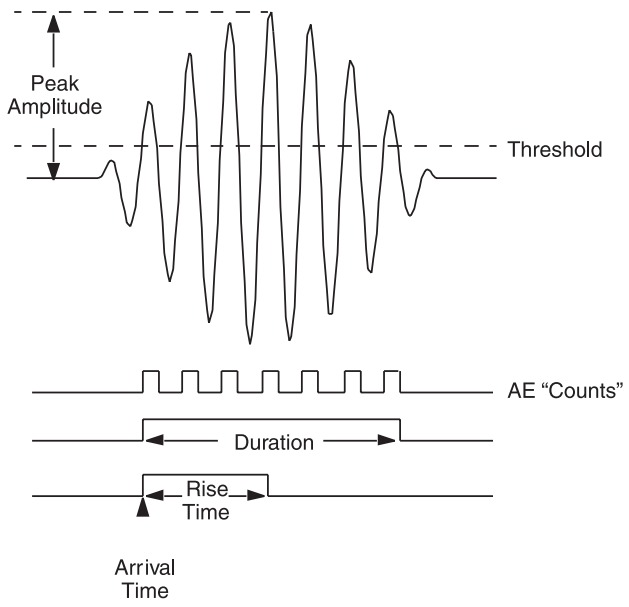


FIGURE 6.14 Illustration of AE signal features on simplified model AE signal.

Copyright © Marcel Dekker, Inc. All rights reserved.

There are difficulties in measuring the end, or last threshold crossing on an AE signal that should also be mentioned. First, AE signals may vary in duration from microseconds to milliseconds or even longer, and the amplitude of an AE signal may vary in a very complicated manner as a function of time due to arrivals of multiple modes and/or reflections. Furthermore, AE signals are asynchronous, or not equally spaced in time. This is because events causing AE signals may occur at widely varying intervals depending on loading conditions and the localized strain conditions at the source of emission. Thus, the trick is to determine the appropriate end of the signal without counting additional modes or reflections as separate signals, while at the same time not including later signals with the previous signal. The usual approach to this problem is to use a preset window time, known as the hit lockout time. For the determination of new signals, additional threshold crossings after the initial threshold crossing are ignored if they occur within the hit lockout time. This hit lockout time is a moving window and is reset for each new threshold crossing, until the signal decreases in amplitude such that no threshold crossings are detected for the duration of the lockout time. At this point, the signal is considered to have ended and the next threshold crossing is taken as the start of a new signal. This process is illustrated in Figure 6.15. Experimentation with a particular material and transducer, and gain and filter settings, is used to determine appropriate threshold voltage and lockout times.

One advantage of feature based systems is that they provide high event rates which are necessary for some AE testing. Very fast analog circuitry has been developed for AE signal counting and feature determination. The high acquisition rates and computer processing and storage can provide near real-time results. This is important in some testing for preventing catastrophic failure of an expensive specimen. Another advantage of feature measurement systems is that the circuits can measure signals over a wide dynamic range, which is often quoted to be as high as 80–100 dB. However, the real dynamic range may be less because of the smaller dynamic range of signal conditioning elements including preamplifiers and amplifiers.

Feature based AE measurement systems are still widely used and there are numerous approaches for analyzing such data. Approximate source locations can be obtained from the triangulation of arrival times for a signal detected by multiple sensors at different locations. Attempts can also made to correlate one or more signal features with different source mechanisms. Some of the these approaches will be discussed later in the techniques section.

Waveform Digitization. High speed analog-to-digital (A/D) converters have been combined with low cost, high speed personal computers to provide AE systems that can record AE waveforms. In capturing digitized recordings of AE signals, these systems more closely resemble the ideal AE measurement system

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

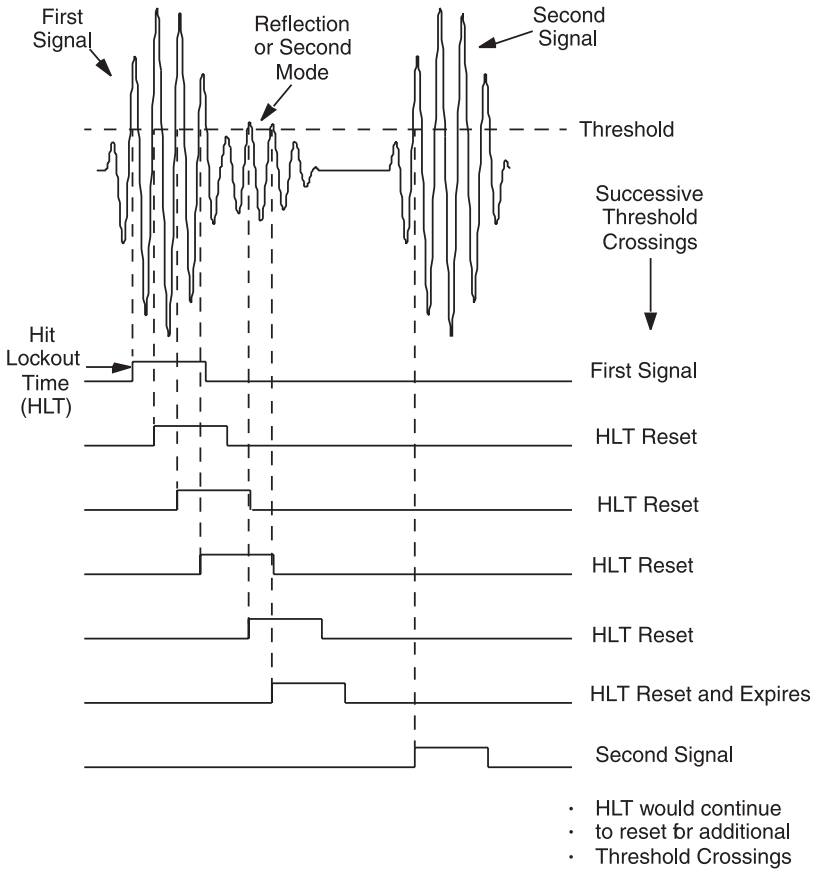


FIGURE 6.15 Illustration of Hit Lockout Time on simplified model of two AE signals.

than feature measurement systems. Analysis of the waveforms can provide additional information about the AE source in comparison to feature set data. More accurate arrival time determination can be made, which results in improved source location. Waveform analysis can also lead to enhanced noise discrimination and superior source characterization. Another advantage is that once the waveforms are recorded, they can be analyzed multiple times with different analysis techniques.

One current disadvantage of waveform digitizing AE systems is that the ideal capability of continuous recording is not possible. This is because of computer storage memory limitations. Instead, a segmented capture approach is

Copyright © Marcel Dekker, Inc. All rights reserved.

used. Similar to feature based AE systems, a trigger is generated when the AE signal exceeds a preset threshold. The trigger initiates the acquisition of a preset number of digitized points, or equivalently, period of time. A circular memory buffer system is also used to allow pretrigger signal to be recorded. Pretrigger information is of importance in more accurately determining the arrival time of an AE signal. Often, the first arriving directly propagating signals are of smaller amplitude, especially at sensors more remote from the source. These small amplitude signal components may not exceed the threshold. Of course, more accurate arrival time determination provides more accurate source location.

A disadvantage of recording only a segment or preset number of points is that the end portion of signals with very long durations may be missed. There is a tradeoff that must be considered in extending the number of points recorded for each signal versus how much storage memory is available. Fortunately, the most important information in AE signals, such as the arrival time and amplitude of directly propagating modes, is usually contained in the early time period. The latter part of most AE signals is dominated by reflections. These reflections may provide much information about the geometry of the test specimen, but little about the source mechanisms.

In addition to limited storage media, limited computer bus speeds play a role in preventing continuous recording of signals from AE transducers, and in limiting the rate of signals that can be acquired when threshold triggering and finite window sampling is used. This has probably been one of the biggest issues in transitioning AE technology from feature-based systems to waveform based systems. Feature based systems can measure and record data at rates of thousands of events/second which, for example, may occur in testing composite materials. Current waveform based systems may handle only hundreds of events/second. However, advances in low cost computer storage memory and improved bus speeds are anticipated that will resolve this problem. Other major issues in transitioning from feature-based systems to waveform based systems include the cost to replace existing instruments, the required transition of existing standards, codes, and practices, and the lack of automated analysis software for waveform based systems.

For waveform digitizing AE systems, the sampling frequency and vertical resolution are two parameters of importance. A sampling frequency of at least twice the highest frequency component of the signal is required to meet the Nyquist criteria to avoid aliasing. Even higher sampling frequencies are desirable to more accurately record the signals, and to provide improved time resolution for source location analysis. For AE waveforms detected with a high fidelity sensor, the frequency content may exceed 1 MHz. Sampling frequencies from 5 MHz to as high as 30 MHz are commonly used.

Vertical resolution is determined from the number of voltage levels digitized by the A/D device. It is often quoted as a number of measurement

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

bits (N) where the number of voltage levels is 2^N. The theoretical maximum *dynamic range* (DR) of an A/D device is given by

$$DR \text{ dB} = 20 \log(2^N) \tag{6.21}$$

In practice, the number of effective bits, and thus real dynamic range of a digitizer is somewhat less than the maximum specified. Early A/D systems were limited to 8 bits, which provided a theoretical maximum dynamic range of only 48 dB. This may be inadequate for practical AE testing. More recent A/D converters provide 12 bit or higher vertical resolution, offering 72 dB or higher theoretical maximum dynamic range, at sufficiently high sampling frequencies for AE measurement.

6.3 ANALYSIS TECHNIQUES

There have been numerous techniques developed to interpret AE data. The most appropriate approach for any given AE testing situation is dependent on a number of factors. These include the specimen and loading conditions, acoustic and EMI noise environment, type and number of sensors used, type of measurement system, applicable standards and codes, and the goals or purposes for AE monitoring. Some of the more commonly used analysis methods for data obtained with "activity," feature, and waveform AE systems will be discussed in the following.

6.3.1 AE "Activity"

As discussed previously, measurements of AE "activity" such as with a RMS voltmeter are a simple approach to recording AE. With this approach, information about individual events is not measured, and thus opportunities to characterize AE sources or determine source location are lost. Also, without the ability to analyze individual signal waveforms or feature sets, little can be done after the test to discriminate or eliminate noise. Thus, noise sources must be carefully eliminated before the data is acquired. Monitoring AE "activity" on a structure with all instruments and test systems powered, but without applying any load, will provide indications of some noise sources. Any AE detected under these conditions will probably correspond to noise, which should be eliminated. Other noise sources such as friction of loading fixtures or grips on test specimens, may only be active during loading conditions. These noise sources are much more difficult to detect and eliminate. In general, great care should be applied when using the results from analysis of AE data if there is a significant possibility that the data is corrupted by noise.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Activity Curve Shape Analysis

After noise sources have been eliminated, and AE activity data acquired, there are a number of methods for analysis. One of the simplest is the analysis of the shape of the AE "activity" curve. For this, the AE measured variable (i.e. RMS voltage or counts) is plotted against another important test parameter such as load, stress, strain, temperature, or time of test. Often, the slope of the AE activity curve increases as significant damage occurs, as shown in Figure 6.16. The point at which this significant increase in slope occurs is associated with the release of AE from a number of sources, and has been called the "knee" of the curve. By testing a large number of nominally identical specimens and determining the load relative to ultimate failure at which the "knee" occurs, attempts have been made to use this phenomenon to quantitatively predict failure.

There are a number of limitations in using this approach. One of the most significant is that the requirement to test large numbers of specimens is time consuming and may be very expensive. Also important is the fact that any variations in the specimens, the transducers, instrument settings, etc. will nullify results from such repetitive testing. Also, if noise sources were not successfully eliminated, a sudden increase in noise can cause an apparent knee in the curve.

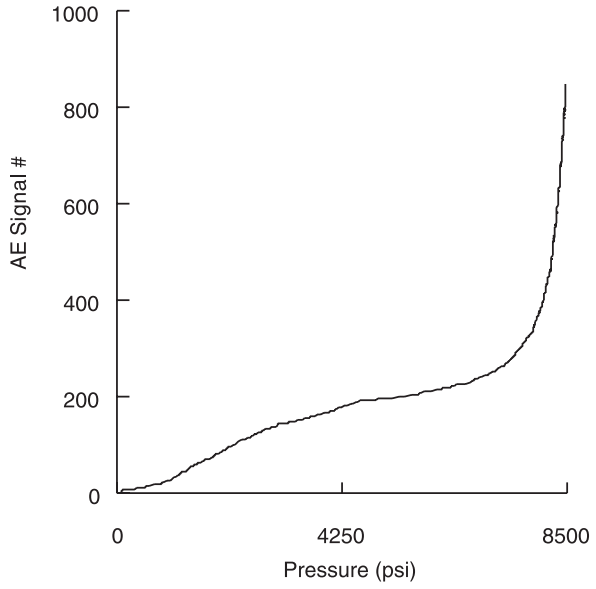


FIGURE 6.16 AE activity curve for test of composite pressure vessel illustrating increasing slope of AE activity at higher pressures.

Copyright © Marcel Dekker, Inc. All rights reserved.

Quantifying when the knee occurs may also be difficult. In some specimens, the change in slope is very gradual. Multiple damage mechanisms or the operation of the same damage mechanisms at different loads or in different locations may cause problems as multiple knees may be present in the activity curves.

Multiple Loading

Another approach is to collect AE "activity" data during repeated loading of the same structure and compare the results. According to the *Kaiser effect*, no significant AE should be produced in a specimen when it is being reloaded until it reaches the previous maximum applied load. The Kaiser effect is illustrated graphically in Figure 6.17a. A measurement of the onset of AE in a proof test of a structure at higher than service loads can give an indication of the previous maximum load that has been applied. The *Dunegan corollary* of the Kaiser effect, is that if AE is observed prior to a previous maximum load, it is an indication that some type of new damage has occurred. This corollary has been widely used in monitoring pressure vessels. AE is monitored during proof tests that are performed at regular intervals using pressures higher than that of normal operation. If no damage occurs during normal operation between proof tests, then no AE should be detected in the repeated proof tests. However, if damage has occurred, then AE may be detected.

However, in many materials, the Kaiser effect does not hold and significant AE is always detected at loads lower than the previous maximum. This is known as the *Felicity effect* and it is illustrated in Figure 6.17b. The *Felicity ratio*, which was first proposed by Fowler (22), is defined as the load at which significant AE reinitiates divided by the previous maximum load. Composite materials are an example of a material that will consistently emit AE at loads below their previous maximum and thus have a Felicity ratio of less than one. Similar to the use of the Kaiser effect, repeated loading can be used with previously determined Felicity ratio values for a given material to obtain information about the previous maximum applied load. A change in the Felicity ratio over time has also been correlated with increased damage in some materials.

There are a number of observations and theories as to why the Felicity effect occurs in some materials. The first is that the Felicity effect most often occurs in materials that exhibit significant time dependent mechanical behavior such as viscoelasticity and stress relaxation (e.g., polymeric matrix composite materials). Another is that some of the AE that occurs prior to previous maximum loads may be caused by friction of existing crack surfaces, rather than by the onset of new damage. It has also been postulated that the Kaiser effect holds on a microscopic rather than the macroscopic level (i.e., when localized stress is the important variable rather than load). Thus, although the previous maximum load may not be exceeded in a structure, a localized stress concentration may be exceeded and responsible for new damage and the associated AE.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

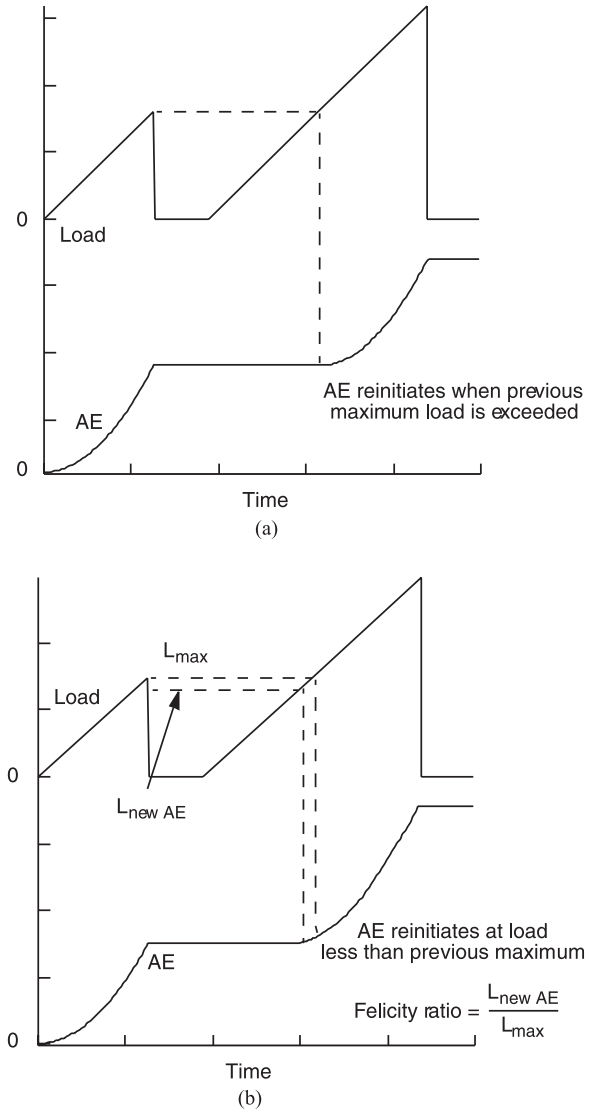


FIGURE 6.17 Illustration of load and AE activity versus time for (a) material in which Kaiser effect is followed and (b) material in which Kaiser effect does not hold and Felicity ratio is defined.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

One last technique for analyzing AE activity during multiple loadings is to consider AE released during load or pressure holds at elevated load or pressure levels, or during the release of load or pressure. A significant release of AE during periods when the load or pressure is not increasing is taken to be an indicator of damage and/or a structural deficiency.

As with "activity" curve analysis, there are also limitations to analysis of AE from multiple loading. Again, noise is a critical issue that can make determination of the Kaiser or Felicity effect difficult. Another problem is in quantifying "significant" AE in defining the Kaiser or Felicity effect. Small levels of AE may occur at very low load levels due to noise, or friction of the surfaces of existing damage. Also, the onset of AE from new damage at higher loads may be gradual.

6.3.2 Feature Set Analysis

Compared to "activity" measurements, AE signal feature measurements provide additional information about individual signals. Feature set information can be analyzed in a variety of ways. Signal arrival times at different sensors can be used to determine the location of the AE source. Other features have been used in attempts to characterize the source of the emission and/or discriminate and eliminate noise signals. Some of the more basic analysis methods are now discussed.

Source Location

The ability to locate the source of emission and thus potential damage sites, is one of the most important capabilities of AE. Even if the exact source mechanism or its severity cannot be determined from AE analysis, the location can be used to guide further nondestructive inspections with alternative techniques. In doing so, the initial inspection with AE may eliminate the need to inspect much of the structure with time consuming, and expensive, point-by-point scanning NDE techniques.

Location analysis is based on the capture of AE signals from the same source at multiple sensor locations. The arrival times of these signals at the different sensors is used along with knowledge of the velocity of acoustic wave propagation to triangulate location. In the following, only methods to locate discrete AE sources will be presented. There are also methods to locate sources of continuous type AE such as from leaks, but they are not presented here. In general, location methods for AE analysis can be broken down into two categories; zonal and point.

Zonal Location. As the name implies, zonal location does not pinpoint the position of the source, but defines it to have originated within a zone or region

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

defined around one or more sensors. Zonal methods are often used in situations when point location is not possible or will not produce reliable results. One such example is when transducers are either too widely spaced or when the material attenuation is too high to allow signals to be detected by a sufficient number of sensors for point location. Point location may also be difficult when the structure permits multiple wave modes with different and/or dispersive velocities, or creates a number of reflections. Zonal location is also often used for anisotropic materials, where accurate knowledge of the velocity variation with direction would be required for point location.

For one-dimensional or linear sensor arrays, the zone is defined as a length on either side of a sensor. For two-dimensional or planar sensor arrays, the zone is an area, and for three-dimensional sensor arrays, the zone is a volume. When the signal is detected by only one sensor in the array, the source is assumed to be located within the region around the sensor such that it is no further away from the source than the midpoint between two sensors. This is illustrated in Figure 6.18a for a planar sensor array.

If the signal is detected by more than one sensor, signal information such as arrival times or amplitudes from the different sensors can be used to further refine the zone. An example in which the order of arrival time at two sensors in a planar sensor array is used for zone refinement is illustrated in Figure 6.18b. In this figure, the ordered pairs correspond to the two sensors that would detect a signal within the refined zone and the order in which the signal would arrive at these two sensors. If arrival times are not reliable because of wave propagation effects such as the presence of multiple or dispersive modes and signal reflections, a feature such as the peak amplitude can also be used. Assuming similar mechanisms of attenuation along different propagation paths to different sensors, the source is assumed to be located closest to the sensor with the largest peak amplitude and second closest to the sensor with the second largest peak amplitude, etc.

Point Location. A number of criteria must be satisfied to permit point location. First, the signal must be detected at a sufficient number of sensors. For linear location along a one-dimensional sensor array, the signal must be detected by at least two sensors. The minimum number of sensors is three for planar location and four for volumetric or three-dimensional location. If the signal is measured at more sensors than the minimum required, then least squares or other optimization techniques can improve the accuracy of the point location. In addition to detection at a minimum number of sensors, an accurate measure of the arrival time for the same mode of propagation at each sensor is needed. In feature measurement systems, this is most often obtained from the first threshold crossing. However, as noted previously, the first threshold crossing is not always reliable. Thus, the time at which the peak amplitude occurs has also been used for the arrival time. Implicit in using the peak amplitude is the assumption that the

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

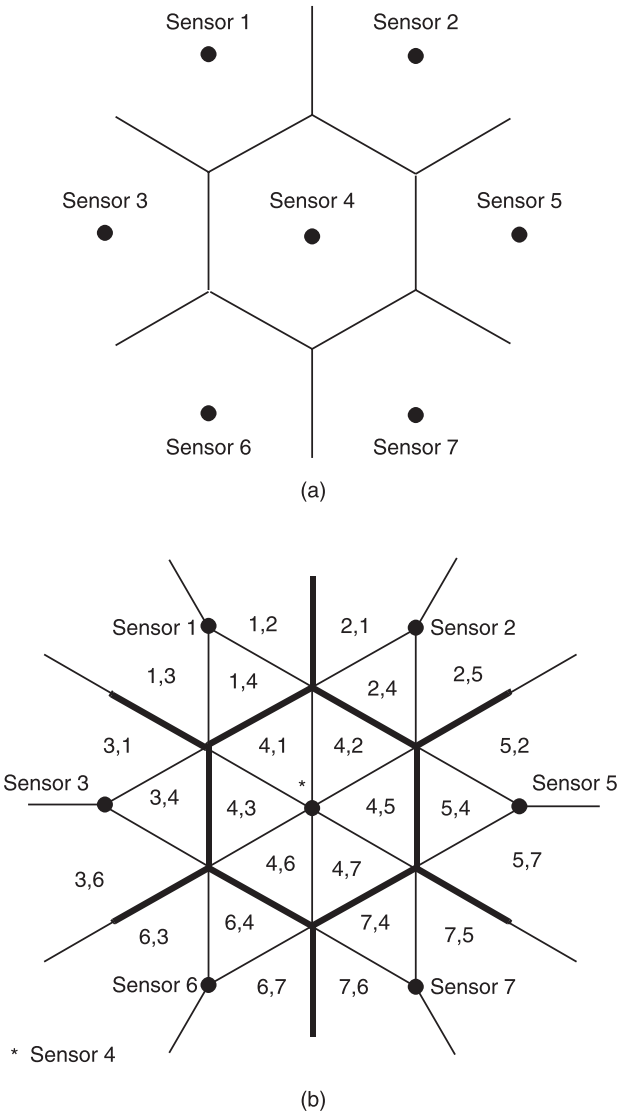


FIGURE 6.18 Illustration of zones used for planar location for two-dimensional sensor array for (a) signal detected on one sensor only and (b) signal detected by at least two sensors. In (b) the ordered pair in each zone indicates the arrival sequence for the two sensors as determined by either arrival times or peak amplitudes.

Copyright © Marcel Dekker, Inc. All rights reserved.

same mode will have the largest peak at all sensors, and that it will occur at or near the same frequency. Another necessary condition for point location is an accurate knowledge of the velocity of acoustic wave propagation in the material. Lastly, accurate measurements of the positions of the sensors are needed.

Equations can then be derived for point location of an AE source. The easiest case to consider is that of linear location along a one-dimensional array of sensors. Such a sensor array might be used to monitor a rod or pipe. For this derivation, the source is assumed to be located between two sensors as illustrated in Figured 6.19 for the case of linear location along a pipe. The distance between the source and the nearest sensor is defined as d, while the distance between the sensors is D. If v is the velocity of the mode of propagation whose arrival time is measured at both sensors, then the time required for the signal to propagate to the closest sensor is given by

$$t_1 = \frac{d}{v} \tag{6.22}$$

Since the time at which the AE source occurred is not known, a signal arrival time measurement is required at a second sensor. The time for propagation to the second sensor is

$$t_2 = \frac{(D-d)}{v} \tag{6.23}$$

The arrival time difference (Δt), is then calculated as

$$\Delta t = t_2 - t_1 = \frac{(D-d)}{v} - \frac{d}{v} \tag{6.24}$$

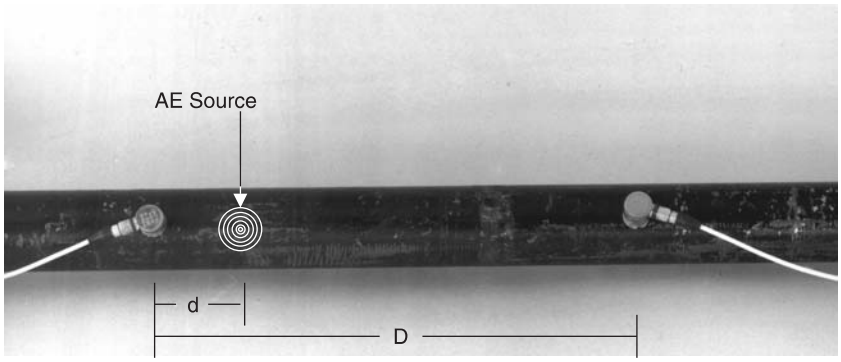


FIGURE 6.19 Sensor spacing distances for linear source location along a pipe.

Copyright © Marcel Dekker, Inc. All rights reserved.

which can be rewritten as

$$\Delta t = \frac{D - 2d}{v} \tag{6.25}$$

Solving this expression for d yields

$$d = \frac{D - \Delta t v}{2} \tag{6.26}$$

It is noted that the source cannot be uniquely located if it occurs outside the region between two sensors. If so, the measured Δt will be the same value, regardless of the distance to the source.

Example 7: Assume that two sensors are spaced 20 cm apart on an aluminum rod (longitudinal wave speed of 6320 m/sec), and that the measured Δt from the arrival of longitudinal waves produced by an AE source between the sensors is 12.66 μsec. Calculate the distance of the source from the closest sensor.

$$d = \frac{D - \Delta t v}{2}$$

$$d = \frac{20 \times 10^{-2}\ \text{m} - 12.66 \times 10^{-6}\ \text{s}\ 6320\ \text{m/s}}{2}$$

$$d = 10 \times 10^{-2}\ \text{m} - 4 \times 10^{-2}\ \text{m} = 6\ \text{cm}$$

Two-dimensional, or planar source location is somewhat more complicated. Assuming the necessary three sensors are located relative to the source as illustrated in Figure 6.20, the origin of the coordinate system is defined to coincide with sensor 0. The location problem is then to determine R, which is the distance from the origin to the source, and θ, the angle to the source with respect to the x-axis. The arrival times at the three sensors are given by

$$t_0 = \frac{R}{v} \qquad t_1 = \frac{r_1}{v} \qquad t_2 = \frac{r_2}{v} \tag{6.27}$$

The arrival time differences are defined relative to the sensor at the origin as

$$\Delta t_1 = t_1 - t_0 \qquad \Delta t_2 = t_2 - t_0 \tag{6.28}$$

Substituting results in the following pair of equations

$$\Delta t_1 v = r_1 - R \qquad \Delta t_2 v = r_2 - R \tag{6.29}$$

It is now noted that

$$Z_1 = R \sin(\theta - \theta_1) \tag{6.30}$$

and

$$Z_1^2 = r_1^2 - (D_1 - R\cos(\theta - \theta_1))^2 \tag{6.31}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

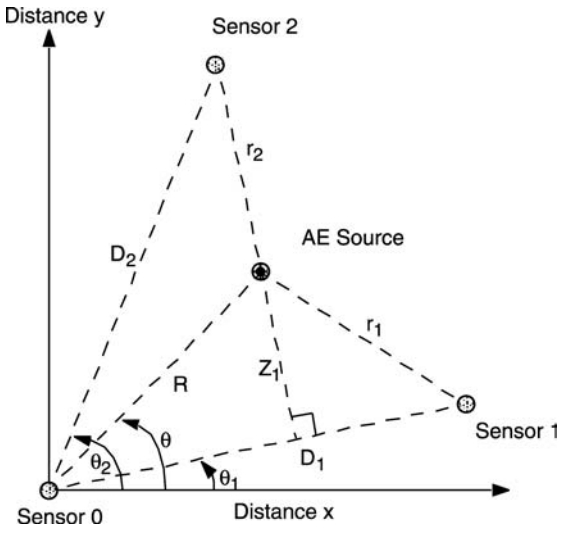


FIGURE 6.20 Sensor array geometry for planar source location.

Combining these equations and reducing terms yields

$$R^2 = r_1^2 - D_1^2 + 2D_1 R\cos(\theta - \theta_1)) \tag{6.32}$$

Inserting

$$r_1 = \Delta t_1 v + R \tag{6.33}$$

and solving for R results in

$$R = \frac{1}{2}\,\frac{D_1^2 - \Delta t_1^2 v^2}{\Delta t_1 v + D_1\cos(\theta - \theta_1))} \tag{6.34}$$

Likewise,

$$R = \frac{1}{2}\,\frac{D_2^2 - \Delta t_2^2 v^2}{\Delta t_2 v + D_2\cos(\theta - \theta_2))} \tag{6.35}$$

These last two equations can be simultaneously solved for R and θ to determine the source location. Using a similar procedure, equations can also be derived for the three-dimensional source location. For the three-dimensional case, four sensors and three arrival time difference calculations are required.

Graphically, planar source location can be represented as shown in Figure 6.21. For the pair of sensors 0 and 1, $\Delta t_1 v$ represents a hyperbola of possible source locations in the plane. Likewise, there is another hyperbola of possible locations from $\Delta t_2 v$ from sensors 0 and 2. The source is located at the intersection

Copyright © Marcel Dekker, Inc. All rights reserved.

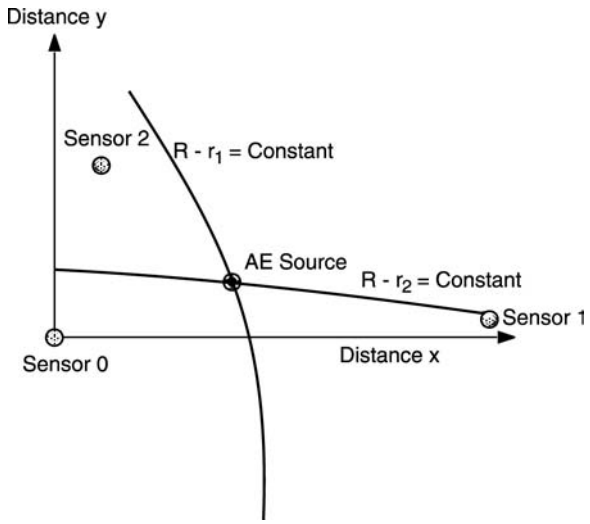


FIGURE 6.21 Intersecting hyperbolas for planar source location.

of these hyperbolas. In theory, the source location can be determined anywhere on the plane on which the three sensors are located. Source location is not restricted to positions within the region of the sensor array as is the case for linear source location. However, practical considerations such as increasing errors in arrival time determination with increased distances of propagation will effectively limit the region that can be monitored by a sensor array. Another problem in planar source location is that it is possible for the two hyperbolas to intersect at two locations. This provides a false source indication, which can be eliminated by adding a fourth sensor and Δt calculation, or by considering the hit sequence at the different sensors.

Example 8: Again assuming an aluminum material with longitudinal wave speed of 6320 m/sec, calculate the location in terms of R and θ for sensors positioned at $(0, 0)$, $(20, 0)$, and $(0, 20)$ cm with measured arrival time difference values of $\Delta t_1 = 15 \times 10^{-6}$ sec and $\Delta t_2 = 8.3 \times 10^{-6}$ sec.

First the distances between sensor pairs must be calculated.

$$D_1 = \sqrt{(y_1 - y_0)^2 + (x_1 - x_0)^2} = \sqrt{(0-0)^2 + (20-0)^2}$$
$$D_1 = 20 \text{ cm}$$

Similarly,

$$D_2 = 20 \text{ cm}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

The angles (θ_1 and θ_2) of the vectors connecting sensor 1 and sensor 2 to the origin with respect to the x_1-axis are then determined from

$$\theta_1 = \sin^{-1}\left(\frac{y_1 - y_0}{D_1}\right) = \sin^{-1}\left(\frac{0-0}{20}\right) = 0 \text{ deg}$$

$$\theta_2 = \sin^{-1}\left(\frac{20-0}{20}\right) = \sin^{-1}(1) = 90 \text{ deg}$$

Substituting into the first equation for R and θ is now made

$$R = \frac{1}{2}\,\frac{D_1^2 - \Delta t_1^2 v^2}{\Delta t_1 v + D_1 \cos(\theta - \theta_1)}$$

$$R = \frac{1}{2}\,\frac{20^2 \text{ cm}^2 - (15 \times 10^{-6})^2 \text{sec}^2 (6.32 \times 10^5)^2 \text{ cm}^2/\text{sec}^2}{(15 \times 10^{-6}) \text{ sec } (6.32 \times 10^5) \text{ cm/sec } + 20 \text{ cm} \cos(\theta - 0)}$$

$$R = \frac{400 - 89.87}{18.96 + 40\cos\theta} \text{ cm} = \frac{310.13}{18.96 + 40\cos\theta} \text{ cm}$$

Substituting into the second equation yields

$$R = \frac{372.48}{10.49 + 40\cos(\theta - 90)} \text{ cm}$$

Solving these two equations produces

$$\frac{310.13}{18.96 + 40\cos\theta} = \frac{72.48}{10.49 + 40\cos(\theta - 90)}$$

which reduces to

$$\sin\theta - 1.2\cos\theta = 0.31$$

This can be solved with an iterative or numerical approach to yield

$$\theta = 61.65 \text{ deg}$$

Substituting this value of θ back into either of the above equations for R results in

$$R = 8.2 \text{ cm}$$

Source location accuracy is affected by a number of factors related to both propagation phenomena and measurement system characteristics. Attenuation, dispersion, multiple modes, and reflections all make it more difficult to accurately determine the arrival time, or to do so on the same mode for all sensors. However, since these propagation phenomena are inherent to the structure under test, little can be done other than to recognize their effects and to optimize the measurement system characteristics such as the number of, and spacing of sensors. Another factor affecting source location accuracy in feature measurement systems is the precision of the clock used to determine arrival times. Such systems typically

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

record the arrival time relative to some digital clock system that operates at a fixed rate. A slower clock rate results in a greater uncertainty in arrival time determination and larger errors in source location.

Source Characterization and/or Noise Discrimination

AE signal feature set data is also used in attempts to characterize AE sources, as well as to discriminate and eliminate noise. Similar to "activity" curve analysis, empirical approaches are often used to correlate measured signal features with observed damage mechanisms. Laboratory tests of specimens that produce only one particular damage mechanism of interest are used to develop such correlations. As an example, transverse matrix cracking is the dominant mode of initial damage in cross-ply composite laminates and has been widely characterized with feature based AE measurements (references contained in Ref. 16). Other laminates at other loading regimes provide opportunities to study other damage mechanisms such as fiber fracture. Although such laboratory studies have shown correlations between AE features and different source mechanisms, great care must be exercised regarding their general applicability. Changes in AE signals, and their measured features in different specimens or with different sensors can be caused by propagation and measurement effects. These may mask any previously determined relationships between features and sources.

Perhaps the single most studied AE signal feature for the characterization of sources has been peak amplitude or the distribution of measured amplitudes. For example, numerous studies have measured amplitudes in composite materials from different source mechanisms. However, there have been discrepancies in the results. Some report that fiber breakage produces large amplitude signals while matrix cracking results in much smaller amplitudes. Delamination was thought to produce medium amplitude signals. Other studies concluded that matrix cracking caused large amplitude signals while fiber breakage produced low amplitudes. The reality is that the amplitude produced at the source depends on a number of factors including the local stress conditions and the energy released. A recent study showed that a very small increment of matrix crack growth produces a much smaller amplitude signal than a large matrix crack (16). In addition, variations in the source orientation with respect to the specimen geometry and sensor positions will lead to differences in what acoustic modes are excited. This will further affect measured signal amplitudes. Other propagation effects such as attenuation and reflection may also lead to significantly different peak amplitudes for the same source mechanism. These factors may explain some of the previous contradictory results and also limit the usefulness in analyzing source mechanisms solely based on amplitude.

Event duration is another feature that has been widely used in attempts to identify sources. Extremely short duration events have been attributed to noise

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

from EMI. However, EMI can also cause long duration events. At the same time, low amplitude events from other sources that barely exceed the system threshold may also have short durations. In composite laminates, long duration events have been attributed to delamination. Recent research has confirmed this for composite plates and shells (23). Wave propagation analysis suggests that this is due to the highly dispersive flexural plate mode of propagation generated preferentially by delaminations, rather than by the time duration of the source mechanism.

Since there have been such difficulties in trying to correlate a single AE signal feature with source mechanisms, multifeature analysis methods have also been devised. For example, to better differentiate EMI, both amplitude and duration have been used. Multiple feature analysis has also been extended to include artificial intelligence (AI) approaches such as pattern recognition and neural networks. While these approaches may have promise, it is still important to realize that many aspects of the measured AE signal features are dominated by propagation and measurement, not the source. Thus, care must be exercised in attempts to identify source mechanisms based solely on waveform feature analysis.

6.3.3 Waveform Analysis

Just as AE feature measurements provide more information than "activity" monitoring, full AE waveform recording affords even further details. As will be discussed below, waveform analysis can be used to provide more accurate source location, quantitative source analysis, and enhanced noise discrimination.

Source Location

The same zonal and point location equations developed previously are generally applicable to data acquired with waveform measurement systems. However, improved methods for signal arrival time determination can be used on waveform data. These provide more accurate arrival time estimates, which result in improved location accuracy, in comparison to first threshold crossing or peak amplitude arrival. Advanced signal processing techniques such as cross correlation methods have been used. Alternative threshold methods including multiple thresholds and noise or peak amplitude dependent thresholds can be used.

Waveforms can also be analyzed to identify different modes of propagation within the AE signals. This enables the arrival time for the same mode to be determined at all sensors and used for location analysis. Threshold based arrival times on different modes of propagation can result in large location errors. In addition, the arrival time of a particular frequency component of a specific mode can be identified in the signals to minimize location errors caused by dispersion. A cross correlation technique has been demonstrated for this purpose on plate mode AE signals (10). The arrival times of multiple modes at a given sensor can

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

also be measured. This additional information can be used with least squares curve fitting methods to improve location or determine location with fewer sensors.

Source Characterization and/or Noise Discrimination

Waveform data also provides an opportunity for more accurate source characterization and noise discrimination. A number of analysis approaches have been developed for this purpose, two of which are discussed below.

Source Inversion. Recalling Eq. (6.1), an AE signal can be expressed mathematically as a convolution of source, propagation, and transducer response functions. One method of characterizing AE sources is to analyze waveform data to determine the source function. This process is known as source inversion and is accomplished by removing, or deconvolving the propagation and transducer response functions. To accomplish this, the spatial derivative of the Green's function and the transducer response function must be accurately known. Both the source time function, and the magnitude and direction of source motion can then be determined with this analysis.

Source inversion has been successfully used to characterize AE source functions in laboratory specimens (14,15). However, there are a number of difficulties in implementing this approach for practical applications. The most significant is that Green's functions have only been derived for very simple geometries such as a half-space, or a thick plate of infinite lateral extent. Green's functions are not available, at least in differentiable analytic form, for complex structural geometries of AE testing interest. Another problem is that absolutely calibrated, high fidelity, broad band sensors are necessary to perform accurate deconvolution. A number of sensors including the NIST conical piezoelectric transducer (21), and capacitive and optical sensors have been developed that provide such a response. However, the sensitivity of these high fidelity sensors is lower than that of more conventional sensors. Another problem is that noise contained in the measured AE signal can lead to errors in the deconvolution process. Thus, source inversion results may not be accuratc for signals with low signal/noise ratios.

Modal Analysis. Fortunately, it is not necessary to absolutely determine source functions to identify AE sources and/or discriminate noise signals. An alternative approach, known as *Modal AE*, has been developed that analyzes the modes of propagation produced by AE sources. Although the structural geometry dictates the acoustic propagation modes that may occur, the type, location, and orientation of the source has a significant influence on the amplitude of the modes contained in the signal. This is illustrated by the waveforms shown in Figure 6.22. These signals were generated by simulated AE sources (pencil lead fractures) in a

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

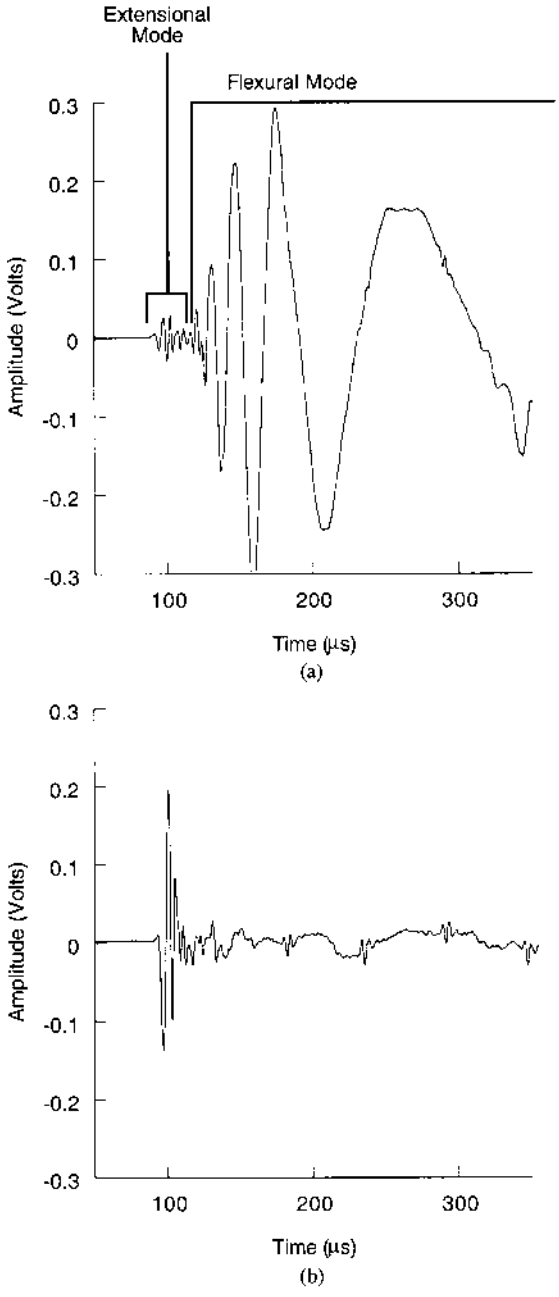


FIGURE 6.22 Waveforms illustrating the effect of source orientation on modal amplitudes for simulated AE source (pencil lead fracture) applied (a) normal to plate surface and (b) on plate edge.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

thin aluminum plate. In Figure 6.22a, the lead fracture source occurred on the surface of the plate. Both the extensional and flexural plate modes are produced by such a source and are identified in this waveform. However, the amplitude of the flexural plate mode component is much larger because of the large out-of-plane source motion. In Figure 6.22b, the simulated AE source was instead positioned along the edge of the plate at the midplane. This in-plane source motion produces an AE signal with a large extensional mode component and little or no flexural mode.

One difficulty in implementing Modal AE is a lack of automated methods to analyze the modes contained in a signal. In some cases, the different modes are comprised of different frequency components such as the plate modes shown in Figure 6.22. In this case, simple frequency analysis and filtering can be used to identify different modes. For more complicated cases where the modes contain overlapping frequency content, other more advanced signal processing techniques are being investigated. These include time-frequency methods such as the short time Fourier transform, wavelets, instantaneous frequency, and Wigner-Ville distribution. AI approaches such as pattern recognition and neural network analysis are also being studied. However, it is noted that one major problem with using AI methods is in developing adequate training data. Experiments to gather the necessary training data may be expensive and time consuming, and it may not he possible to adequately simulate all types of signals from damage mechanisms or noise. Recent developments in modeling AE signals, especially with numerical approaches such as finite element calculations (24,25) may help overcome this problem.

Modal AE analysis offers a number of advantages in comparison to source inversion. Green's functions and absolutely calibrated sensors are not required. However, an understanding of the modes of propagation allowed in the structure of interest and how different sources affect the generation of these modes is necessary. Also, high fidelity, wide band sensors are helpful, but not absolutely necessary, in identifying different modes that may contain widely varying frequency content. In some cases, Modal analysis may even be possible using waveforms acquired with high sensitivity resonant sensors. In addition to applications on thin plates, Modal AE analysis has been demonstrated for bulk modes in thick walled materials, as well as for other geometries such as pipes and shells. Some examples of Modal AE applications are presented later.

6.4 SPECIFIC APPLICATIONS

The application of AE has been attempted for monitoring a wide range of materials and structures. Specimens tested have ranged from laboratory material coupons to large structures such as bridges, petrochemical pressure tanks, and rocket motor cases. Presented below are brief descriptions of several example

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

applications. The complexity of testing ranges from event counting with a feature measurement system to advanced Modal analysis of waveform data.

6.4.1 Acoustic Emission Testing of Metal Pressure Vessels

One widely used application of AE is for the testing and recertification of metal pressure vessels such as those shown in Figure 6.23. The gas tube cylinders in this figure are used for the distribution of industrial gases, and their recertification is controlled by the Department of Transportation. ASTM Standard E 1419, *Standard Test Method for Examination of Seamless, Gas Filled, Pressure Vessels Using Acoustic Emission*, provides guidelines for their testing. Since these vessels can be quite large, scanning using other NDE techniques such as UT, PT, X-ray, etc. can be very time consuming and AE testing provides considerable cost savings. It is estimated that thousands of such gas cylinders are tested with AE each year.

Test Description

Though the dimensions of metal pressure vessels tested with AE vary, the typical dimensions of the trailer tubes in Figure 6.23 are 10 m in length, 50 cm in



FIGURE 6.23 A tube trailer being instrumented for an acoustic emission requalification test. (Courtesy of City Machine and Welding, Inc., Amarillo, TX).

Copyright © Marcel Dekker, Inc. All rights reserved.

diameter, with a 1.27 cm wall thickness. Under the ASTM standard, at least two sensors are placed on each tube, one at each end. The tubes are then pressurized to a predetermined value. This pressure value is determined by a fracture mechanics analysis such that cracks of a critical size will be caused to grow during the test. Linear location analysis is then applied to determine the position along the tube of all detected events. If a predetermined number of events occur at a given location, a follow-up UT inspection is performed at that site for defect verification and sizing.

Results

In these tubes, AE waves propagate as guided modes as shown by the waveform from a lead break simulated AE source in Figure 6.24. If the cylinder radius is greater than 20 times the thickness, then the wave propagation can be modeled as that of a plate. Advanced, Modal AE analysis techniques can be used to provide more accurate source location by making arrival time measurements on a specific

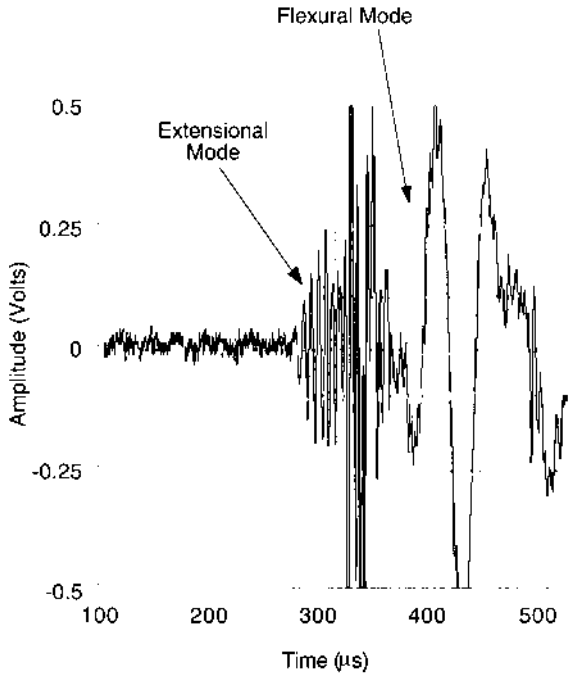


FIGURE 6.24 Detected signal from a lead break on the surface of a tube showing the mode formation. Tube diameter, 50.8 cm, wall thickness, 1.27 cm. The propagation distance is approximately 76.2 cm.

Copyright © Marcel Dekker, Inc. All rights reserved.

mode, and using theoretically calculated velocities based on the material, plate thickness, wave mode, and frequency (10,16). Using this approach, source location accuracy is on the order of 1% of the tube length.

Modal AE analysis of the detected waveforms can also aid in noise discrimination. Visual inspection of the waveforms, which are similar to those from an A-scan in ultrasonics, allows experienced users to quickly identify signals as that created by actual cracks or that from other noise sources. Such an analysis is based on the modes contained in the signal, their frequency content, and the arrival times of the waves at the different sensors, all of which alter the appearance of the waveforms. Post-test processing techniques can be used to automate this analysis using approaches such as wavelect decomposition, filtering, and SONAR algorithms.

6.4.2 Impact Damage in Graphite/Epoxy Composite Pressure Vessels

Acoustic emission has also been applied to the testing of composite pressure vessels. However, composite vessels are more difficult to evaluate because of the multitude of AE sources such as matrix cracking, fiber breaking and debonding, and delamination between multiple composite layers. Composites, especially polymer matrix composites, also typically exhibit higher acoustic attenuation, which makes the implementation of AE more difficult. In the following example application, the capability of AE to detect the presence of, and determine the significance of, impact damage in graphite/epoxy composite pressure vessels was investigated. AE was chosen for this application because of its potential to monitor large areas of composite structures that are being increasingly used on aerospace vehicles. Multiple proof tests were performed on six vessels, of which four had been subjected to controlled impact. An analysis of the number of events (as determined from the AE hits recorded with a feature based AE system), particularly during depressurization proved to be useful in assessing impact damage.

Test Description

The cylindrical vessels (125 mm dia.) were constructed by filament winding a 3 mm thick graphite/epoxy wall onto a 1.3 mm thick aluminum liner. The resulting vessels had an overall length of 457 mm, with a cylindrical section 286 mm long. Each vessel received a single, virgin proof pressure ramp to 58.6 MPa. After the virgin pressurization, four of the vessels received a single, controlled impact of approximately 14.9 J that was applied perpendicular to the cylindrical surface using a 12.7 mm diameter hemispherical tup. Delaminations of the composite were evident at the impact site on radiographs. Four of the six vessels also received thermal exposures including moderate oven heat (107°C)

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

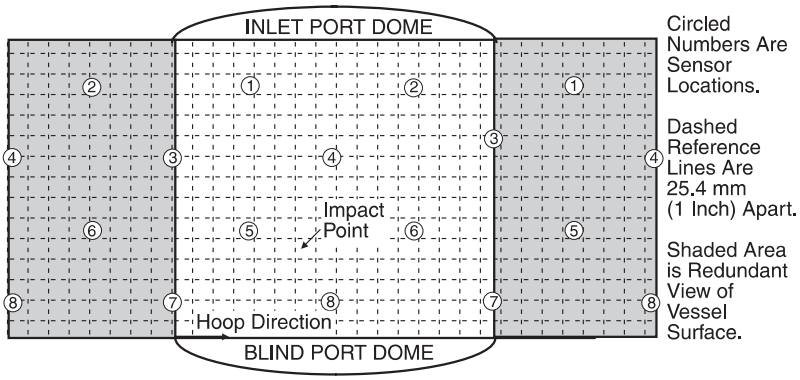


FIGURE 6.25 Planar schematic view of AE sensor positions and impact location on vessel exterior surface.

and/or cryogenic cold (liquid nitrogen at −195.8°C). Eight high fidelity AE sensors (Digital Wave Corporation B1025T) were positioned on the vessel relative to the impact point as shown in Figure 6.25. The sensor locations were selected in a manner to uniformly cover the cylindrical area of the vessel. The proof test fixture, shown in Figure 6.26, was constructed to maximize AE test

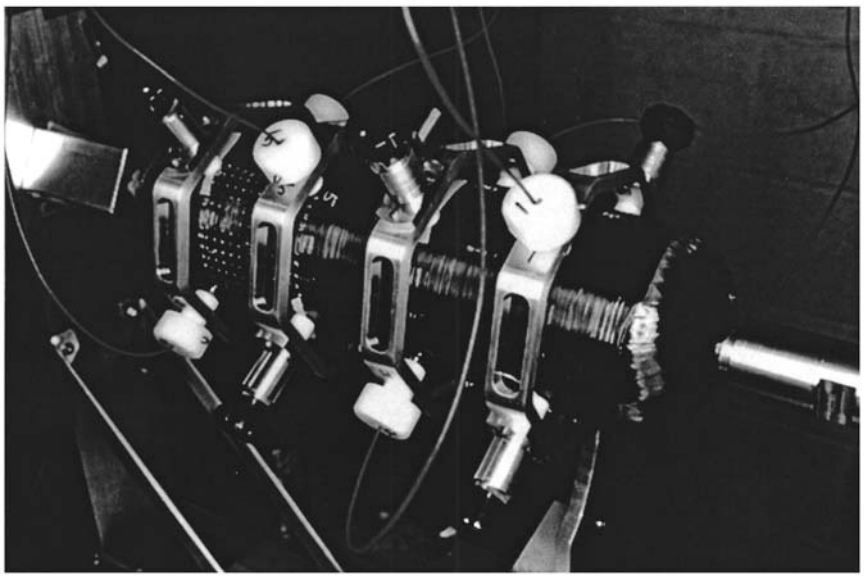


FIGURE 6.26 Photo of typical vessel installed in proof test fixture.

Copyright © Marcel Dekker, Inc. All rights reserved.

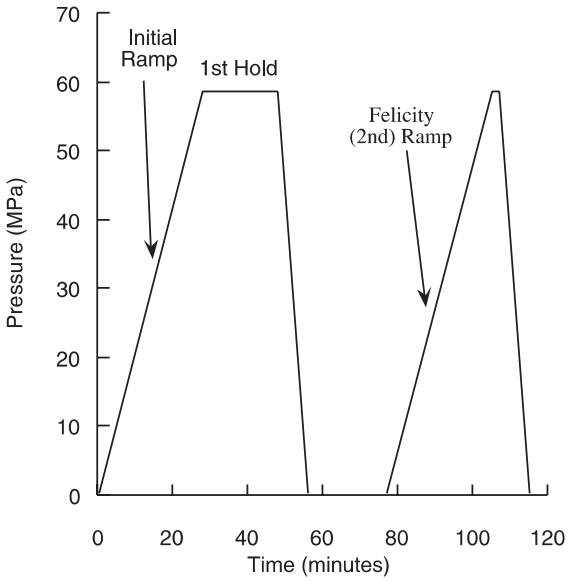


FIGURE 6.27 Schematic of proof pressure profile used in testing.

repeatability. The vessels were pressurized as illustrated in Figure 6.27. AE was recorded (Physical Acoustics Corporation (PAC) Spartan AT) during both ramps, the first hold, and the second unload. A preamplifier (PAC 1220A) gain of 40 dB was used along with a bandpass filter from 85–325 kHz. An additional 36 dB gain was applied by the measurement system and a value of 500 μs was used for the Hit Detection Time.

After proof testing, the vessels were hydraulically pressurized at 20.7 MPa/min until they burst. Both of the nonimpacted vessels burst at the knuckle region nearest the pressure inlet end at an average pressure of 89.26 MPa, and the entire inlet end was essentially blown out. The impacted vessels, however, all clearly showed a failure initiation site near the impact. Their burst strengths were markedly reduced in comparison to the nonimpacted vessels as shown in Table 6.2. A more complete description of the vessels and testing is given elsewhere (26).

Results

After the proof test, the measured AE hit data were segregated into AE events (27). A summary of the AE event data is also given in Table 6.2 for several portions of the proof test sequence. Note the fairly uniform quantity of AE produced during the first hold for all vessels. The Felicity (second) ramp yielded a

Copyright © Marcel Dekker, Inc. All rights reserved.

TABLE 6.2 Composite Pressure Vessel Condition and Strength vs. Sum of AE Events

Vessel ID	Vessel condition	Burst strength (MPa)	Burst strength (% of average nonimpacted strength)	Sum of AE events from 1st hold (last 10 min of 20 min hold)	Sum of AE events from Felicity (2nd) ramp (* = aborted 2nd ramp)	Sum of AE events from 2nd unload (data from 57.2 MPa to 0.69 MPa)
V- 7	Impacted	68.19	77.1	4094	5738	1880
V- 3	Impacted	75.02	84.0	1392	2078	978
V- 6	Impacted	76.33	85.5	2092	2690	612
V- 5	Impacted	78.40	87.8	1732	$262^{*} + 3723$	1055
V- 4	Nonimpacted	86.26	96.6	1198	752	41
V- 2	Nonimpacted	92.26	103.4	1315	786	26

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016



Copyright © Marcel Dekker, Inc. All rights reserved.

minimum threefold greater total AE for impacted vs. nonimpacted vessels. The second unload ramp provided a minimum order of magnitude difference in AE for impacted vs. nonimpacted vessels, it is also noted that (except for V-5) the overall number of unload events properly ranks the burst strengths of the vessels, both impacted and nonimpacted.

To analyze the AE during pressure unloading, a new characteristic, the *Shelby countup ratio* (SCR) (27), was developed. A Shelby ratio is defined as "the load at which a particular evaluation level of acoustic emission activity has been reached during a depressurization or unload cycle, divided by the previous maximum load" the *countup* version of the ratio adds the caveat that the AE activity is assessed starting near the lowest pressure of the unload and worked backwards with respect to time toward the highest pressure. By this definition the SCR has a value ranging from zero to exactly 1.0. If the AE activity never reaches the preset level, then the SCR is also defined to be 1.0. This is somewhat analogous to the Felicity ratio that is instead computed for AE data from an increasing load cycle. In the SCR values presented here, "AE activity" was defined as cumulative AE events.

Two correlations of burst strength with average SCRs are shown in Figure 6.28. The averaging methodology and rationale is explained elsewhere (26). In Figure 6.28a, AE data from the entire vessels were used. For this case, the average Shelby ratio values demonstrate a generally linear correlation with burst strength, and provide a distinct separation between impacted and nonimpacted vessels. In Figure 6.28b, only the AE events originating from a small region around the vessels' failure location were used in calculating the average Shelby ratio. As might be expected, a slightly better correlation with burst strength is demonstrated.

To summarize, the cumulative quantity of AE events recorded during depressurization of graphite/epoxy pressure vessels was an order of magnitude larger for vessels having impact damage as compared to those with no impact damage. In contrast, AE event data recorded from Felicity ramps or holds only displayed a two- to fourfold distinction between impacted and nonimpacted vessels. The newly defined Shelby ratio provided a methodology for assessing unload AE data in a fashion somewhat analogous to the Felicity ratio. Approximately linear correlations were obtained between average Shelby ratios and vessel residual strength.*

6.4.3 Materials Testing—TMC Study in Composites

This next application example demonstrates the capability of AE as a laboratory tool to study the development of damage in composite materials. Waveform

* Reproduced with permission from Technomic Publishing Co., Inc., copyright 1998.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

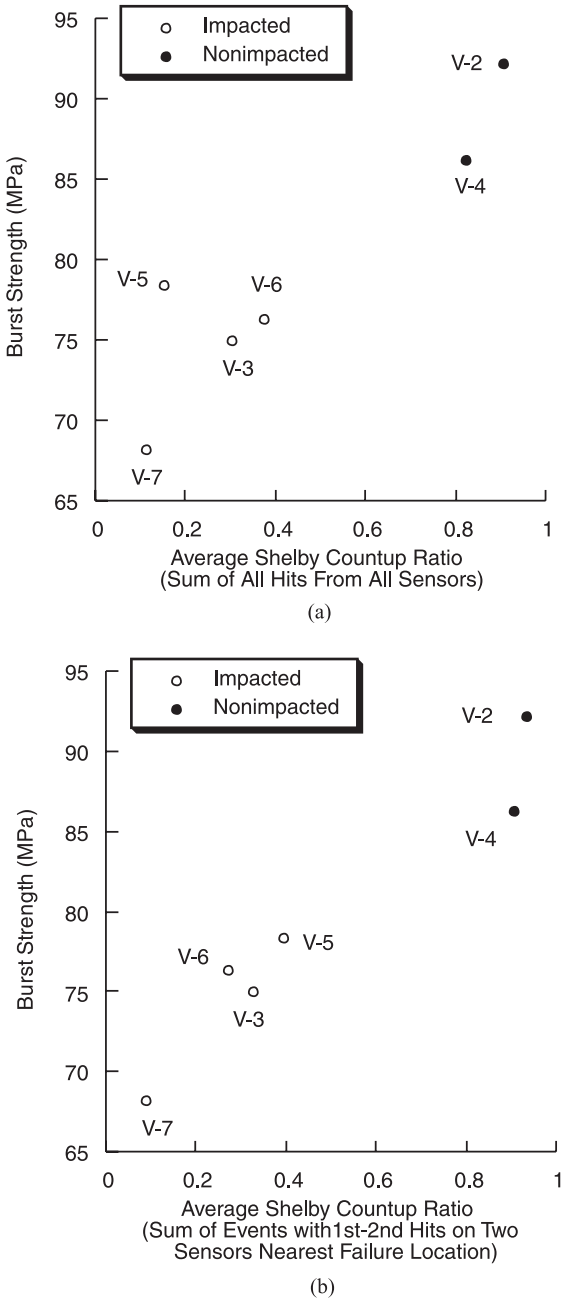


FIGURE 6.28 Correlation of burst strength with average Shelby countup ratio for second unload for hits from (a) all sensors, and (b) sensors nearest impact damage site.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

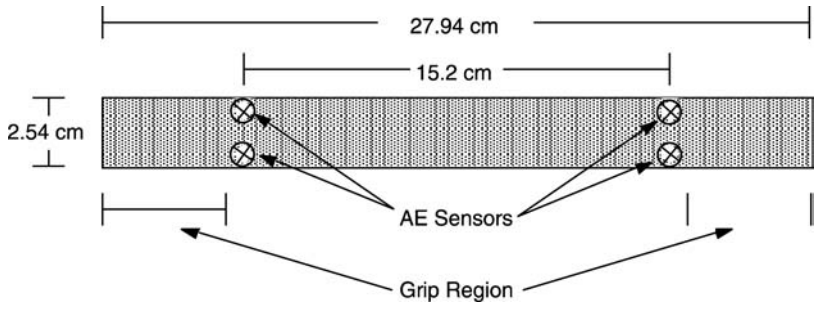


FIGURE 6.29 Diagram of specimen and sensor positions used for AE detection of transverse matrix cracking.

acquisition and Modal AE analysis were used to determine the loads and stresses at which transverse matrix cracks initiated in different cross-ply graphite/epoxy laminates. AE also provided information about the dynamics of crack initiation and progression. AE was chosen for this application because of its ability to continuously monitor specimens in situ, while they were being tested in the load frame. Other techniques that had been used previously, such as penetrant enhanced radiography or edge microscopy, required either that the specimens be removed from the load frame, or that measurements be made only at discrete load intervals.

Test Description

Tensile coupon specimens (2.54 cm wide by 27.94 cm long) of AS4/3502 graphite/epoxy composite material were tested under stroke control loading (0.127 mm/min). Specimens from six different cross-ply laminates were tested. The stacking sequences were $[0_n, 90_n, 0_n]^*$ where n ranged from 1–6. Thus, the samples varied in thickness from 3–18 plies. Specimen end tabs, which are often used in AE tests to reduce noise signals created by grip damage, were not used in these tests. Noise signals were identified and eliminated based on waveform Modal analysis.

Broadband, high fidelity sensors (Digital Wave Corporation (DWC) B1000) were used to detect the waveforms. Rather than a single sensor at either end of the specimen as in most previous AE coupon studies, four sensors were used. At either end of the nominally 152 mm specimen gage length, a pair of sensors were positioned. A diagram of a specimen showing the sensor positions and the grip regions is shown in Figure 6.29. In addition to providing the linear location of the

* $[0_n, 90_n, 0_n]$; the numbers indicate the orientation of the plies in degrees and the subscripts indicate the number of plies in the indicated orientation.

Copyright © Marcel Dekker, Inc. All rights reserved.

position of cracks, this sensor arrangement also allowed determination of the lateral site of crack initiation (i.e., provided 2-D planar location). In addition to digitizing the detected AE waveforms, the measurement system (DWC F4000) recorded the voltage from the load cell output to determine the load and stress at which each signal occurred.

Upon acquisition of one or more transverse matrix crack AE signals, the load test was stopped and the specimen removed from the test machine. One edge of the specimen, which had been polished prior to testing, was examined under an optical microscope. The specimen was mounted on a linear translation stage to allow measurement of crack locations for comparison with the AE data. Other techniques such as backscatter ultrasonics, penetrant enhanced radiography, and destructive sectioning and microscopy were also used to confirm further the crack locations and to provide information about the lateral extent of the cracks. Further details about the testing and analysis can be found elsewhere (16).

Results

A typical waveform from a transverse matrix crack source is shown in Figure 6.30a while that from a noise source is shown in Figure 6.30b. These signals are quite complicated by the presence of multiple signal reflection components that have propagated back and forth across the narrow width of the coupon specimen. However, the crack signal contains high frequency extensional mode components while the grip noise signal contains low frequency flexural mode components.

Planar location analysis showed that all cracks initiated at the specimen edges. A typical set of four waveforms from a crack is shown in Figure 6.31. These signals clearly show the time delay between the sensor pairs associated with the crack initiation site being located along the edge. Furthermore, differences in amplitudes between the sensors within a pair indicate the increased attenuation for signal propagation across the specimen. The differences in signal amplitudes and frequency content for signals detected at opposite ends of the specimen (and thus different distances of propagation distances) should also be noted. These differences, which are caused by attenuation and dispersion, can have significant effects on location accuracy and conventional amplitude distribution analysis.

To illustrate the effects that attenuation and dispersion can have on location accuracy, the linear location results were compared for the two pairs of sensors on either edge, and for cracks located at different positions along the specimen. The most accurate linear location was obtained by using the two sensors on the same edge as the crack initiation site. The pair of sensors on the opposite edge produced much more inaccurate locations because of several effects. First, the direct waves traveled an increased distance crossing the width of the specimen. Also, because of the anisotropy of the composite material, the velocity of the

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

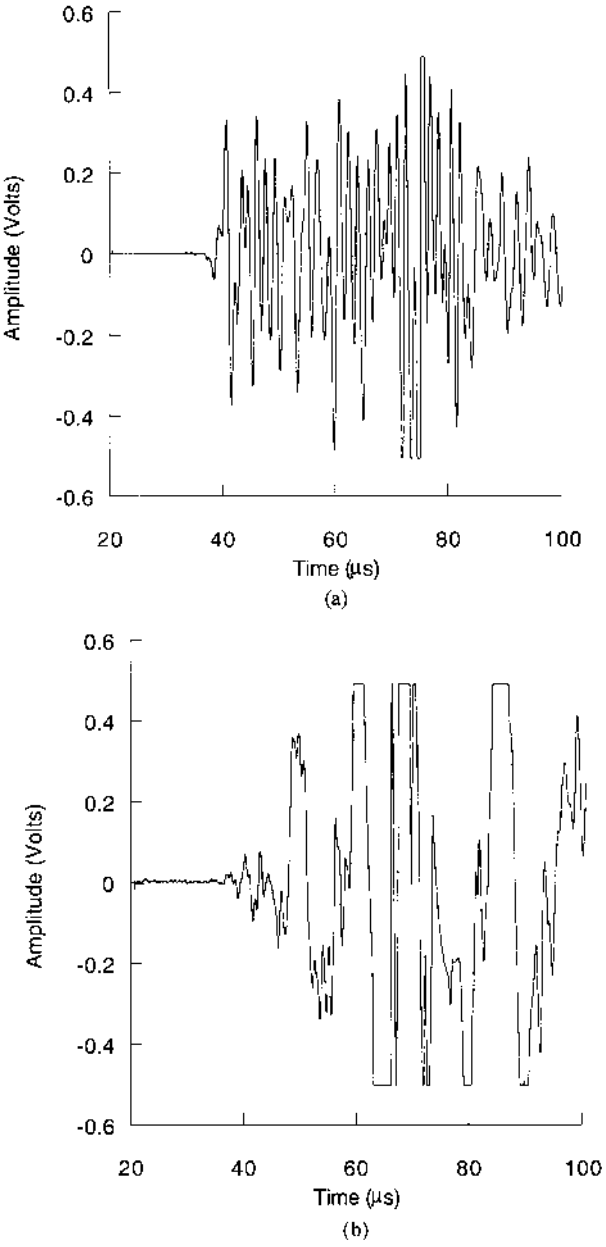


FIGURE 6.30 Typical waveforms from (a) transverse matrix crack source and (b) noise source caused by grip.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

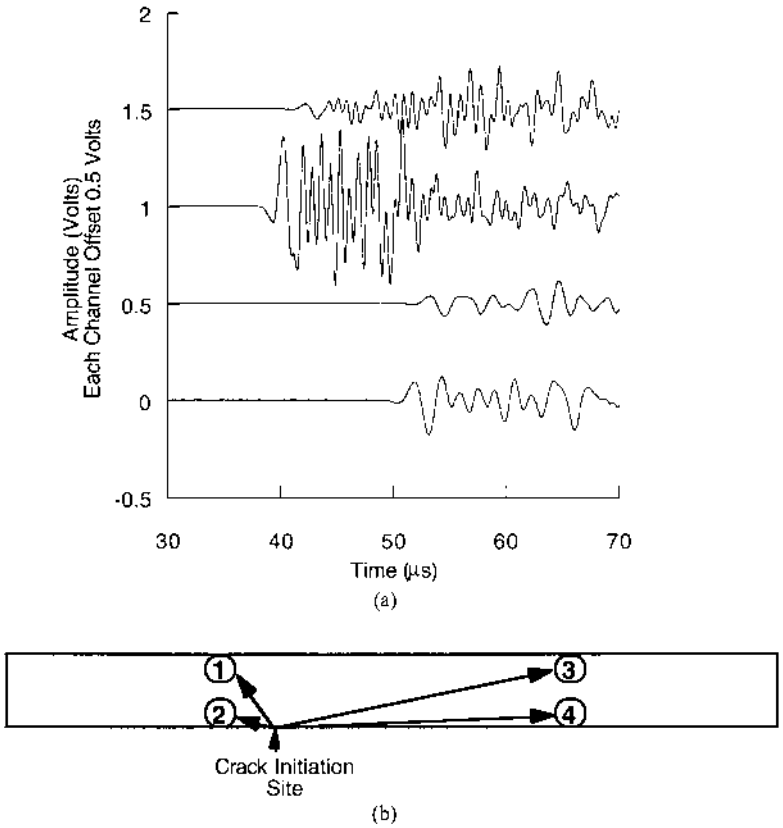


FIGURE 6.31 (a) Set of four channel waveforms indicating crack initiation along the specimen edge and (b) diagram showing sensor positions, crack initiation site, and rays of direct propagation for the AE signal.

waves traveling at angles across the specimen is different (slower) than the signals propagating directly along the specimen edge. Finally, the increased distance of propagation and the more attenuating direction of propagation contributed to reduce the amplitude, making arrival time determination more difficult. These effects would also affect location measurements made from transducers mounted in the center of the specimen width as is most commonly used. Using the results from the sensors on the same edge as the crack initiation site, the average of the absolute value of the difference in crack locations from AE and microscopy was 3.2 mm for a nominal sensor gage length of 152 mm. Typically, the cracks located toward the center of the gage length were more accurately located with AE while

Copyright © Marcel Dekker, Inc. All rights reserved.

those closer toward one end of the specimen showed larger errors. This is to be expected as the signals are more equally affected by attenuation and dispersion when the distance of propagation is nearly the same, which occurs for cracks near the center of the gage length. Figure 6.32 shows the location results as compared to microscopy locations for a specimen that was loaded until nine cracks within the gage length were detected.

A difference in crack initiation and propagation behavior was observed between the specimens with thick 90° layers ($n = 3$ or larger) and those with thin 90° layers. For thick 90° layers, the cracks initiated at the specimen edge and rapidly propagated across the width of the coupon. The AE signals produced by such crack growth were generally large in amplitude with the amplitude of the signals being proportional to the 90° layer thickness. For the specimens with $n = 1$ or 2, the cracks produced AE signals with such smaller amplitudes that some cases were not detectable. Analysis with other techniques such as radiography showed that cracks in these specimens did not rapidly grow across the specimen width after initiation. Thus it was shown that the *same* source mechanism transverse matrix cracking, can produce a variety of AE signal amplitudes which depend on the length and width of crack, and thus the amount of energy released.

In summary, this study showed the value of using AE to characterize the development of damage in composite specimens. Modal analysis of the AE signals provided a method to discriminate and eliminate noise signals. Planar location analysis provided information about the sites of crack initiation. The effects of attenuation and dispersion on location accuracy were illustrated. Also, the variability in signal amplitudes from the same source mechanism was demonstrated.

6.4.4 Fatigue Crack Detection in Aerospace Structures

In this final application example, Modal AE analysis of waveform data was used for crack growth detection during the fatigue testing of an F-16 bulkhead (FS479)

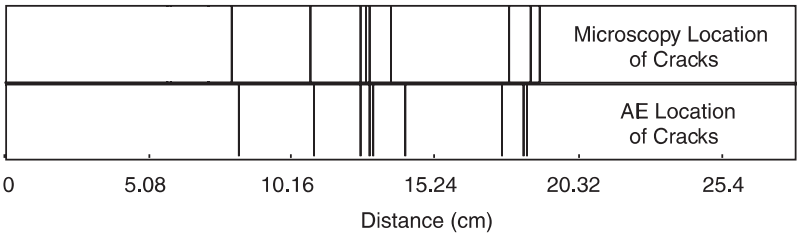


FIGURE 6.32 Location results for transverse matrix cracking in a specimen loaded until nine cracks were detected.

Copyright © Marcel Dekker, Inc. All rights reserved.

subcomponent (28). This particular bulkhead supplies the attachment points for the vertical stabilizer of the F-16. In the actual FS479 bulkhead, fatigue crack growth occurs at the base of the attach pads, due to the extreme loading of the vertical stabilizer. Inspection of this component is difficult, since access to the bulkhead is limited. Also, NDE techniques such as dye penetrant and eddy current can only give surface crack growth information, but no depth information. As illustrated in this study, Modal AE can provide the full crack profile, which can be used to develop improved life prediction models. An assessment of the AE noise environment on the F-16 with engines at full throttle and avionics systems operating suggests that AE can also be used for in-flight crack detection.

Test Description/Results

Eight AE sensors were arrayed around the attach pad area, as shown in Figure 6.33. Sensors were positioned so that the detected wave did not propagate around any bolt attachment holes or sharp corners. This would have distorted the wave, resulting in incorrect source location analysis. Because of the thickness of the attach pad area, the waves detected were bulk longitudinal waves instead of plate waves discussed in the last example. A set of waveforms from a crack source are shown in Figure 6.34. Arrival times of the longitudinal wave at the various sensors were used to triangulate the source using a nonlinear least squares fit. The test results showed that crack growth could be detected during the testing, and separated from noise sources based on the wave modes observed in the captured waveforms. Three-dimensional source location showed that accurate crack

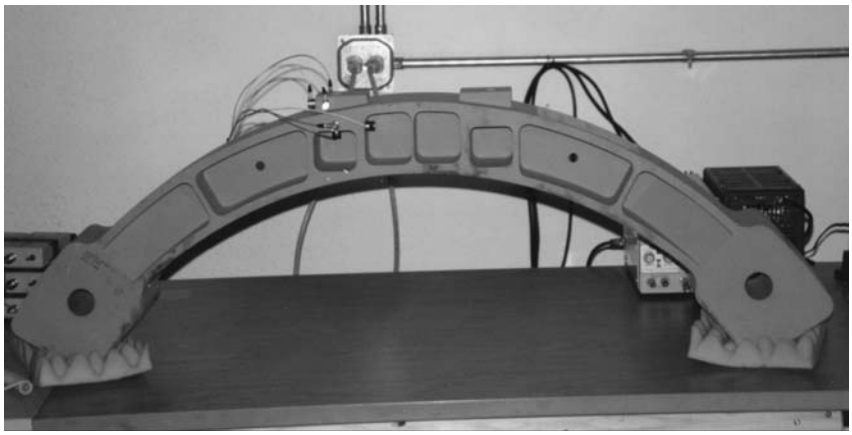


FIGURE 6.33 F-16 bulkhead subcomponent. The vertical stabilizer of the F-16 is bolted to the attach pads at the top of the bulkhead.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

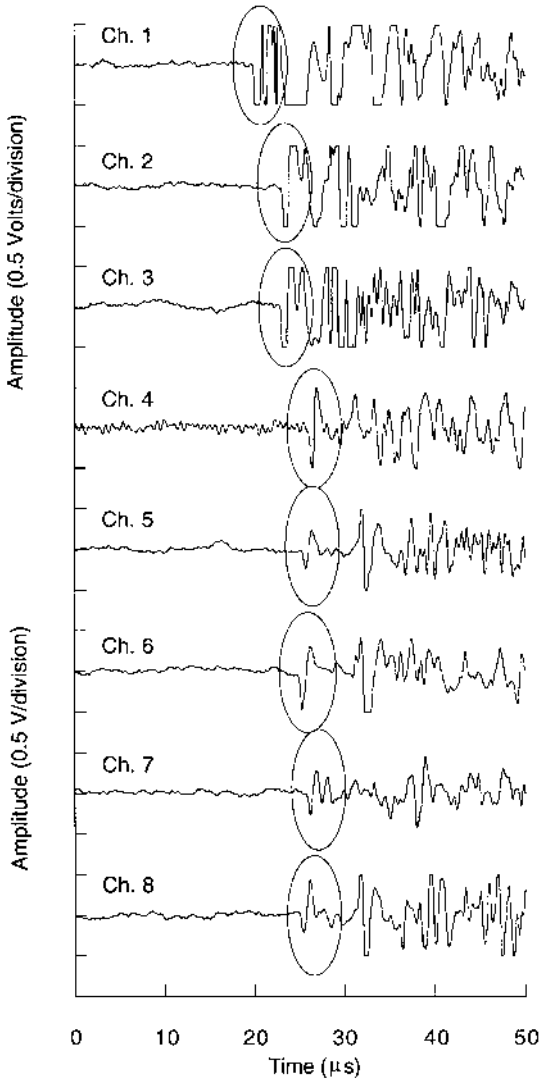
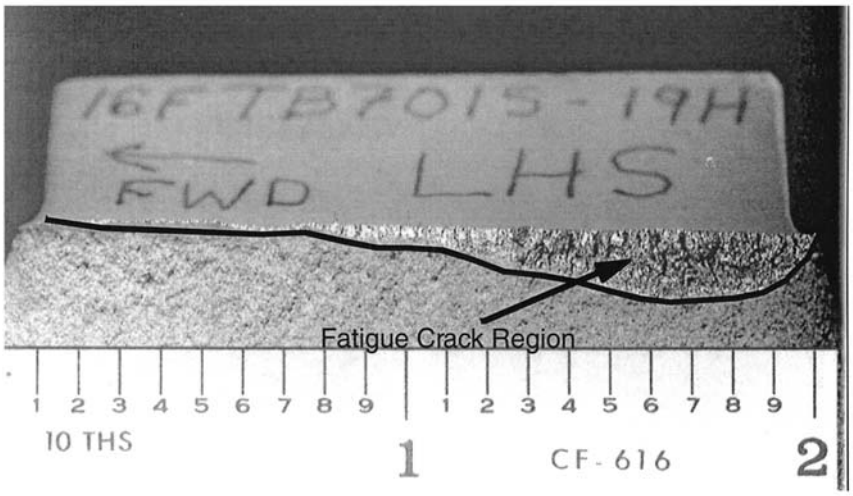


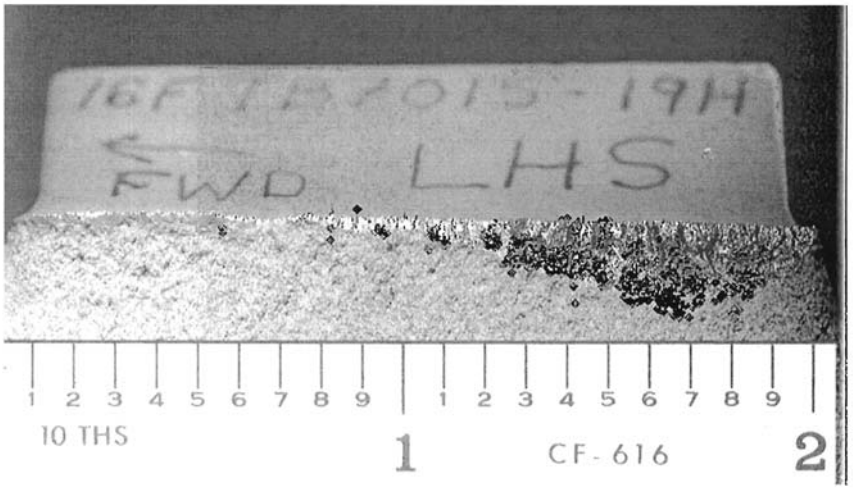
FIGURE 6.34 Waveforms from fatigue crack growth on the F-16 bulkhead. Longitudinal wave arrivals are circled.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016



(a)



(b)

FIGURE 6.35 (a) Photograph of cracked specimen with extent of fatigue crack prior to failure outlined and (b) same photograph of crack with AE source locations superimposed.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

profiling could be performed, and correlated to the metallurgical analysis of the fatigue crack growth. Figure 6.35a shows a photograph of the cracked specimen with the approximate extent of the fatigue crack prior to ultimate failure outlined. In Figure 6.35b the specimen is again shown with symbols for individual AE source locations superimposed on the photograph.

Testing was also performed to determine if noise sources from avionics, engines or control surfaces would interfere with the detection of crack growth. An F-16 was instrumented with AE sensors, as shown in Figure 6.36, and then run to full afterburners on the ground. Waveform AE data was acquired during all



FIGURE 6.36 Instrumented F-16 for engine noise tests.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

stages, from engine startup to full afterburners. The testing showed that the noise from the engines, avionics and control surfaces was of much lower frequency content than that from the crack growth, indicating that these noises could he filtered from the crack growth data.

6.5 SUMMARY

Acoustic emission monitoring is, in a number of ways, unique in comparison to other more conventional NDE techniques. In techniques such as ultrasonic testing and radiography, a low level of some form of energy is input into the material or structure. An evaluation of material properties and an identification of any flaws is made based on how that energy propagates through, and interacts with the material. AE, on the other hand, is based on passive monitoring of acoustic energy released by the material or structure itself while under load. The initiation and/or propagation of damage mechanisms produce the acoustic waves. Analysis of AE signals can be used to detect, locate, and assess these damage mechanisms. With this approach, AE provides a number of advantages. An array of acoustic sensors can provide global monitoring of a region of a structure. Time consuming, point-by-point scanning is not required. Also, it may be possible to monitor a structure in situ rather than having to remove it from service for inspections.

However, a number of aspects must be considered to develop successful applications of the AE technique. Probably most important is the recognition that acoustic signals may also be produced by sources other than the damage mechanisms of interest. Such sources need to either be eliminated during testing, or the resulting signals must be discriminated and eliminated during data analysis. Important for data analysis is the understanding that the measured acoustic waveforms are a result of the convolution of source, propagation, and measurement effects. As such, the propagation and measurement effects can significantly affect the signal and mask effects related to the damage mechanism. Propagation effects may include attenuation, velocity dispersion, multiple modes of propagation, reflections, and mode conversions. Measurement effects include sensor response and acquisition filtering and electronics.

There are also a number of measurement approaches and analysis techniques for AE. Measurements range from simple AE activity monitoring (i.e., acoustic signal counting) up to full waveform digitization and recording. Analysis approaches range from simple curve analysis and empirical correlations, to artificial intelligence methods and neural networks. The appropriate measurement system and analysis approach for a given application depends on a number of factors including the test objectives, methods available for loading the structure, applicable standards and codes, and available instrumentation and sensors.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

ACKNOWLEDGMENTS

The author would like to acknowledge the efforts of a number of contributors to this work including Dr. Marvin Hamstad, Karyn Downs, Dr. Steve Ziola, Capt. Scott May, Ian Searl, and Neal Phelps. The author would also like to acknowledge the love and support of his family, Karen, Drew, and Kayla, during the preparation of this work.

PROBLEMS

1. Derive equations for the bulk longitudinal and shear wave velocities in terms of the density (ρ), Young's modulus (E), and Poisson's ration (v).
2. Derive equations for the flexural and extensional plate velocities in terms of the density (ρ), and the Lamé elastic constants (λ and μ).
3. Calculate the longitudinal and shear velocities for carbon steel, which has a density of 7800 kg/m^3 and elastic properties of $E = 200$ GPa and $v = 0.27$.
4. Calculate the approximate Rayleigh velocity for carbon steel that has the material properties given in Problem 3.
5. Calculate the plate theory extensional velocity for carbon steel that has the material properties given in Problem 3.
6. Calculate the plate theory flexural velocity for a 2.0 mm thick carbon steel plate as a function of frequency and for specific frequency values of 50 kHz and 200 kHz.
7. A unidirectional composite material with a thickness of 2.26 mm and a density of 1550 kg/m^3 has the following in-plane stiffness properties. Calculate the extensional plate mode velocity along 0, 60, and 90° propagation directions.

$$A(N/m) = \begin{bmatrix} 3.29 \times 10^8 & 6.58 \times 10^6 & 0 \\ 6.58 \times 10^6 & 2.19 \times 10^7 & 0 \\ 0 & 0 & 1.35 \times 10^7 \end{bmatrix}$$

8. The same unidirectional composite material as in Problem 7 has the following bending stiffness properties. Calculate the flexural plate mode velocity as a function of frequency, and for a frequency value of 50 kHz, along 0, 60, and 90° propagation directions.

$$D(Nm) = \begin{bmatrix} 140.0 & 2.80 & 0 \\ 2.80 & 9.35 & 0 \\ 0 & 0 & 5.75 \end{bmatrix}$$

9. For a constant distance of propagation (r) and force input (F_3), calculate the ratio of the peak longitudinal displacement to the peak shear displacement for

Copyright © Marcel Dekker, Inc. All rights reserved.

propagation at 30° with respect to the source direction for carbon steel with elastic properties given in Problem 3.

10. Calculate the distance as a function of an absorption attenuation coefficient (α_A) such that attenuation due to geometric spreading (α_{gs}) and absorption are equal for plate mode propagation.
11. Calculate the maximum allowable sensor spacing such that the signal amplitude decrease is no more than a factor of 5 between sensors when the attenuation is (a) 10 dB/m and (b) 35 dB/m.
12. (a) Calculate the theoretical dynamic range for a 12 bit A/D converter. (b) Calculate the reduction in dynamic range if the A/D converter only provides 10 effective bits. (c) Calculate the dynamic range of A/D converters per bit.
13. Assume a linear array of two sensors spaced 53 cm apart on a carbon steel structure (density = 7800 kg/m^3, $E = 200$ GPa and $v = 0.27$) and a source that occurs between the sensors at a distance of 14 cm from the nearest sensor. Calculate the Δt that will be measured if both sensors detect the arrival of the longitudinal wave.
14. For the sensor spacing, source location, and material properties given in Problem 13, calculate the error in location position that will occur if the Δt is instead measured on the shear wave arrivals, but the longitudinal velocity is incorrectly used.
15. Again assuming a carbon steel plate material with the properties specified in Problem 13, calculate the location in terms of R and θ for sensors positioned at (0, 0), (20, 5), and (7, 20) cm with measured extensional plate modal arrival time difference values of $\Delta t_1 = 7.2 \times 10^{-6}$ sec and $\Delta t_2 = 12.4 \times 10^{-6}$ sec.
16. Now assume that the measured Δts in Problem 15 are erroneous due to noise and digitizer quantization error such that they are both smaller by 2.0 μsec, calculate the error in source location.

GLOSSARY

Acoustic emission (AE): (1) The release of transient elastic (acoustic) waves produced by a rapid redistribution of stress in a material, (2) transient elastic waves produced by a rapid redistribution of stress in a material, and (3) the method of monitoring transient elastic waves produced by a rapid redistribution of stress in a material to nondestructively detect damage and assess structural integrity.

Attenuation: The decrease in amplitude of an acoustic wave as a function of propagation distance, which is caused by (a) absorption, (b) geometric spreading, (c) dissipation of acoustic energy into adjacent media, and (d) dispersion.

Burst emission: A qualitative description of acoustic emission in which signals from discrete individual AE sources are observable.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Continuous emission: A qualitative description of acoustic emission in which numerous, rapidly occurring AE sources produce complex, overlapping signals in which signals from individual AE sources are no longer discernable.

Dispersion: Variation in velocity of propagation as a function of frequency.

Felicity effect: During repeated loading of a material, the presence of significant acoustic emission at loads below the previous maximum applied load.

Felicity ratio: A quantitative measurement of the felicity effect that is defined as the load at which significant AE initiates divided by the previous maximum applied load.

High fidelity transducer: A transducer whose response as a function of frequency has little or no variation over a wide frequency bandwidth, and that produces an output that accurately represents the measured variable (i.e., in acoustic emission, the surface displacement, velocity, or acceleration).

Hsu-Nielsen source: The fracture of a pencil lead against the surface of a material, which is used to simulate an AE source. The resulting acoustic wave is often used for such purposes as the evaluation of transducer coupling and response, the measurement of acoustic wave velocity or attenuation, and the determination of AE source location accuracy.

Kaiser effect: A demonstration of the irreversible nature of the acoustic emission phenomenon in which no AE is detected in a material undergoing repeated loading until the applied load exceeds the previous maximum.

Pencil lead break: See Hsu-Nielsen source.

Resonant transducer: A transducer that uses the amplification caused by a mechanical resonance (or multiple resonances) to provide increased sensitivity.

Sensor: See transducer.

Transducer: A device that converts one form of energy to another, in acoustic emission, this is typically a (piezoelectric) device that converts the mechanical energy of the AE wave into electrical energy.

VARIABLES

A	Signal amplitude
V	Voltage
t	Time
f	Frequency
ω	Angular frequency ($2\pi f$)
T	Transducer response function
S_T	Transducer area
ρ	Density

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

r	Receiver position or distance
r_0	Source position
$G_{ij,k}$	Spatial derivative of Green's tensor
σ_{jk}	Component of stress tensor
λ, μ	Lamé elastic coefficients for isotropic material
E	Young's elastic modulus for isotropic material
ν	Poisson's ratio for isotropic material
D	Bending stiffness for isotropic material (also used to denote distance between sensors)
d	Distance between AE source and sensor
x, y, z	Coordinate axes
u, v, w	Displacements along x, y, z coordinate axes
h	Plate thickness
F	Force
θ	Angle
v	Velocity for unspecified acoustic mode
v_l	Bulk longitudinal wave velocity
v_s	Bulk shear wave velocity
v_R	Rayleigh wave velocity
v_e	Extensional plate mode velocity
v_f	Flexural plate mode velocity
α	Attenuation coefficient (also used to denote ratio of Rayleigh velocity to bulk shear velocity)
κ	Ratio of longitudinal wave bulk velocity to shear wave bulk velocity
N_x, N_y, N_{xy}	Laminated plate theory (LPT) in plane stress components
M_x, M_y, M_{xy}	LPT bending moments
ε_x^0, ε_y^0, ε_{xy}^0	LPT midplane strains
κ_x, κ_y, κ_{xy}	LPT plate curvatures
A_{ij}	LPT in-plane stiffnesses for anisotropic material
D_{ij}	LPT bending stiffnesses for anisotropic material

REFERENCES

1. RK Miller, P McIntire, eds. Nondestructive Testing Handbook, Vol. 5, Acoustic Emission Testing. American Society for Nondestructive Testing, 1987, pp 2–6.
2. F Forster, E Schiel. Akustische Untersuchung der Bildung von Martensitnadeln (Acoustic Study of the Formation of Martensite Needles). Zeitschrift fur Metallkunde. 29(9):245–247, 1936.
3. W Mason, HJ McSkimin, W Shockley. Ultrasonic Observation of Twinning in Tin. Phys Rev 73(10):1213–1214, 1948.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

4. D Millard. Twinning in Single Crystals of Cadmium. PhD dissertation, University of Bristol, England, 1950.
5. J Kaiser. Untersuchungen uber das Auftreten Geraushen beim Zugversuch (An Investigation into the Occurrence of Noises in Tensile Tests). PhD dissertation, Technische Hochschule, Munich Germany, 1950.
6. T Drouillard. Acoustic Emission: A Bibliography with Abstracts. New York: Plenum Data Co., 1979.
7. RG Liptai, DO Harris, RB Engle, CA Tatro. Acoustic Emission Techniques in Materials Research. Int J Nondestructive Testing 3(3):215–275, 1971.
8. DG Eitzen, FR Breckenridge, RB Clough, ER Fuller, NN Hsu, JA Simmons. Fundamental Developments for Quantitative Acoustic Emission Measurements. Interim Report for the Electric Power Research Institute NP-2089, Research Project 608-1, 1981.
9. MR Gorman, WH Prosser. AE Source Orientation by Plate Wave Analysis. J Acoust Emission 9(4):283–288, 1991.
10. SM Ziola, MR Gorman. Source Location in Thin Plates Using Cross-Correlation. J Acoust Soc Am 90(5):2551–2556, 1991.
11. JL Walker, E Hill, GL Workman, SS Russell. A Neural Network/Acoustic Emission Analysis of Impact Damaged Graphite/Epoxy Pressure Vessels. Proceedings of the ASNT 1995 Spring Conference, Las Vegas, 1995, pp 106–108.
12. K Aki, PG Richards. Quantitative Seismology: Theory and Methods. New York: W. H. Freemen and Company, 1980.
13. CB Scruby. Quantitative Acoustic Emission Techniques. Nondestructive Testing, Vol. 8. London: Academic Press, 1985, pp 158–159.
14. KY Kim, W Sachse. Characteristics of an Acoustic Emission Source from a Thermal Crack in Glass. Int J Fracture 31:211–231, 1986.
15. M Ohtsu. Source Mechanism and Waveform Analysis of Acoustic Emission in Concrete. J Acoust Emission 1(2):103–112, 1982.
16. WH Prosser, KE Jackson, S Kellas, BT Smith, J McKeon, A Friedman. Advanced, Waveform Based Acoustic Emission Detection of Matrix Cracking in Composites. Mat Eval 53(9):1052–1058, 1995.
17. JW Whitney. Structural Analysis of Laminated Anisotropic Plates. Lancaster, Pennsylvania: Technomic Publishing Co., 1987.
18. SW Tsai, HT Hahn. Introduction to Composite Materials. Lancaster, Pennsylvania: Technomic Publishing Co., 1980.
19. WH Prosser. Applications of Advanced Waveform-Based AE Techniques for Testing Composite Materials. Proceedings of the SPIE Conference on Nondestructive Evaluation of Materials and Composites, Scottsdale, 1996, pp 146–153.
20. WH Prosser. The Propagation Characteristics of the Plate Modes of Acoustic Emission Waves in Thin Aluminum Plates and Thin Graphite/Epoxy Composite Plates and Tubes. NASA Technical Memorandum 104187, 1991.
21. TM Proctor Jr. An Improved Piezoelectric Acoustic Emission Transducer. J Acoust Soc Am 71:1163–1168, 1982.
22. TJ Fowler. Acoustic Emission Testing of Fiber Reinforced Plastics. Preprint 3092, American Society of Civil Engineers, 1977.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

23. WH Prosser. Waveform Analysis of AE from Composites. Proceedings of the Sixth International Symposium on Acoustic Emission from Composite Materials (AECM-6), San Antonio, 1998, pp 61–70.
24. MA Hamstad, J Gary, A. O'Gallagher. Far-Field Acoustic Emission Waves by Three-Dimensional Finite Element Modeling of Pencil-Lead Breaks on a Thick Plate. J Acoust Emission 14(2):103–114, 1996.
25. WH Prosser, MA Hamstad, J Gary, A O'Gallagher. Reflections of AE Waves in Finite Plates: Finite Element Modeling and Experimental Measurements. J Acoust Emission, 17(1–2):37–47, 1999.
26. KS Downs, MA Hamstad. Acoustic Emission from Depressurization to Detect/Evaluate Significance of Impact Damage to Graphite/Epoxy Pressure Vessels. J Comp Mat 32(3):258–307, 1998.
27. KS Downs, MA Hamstad. Correlation of Acoustic Emission Felicity Ratios and Hold-Based Rate Moments with Burst Strengths of Spherical Graphite/Epoxy Pressure Vessels. J Acoust Emission 13(3–4):45–55, 1995.
28. S May, I Searle, S Ziola, N Phelps. F-16 Structural Crack Monitoring by Means of Broadband Acoustic Emission. Proceedings of the First Joint DOD/FAA/NASA Conference on Aging Aircraft, Volume II. Ogden, Utah, July 8–10, 1997, pp 1469–1492.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

7

Radiology

Harry E. Martz, Jr., and Clinton M. Logan*
University of California, Livermore, California

Peter J. Shull
The Pennsylvania State University, Altoona, Pennsylvania

Radiographic techniques, and specifically X-rays, are historically and currently one of the most common methods of NDE. Radiography is one of the few NDE methods that can examine the interior of the object and the only NDE method that works on all materials. In addition, radiography is a mature method that has been in use for over a hundred years.

This chapter is an overview of radiology both in the more common two-dimensional (2-D) radiography method and the burgeoning fields of computed tomography (CT) and other three-dimensional (3-D) methods. The basic theory behind the absorption and creation of X-rays will be covered as well as the general equipment and procedures used in 2-D radiography and 3-D CT. Finally,

*Retired
Depth Informing Techniques (other than CT) in Sec. 7.4.5 submitted by Harold Berger and Tomas S. Jones, Industrial Quality, Inc., Gaithersburg, Maryland. Radiation Safety (Sec. 7.6.1) submitted by Frank A. Iddings, Professor Emeritus, Louisiana State University, Baton Rouge, Louisiana.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

selected applications from both 2-D radiography and CT imaging will be discussed.

7.1 INTRODUCTION TO X-RAYS

X-rays are electromagnetic waves with an energy that is 100–100,000 times higher than visible light (Table 7.1) (1). Like visible light, X-rays can be scattered and absorbed by materials; however, since X-rays have a much higher energy, these rays tend to pass through materials that block visible light. Therefore, X-rays can "see through" objects that are opaque to the naked eye. The amount of X-rays that pass through a material is dependent on the elemental composition, density, and thickness of the material, and the energy and amount of X-rays. Thus, X-rays are often very useful in determining change in material density, thickness, or composition. Also, the energy of an X-ray produced from a material depends on the atoms of the material, and can thus be used to determine its anatomical makeup.

X-rays can be used to yield data in one-dimension (1-D) called a gauge measurement, two-dimensions (2-D), referred to as a radiograph or projection radiograph, two-dimensional slices (2-D slice), or all three (3-D) spatial dimensions, and is referred to as computed tomography. A single measurement of the attenuation along a well-*collimated* (straight) beam path is usually called *X-ray gauging*, shown schematically in Figure 7.1. This method provides a single data point. The reader may wonder what possible use could be made of this method, but for situations where other data is available this type of measurement can complete the characterization of a material. For example, if the composition and

TABLE 7.1 Approximate Values of Wavelength, Frequency, and Energy of Various Portions of the Electromagnetic Spectrum

Region	Wavelength (Angstroms)	Wavelength (centimeters)	Frequency (Hz)	Energy (eV)
Radio	$> 10^9$	> 10	$< 3 \times 10^9$	$< 10^{-5}$
Microwave	10^9–10^6	10–0.01	3×10^9–3×10^{12}	10^{-5}–0.01
Infrared	10^6–7000	0.01–7×10^{-5}	3×10^{12}–4.3×10^{14}	0.01–2
Visible	7000–4000	7×10^{-5}–4×10^{-5}	4.3×10^{14}–7.5×10^{14}	2–3
Ultraviolet	4000–10	4×10^{-5}–10^{-7}	7.5×10^{14}–3×10^{17}	3–10^3
X- and gamma rays	10–< 0.1	10^{-7}–$< 10^{-9}$	3×10^{17}–$> 3 \times 10^{19}$	10^3–$> 10^5$

Although continuous, the spectrum is commonly separated into categories based on origin or function of the radiation.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

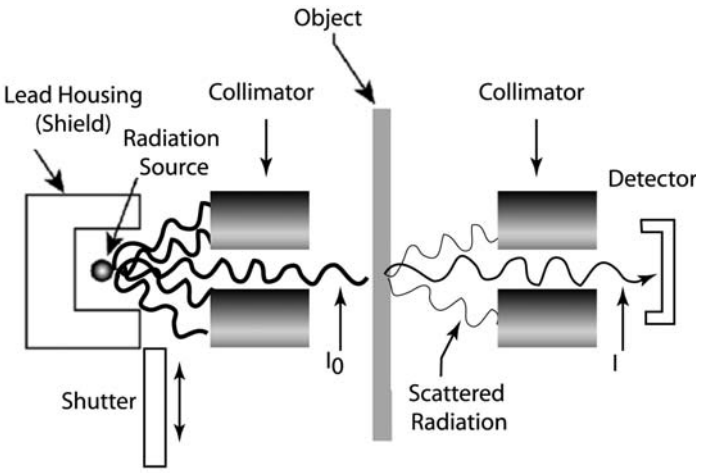


FIGURE 7.1 Schematic of a radiation gauge set up. This method is for example used to measure sheet steel thickness. I_0 refers to the incident intensity of the beam while I is the transmitted intensity.

thickness of the material are known, a single measurement of attenuation can define the average mass density along the radiation path (2, 3). If, on the other hand, both the composition and the mass density of the material are known, a measurement of attenuation can be used as a non-contact method to define the material thickness (4, 5).

Two-dimensional *projection radiography* (Figure 7.2) is the most widely employed imaging method in both medicine and NDE, usually with film as the recording and storage medium. In fact 2-D radiography is so common that it is often referred to as simply *radiography*. Nearly everyone has personal experience with a dental X-ray such as shown in Figure 7.3, chest X-ray, or a radiograph of a broken bone. This is an especially powerful technique for bone imaging because bone attenuates X-rays much more strongly than does soft tissue. Note that X-rays projected onto film show bright spots in the areas where the X-ray had high absorption (e.g., bone and metal), and dark spots where the X-ray had low absorption (e.g., soft tissue). Therefore, 2-D X-rays are really just shadows of the objects that they are imaging. Radiographs are a negative image just like a photographic negative produced from an everyday camera.

With the advent of computers, radiographic methods have been developed to image all three dimensions of an object—*computer tomography* (CT). Recall that a projection radiograph is a shadow or projection of the image where only the two lateral dimensions are preserved and the depth information is lost. (The depth direction is defined to be along the direction of the radiation travel.) Three-

Copyright © Marcel Dekker, Inc. All rights reserved.

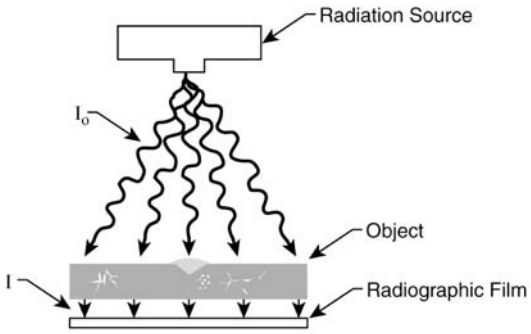


FIGURE 7.2 Schematic drawing of the geometry for industrial film radiographic imaging. This method is used for example to radiograph two plates welded together.

dimensional imaging is created by acquiring a large number (hundreds to thousands) of projection radiographs at different angular orientations to the object. Computer analysis correlates all projection images to create the depth dimension in addition to the two lateral dimensions. While a three-dimensional

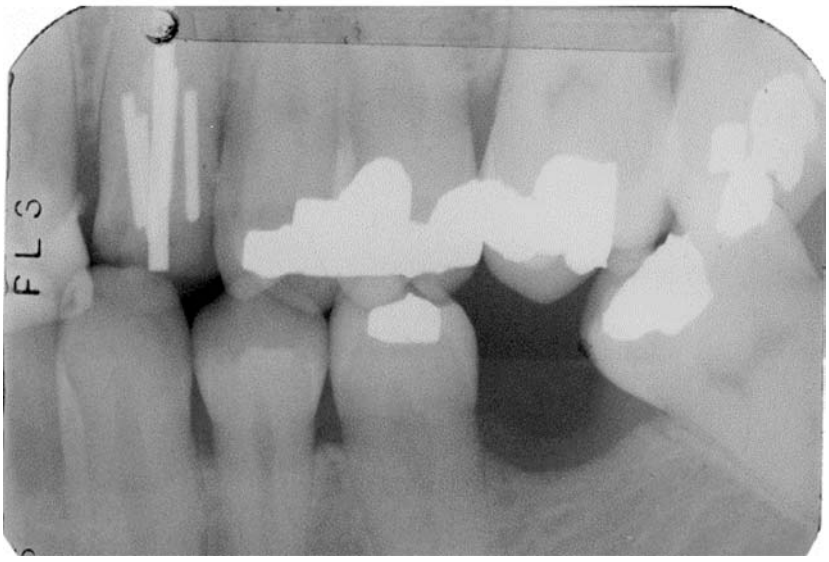


FIGURE 7.3 Digitized film radiograph of a person's teeth. Note the white areas are high attenuating metal fillings and pins for a root canal. Cavities show up as dark spots within the teeth. (Courtesy of Dr. Takemoto, Livermore, CA.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

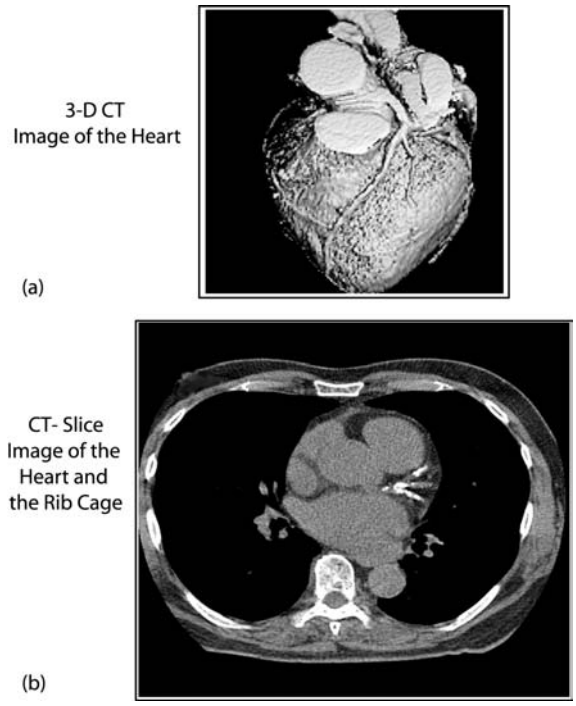


FIGURE 7.4 CAT scan of a human heart (a) a full 3-D rendering of the heart and (b) a scan slice of the heart and rib cage. (Courtesy of Imatron, Inc.)

image can be created, the most common CT image is the *2-D slice*. A CT slice is a two-dimensional image that contains the depth dimension and both lateral dimensions. Stacked CT slices create the 3-D image. In the medical industry CAT (computer aided tomography)* scans are usually viewed as slices (Figure 7.4).

7.1.1 History of X-Rays

In 1785 Benjamin Franklin was a diplomat in England and visited a Welsh scientist named William Morgan. Morgan was studying electrical discharges in a vacuum when mercury was boiled and found that the color of light was a function of its energy, and that at a high energy the light became invisible, "according to

*CAT can also stand for Computer Axial Tomography, where axial refers to the fact that the scan is around the center axis of the object. In general, tomography of humans are called CAT scans, while tomography of inanimate objects are called CT scans.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

the length of time during which the mercury was boiled, the 'electric' light turned green, then purple, then a beautiful violet ... and then the light became invisible." This is the first recorded production of X-rays, but the "discovery" of X-rays would require an additional 100 years (6).

In the mid-1850s many scientists around the world were examining a new device: a low-pressure glass tube with a high voltage filament attached (Figure 7.5) eventually called a *cathode ray tube*. Scientists soon found that rays were coming from the filament and creating light when it hit the glass tube. These rays were called "*cathode rays*" and were eventually discovered to be electrons. What the scientists didn't know was that when the electrons hit the glass surface they were creating X-rays as well as visible light. In fact, several scientists complained that their photographic film was slightly exposed (by X-rays) while in a drawer in a laboratory with cathode tubes and tried to get their money back (7)!

On November 8, 1895, Wilheim Röntgen, a relatively obscure professor of physics at Julius Maximilian University in Würzburg, Germany, completed the final test on his year long examination of a cathode ray tube. While the tube was covered with a heavy black cardboard he noticed that a florescent plate outside the cardboard glimmered slightly. These rays were not cathode rays, so as they were unknown, he used the mathematical letter for unknown, *X*, and called them *X-rays*. Immediately he began investigating the penetrating nature of the rays through metal and glass. He also placed his wife's hand in front of the plate and found to his amazement that he could distinguish between her bones, her flesh, and her metal ring. After two months of intensive study, Röntgen published "A

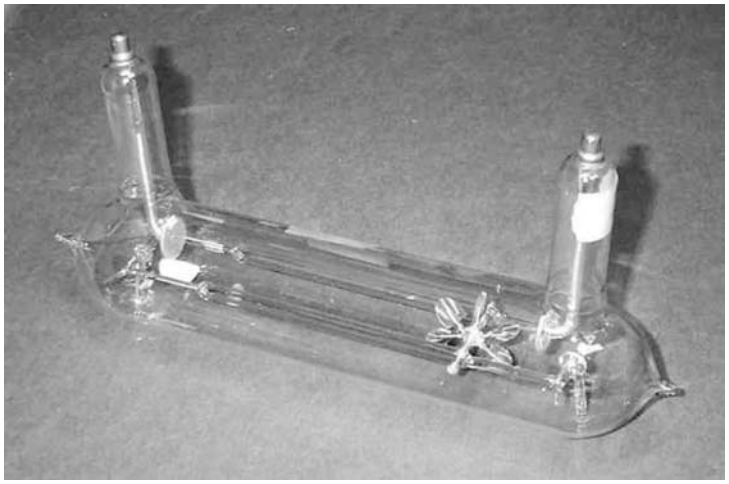


FIGURE 7.5 Cathode ray tube. (Courtesy of Oak Ridge Associated Universities.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

New Kind of Ray; Preliminary Communication," which was an immediate international sensation (8).

The impact of Röntgen's discovery is hard to imagine. Röntgen was showered with innumerable awards and honors, including the first Nobel Prize in Physics in 1901, as well as having several city streets changed to Röntgen (9). For a long time X-rays were called *Röntgen rays* (as they still are in Germany), X-ray pictures were referred to as Röntgenograms, and X-ray technicians as röntgenographers—despite the inherent difficulty in pronunciation. Newspapers around the world described this amazing discovery on the front page often with drawings or copies of the quickly famous X-ray of Mrs. Röntgen's hand. A typical report from the *London Standard* on January 7, 1896 ended with the remark that "there is no joke or humbug in the matter. It is a serious discovery by a serious German Professor" (10). In the first year of his discovery alone, several books and nearly one thousand scientific papers were written on the subject (11). Despite all of his fame, Röntgen did not profit from his findings. He never patented his discovery, he gave the money from his Nobel prize to his university, and died nearly penniless from the inflation after World War I.

The practical applications of X-rays, especially for medical uses, were immediately appreciated. Within months of Röntgen's discovery, X-rays had been applied on countless bodies and even used against breast cancer (12). In 1898, about two years after its discovery, a British surgeon carried out the first X-ray radiography during a battle (13). The radiography was performed in a mud hut with electricity powered by a bicycle. Soon doctors started using X-rays as a magic pill for everything from acne, leukemia, and epilepsy to "curing" lesbianism! Figure 7.6 displays an advertisement and a set of chairs from the London Hospital from 1905 that were used to treat ringworm. In the first year after Röntgen's discovery, 1896, there are even reports of X-rays being used in industrial NDE (14).

X-rays were not only used for medical uses. Portrait studios, traveling "amusement machines" where you could see the bones of your hand for a nickel, and X-ray clubs were popular overnight. Thomas Edison thought that X-ray machines would soon be in every home. He envisioned a time when people could X-ray their own injury and then send their pictures to their doctor. Companies advertised their X-ray machines as being "so easy a child could do it" (15). Every quality shoe store in the United States contained a X-ray machine for a "scientific fit" (Figure 7.7).

Today it is easy to forget what an amazing and frightening discovery X-rays were in 1896. At the time the human body, especially the female body, was rarely examined. Doctors would often examine the patients purely by touch, even for births. To be able to see inside a person was considered shocking and occasionally obscene as the following, admittedly horrible, poem published in April of 1896 demonstrates:

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

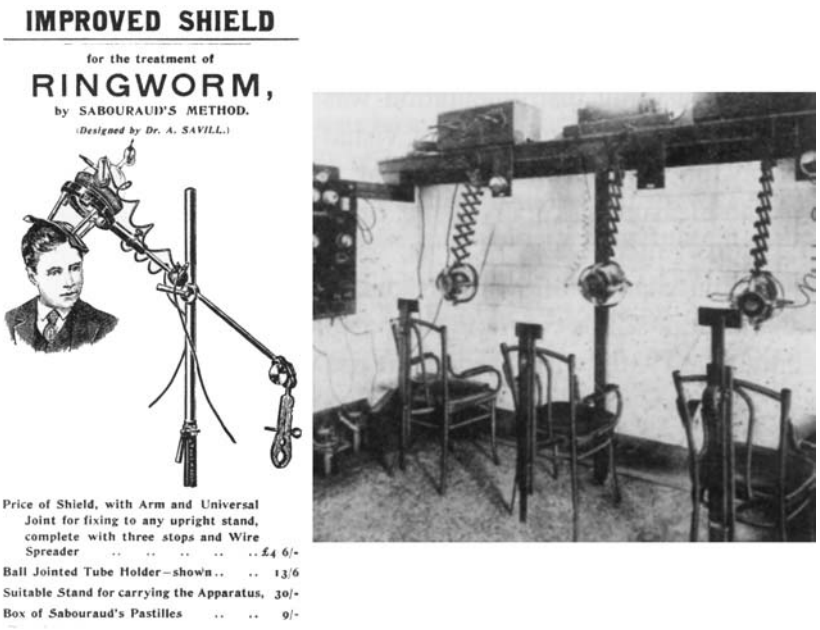


FIGURE 7.6 Advertisement for an X-ray machine to cure ringworm and an actual hospital X-ray room implementing such machines. (Reproduced with permission of American College of Radiology, Reston, Virginia.)

X-actly So!

The Roentgen Rays, the Roentgen Rays,
What is this craze?
The town's ablaze
With the new phase
Of X-ray's ways.

I'm full of daze
Shock and amaze;
For nowadays
I hear they'll gaze
Thro' cloak and gown—and even stays
These naughty, naughty Röntgen Rays [16]

Viewing the human body shape is still an issue today. As a consequence, full-body X-ray scans to reveal bombs and weapons concealed under an airline passenger's clothing is considered an unacceptable invasion of privacy (17).

Copyright © Marcel Dekker, Inc. All rights reserved.



FIGURE 7.7 Shoe fitting X-ray machine commonly used in the 1950s. (Courtesy of Oak Ridge Associated Universities.)

An interesting side note to the history of X-rays is the role of women. Women were the first to be examined with X-rays: a woman's hand was the first to be X-rayed, and breast cancer was the first disease to be cured with radiography. But women were not only the recipients of X-rays but also the technicians. Perhaps the novelty or lack of knowledge about X-ray radiography caused many doctors to consider it less threatening than general practice. Or perhaps it was because lay people could become X-ray technicians without formal male-dominated schooling. For whatever reason, many women were administrating X-rays in this emerging field far before the majority of nurses were female (18).

Although the popularity of X-rays was astounding, the public and especially the scientific community were much more reluctant to comprehend the dangers of X-rays. Very soon after X-rays were discovered, X-ray related injuries were recorded. However, the scientific community tended to ignore these reactions as unrelated or "X-ray allergies." Added to these facts was the common practice of

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016



FIGURE 7.8 A women X-ray technician testing the strength of the radiation source before applying them to the patient. (Reproduced with permission of American College of Radiology, Reston, Virginia.)

X-ray technicians to test the intensity of the machine on their body before examining a patient (Figure 7.8). Soon technicians and radiography scientists were dying in extraordinary numbers. Only when Edison's assistant Clarence Dally died in 1904 after a very public and detailed documentation of his burns, multiple amputations and lymph node aberrations did the majority of scientists began to believe that this new ray could be dangerous (19, 20).

Despite this knowledge of the hazards of X-rays, in the 1910s, 1920s, and 1930s entrepreneurs marketed radiation as containing miraculous healing powers. Radiation spas were common. Figure 7.9 shows some of the devices such as radioactive water crocks for drinking water, radioactive pads to be placed under your pillow, and radioactive cigarette holder to inhale radiation into the lungs with the smoke. In each of these "healing" devices a small amount of radioactive material is placed in the instrument. Radioactive beauty creams, toothpaste (radon was thought to fight dental decay and improve the digestion), earplugs, chocolate bars, soap, suppositories, and even contraceptives were marketed. The close of this era is marked by the radiation related death of Eben Byers, the well-known Pittsburgh industrialist, U.S. amateur golf champion, and a man-about-town. He was so convinced of the healing powers of Radithor—a liquid containing 2μCi of radium—that he averaged three bottles a day, until he died of radium poisoning in April 1932 (21).

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016



FIGURE 7.9 Devices used to expose the body to the ''healing'' effects of radiation: (a) crock used to dispense radioactive drinking water, (b) radioactive pad for sleeping, and (c) radioactive cigarette holder used to inhale radioactive smoke. (Courtesy of Oak Ridge Associated Universities.)

Although the accuracy and safety for radiography has improved considerably in the last century, the basic principles of 2-D X-rays are the same as when Mrs. Röntgen placed her hand in front of the fluorescent plate in 1895.

The advent of computers allowed for a completely new method of X-ray imaging. In 1972, a British engineer named Godfrey Hounsfield of EMI Laboratories, and a South African physicist Allan Cormack of Tuffs University simultaneously and independently invented CT scanning (22). Hounsfield was knighted and Hounsfield and Cormack were given the Nobel prize for their work in 1979. By the 1980s CAT scans became a familiar site in hospitals. Faster computers and simpler operating systems have caused CT scans to be used to study systems for many NDE and medical purposes.

7.1.2 Advantages and Disadvantages

The radiation imaging method of NDE enjoys an advantage over many other NDE methods in that it is inherently pictorial and interpretation is to some extent intuitive (Table 7.2). Analyzing and interpreting the images requires skill and experience but the casual user of radiation imaging services can easily recognize the item being imaged and can often recognize discontinuities without expert interpretation. For example, in an X-ray radiograph's image of a hand (Figure

Copyright © Marcel Dekker, Inc. All rights reserved.

TABLE 7.2 Advantages and Disadvantages of Radiographic NDE

Advantages	Disadvantages
Visual	Dangerous
Internal detection	Expensive
Can give material, density and/or thickness measurements	Imaging artifacts can make analysis difficult
Can examine almost all types of materials	Requires highly skilled operator
Rapid area inspection	Cannot detect closed cracks

7.10) you do not have to be an expert radiographer to distinguish hard tissue (bone) from soft tissue (flesh). This simple example reveals the potential of the radiographic technique to inspect or characterize objects (from humans to engine blocks) for internal details. Also, X-ray NDE is not as limited to the type of material it can study, unlike other NDE methods.

Radiation methods are suitable for sensing changes in elemental composition and/or mass or volume density and/or thickness. They are often applied to confirm the presence or configuration of internal features and components. It is especially applicable to finding voids, inclusions and open cracks and is often the method of choice for verification of internal assembly details. Castings are susceptible to internal porosity and cracks. Figure 7.11 shows an example of porosity detection in cast Mg tensile specimens.

Computed tomography excels where information is needed in three spatial dimensions, such as the hand and reinforced concrete columns (Figure 7.12). X-ray CT have been applied to objects ranging in size from 5×10^{-6} m to 2 m (23, 24).

Despite their advantages, radiation methods create serious safety concerns. Not only is the radiation dangerous, but the high voltage needed to generate most X-rays can be dangerous as well as the difficulty in using heavy shielding materials. Also, radiography is limited in utility for detecting cracks. For a crack to affect the transmission of radiation, there must be an opening resulting in a local absence of material. A closed crack is not detectable using radiation. In addition, even when the crack has a finite opening, it will generally only be detectable in a radiograph at certain orientations. Ideally the long dimension of the crack is *parallel* to the direction of radiation travel, i.e., this maximizes the radiation-crack interaction. (This orientation is in contrast to ultrasound where a reflection from a crack *perpendicular* to the UT path maximizes the response.). Surface defects are often hard to distinguish with 2-D radiography. Finally, X-ray machines, especially CT systems, are very expensive and time consuming and

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

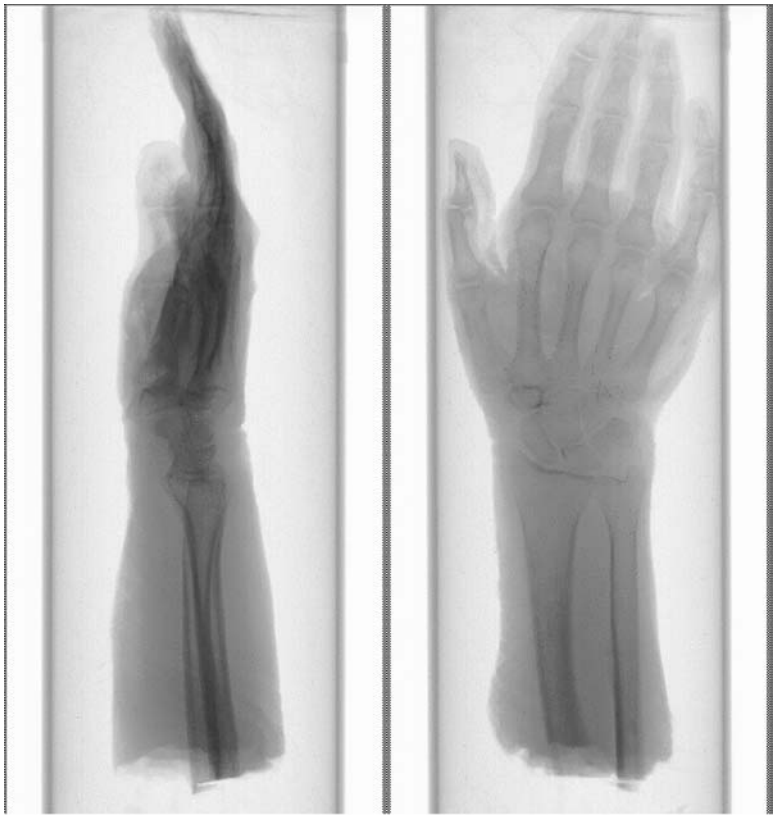


FIGURE 7.10 Digital radiographs of a women's hand at two angles 0° (left) and 90° (right). In these radiographs both the hard (bone, dark in the images) and soft (flesh, light in the images) tissues are observable. Note that these images are the inverse of film (negative) images.

require the use of highly trained safety conscious engineers, scientists or technicians.

7.2 RADIATION FUNDAMENTALS

7.2.1 Introduction

The most familiar form of radiographic energy for NDE is electromagnetic in nature, referred to as *photons*. Electromagnetic radiation can be described by the energy of the individual photons. The SI unit for energy is the joule, J, while the

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

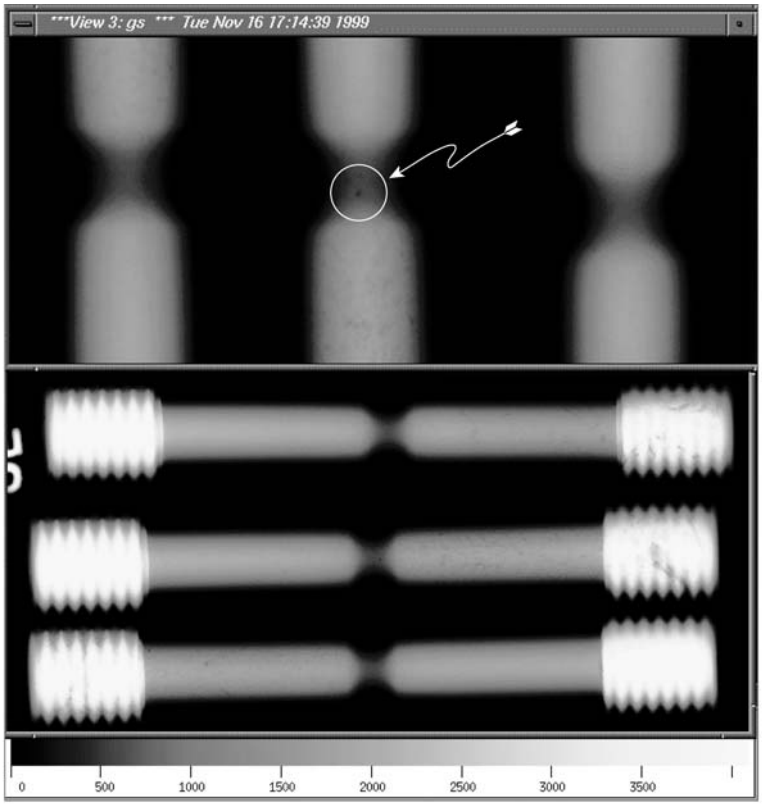


FIGURE 7.11 Digitized film radiographs of three magnesium, notched-tensile bars. The upper images are blown up to show more detail of the tensile bars given in the lower radiograph. In the upper central tensile bar near the middle the dark spot in the notch is a pore (dark area). In the lower radiograph the middle tensile bar has a crack (dark area) in the grip on the right.

traditional unit for radiation energy is the electron volt or eV. One eV is the kinetic energy imparted to an electron when the electron is accelerated across a one volt potential. The traditional unit eV ($1\,\text{eV} = 1.602 \times 10^{-19}\,\text{J}$) is most often used. Photons in the energy range from 1 keV to 10 MeV (10^{-9} –10^{-12} m wavelength) is the portion of the electromagnetic spectrum (Table 7.1) usually applied for NDE and are either γ-rays or X-rays depending on whether they originate within the nucleus or the orbital electrons of an atom, respectively. In addition to photons, electrons, protons, and especially neutrons are also used for NDE. All of the sources used for radiation-based NDE are directly or indirectly produced with

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

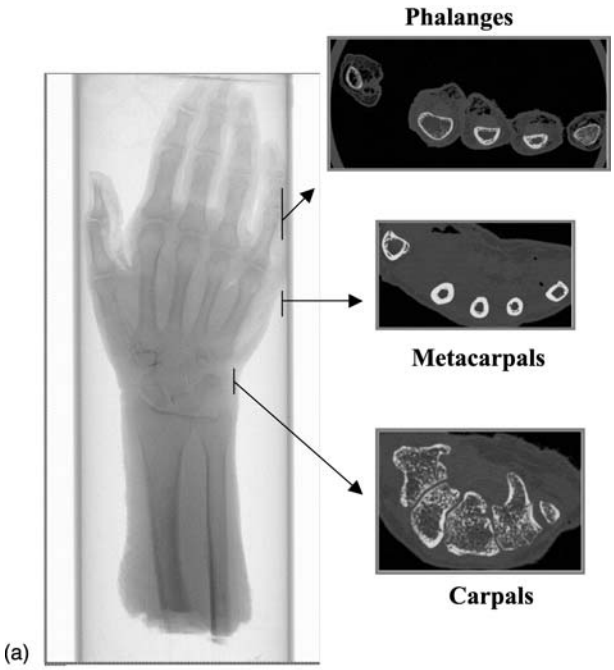


FIGURE 7.12 DR/CT images: (a) digital radiograph with some exemplar computed tomography cross-sectional images that reveal the true 3-D nature of computed tomography and (b) top left: Photograph of a large highway reinforced concrete column. Bottom left: Radiograph of the column shows internal details but are difficult to interpret due to the superposition of the concrete and rebar. Right: 3-D rendered image of tomograms of the column. The concrete is transparent and the rebar is opaque. When this data is rotated in a computer the internal details are revealed in all three dimensions.

machines. From these machines, radiation is either used directly or radioisotopes are produced which are then used as the source of radiation. Naturally occurring radioisotopes exist, but none are important NDE radiation sources.

Nearly all NDE radiation applications assume that the radiation travels in a straight line, ray, or path from the source to the detector attenuating as it passes through the test object. The varying degree of attenuation within the object creates the image at the detector. Each element has a unique atomic structure which gives rise to unique X-ray attenuation properties and unique X-ray emission characteristics when excited. Thus as the radiation passes through the object it will be attenuated according to the composition of the object. However,

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

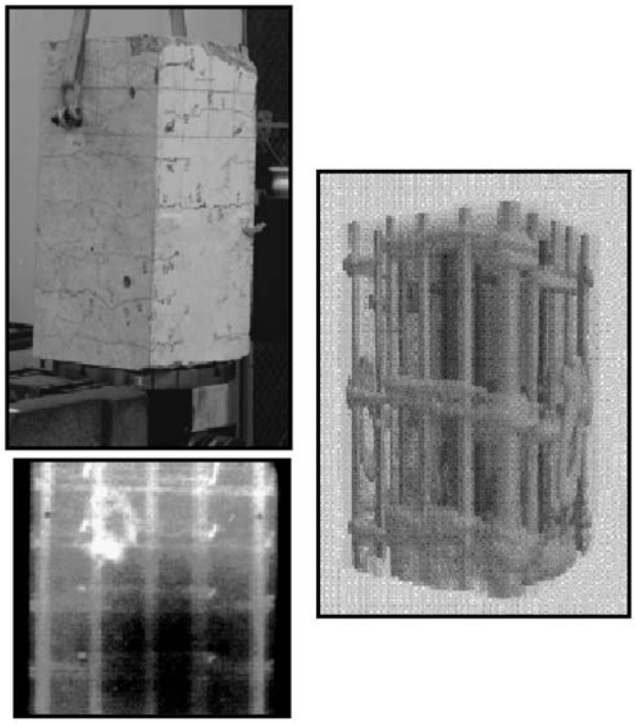


FIGURE 7.12 (*continued*)

this idealization of straight-line travel is not exactly true. The radiation is often scattered from its original straight-line path. For typical radiographic and tomographic imaging, if this scattered radiation reaches the detector, it will carry the wrong information (noise) and will degrade the image quality by reducing the signal-to-noise ratio (SNR). (Note that Monte Carlo based computer programs can account for scattering statistically and improve the SNR, but this is rarely done.)

Photons, the most commonly used radiation, are attenuated as they travel though matter via five distinct processes, *Rayleigh* (also called coherent or elastic) scattering, *photoelectric* effect, *Compton* scattering, *pair production*, and *photo-disintegration*. Only the photoelectric effect causes a photon to disappear with no significant re-radiation of energy (*secondary radiation*). The probability of these processes occurring over a given photon path length depends on the energy of the photon, the atomic number(s) of the matter or object, the density and thickness of the matter and nothing else, so long as the matter is not a plasma*.

*Plasmas—a completely ionized gas, composed of positive ions and free electrons—have unique radiation properties that will not be addressed here.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

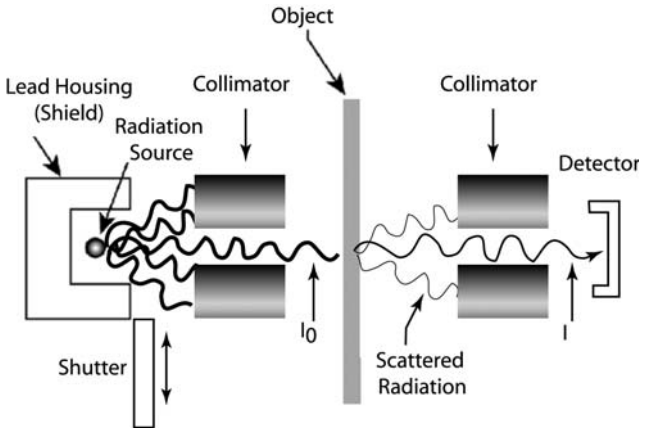


FIGURE 7.13 Gauging—one-dimensional radiation measurement. The beam is commonly referred to as a pencil beam.

Gauging is a one-dimensional radiation measurement along a single line called a ray. The source and detector are usually fitted with collimators that direct the radiation in a straight-line path and minimize the effect of scattered radiation (Figure 7.13). The object is placed between source and detector so as to determine attenuation through the object.

Radiography is a term used to describe a two-dimensional image of a three-dimensional object made with penetrating radiation. It is a type of shadowgraph made with photons or particles that partially penetrate the object under study. Radiographs superimpose information from along the entire path of the radiation onto the detector. Thus radiographs are sometimes called *projection radiographs*. In the usual case, the radiation source is either small with respect to the image or the source is configured such that the radiation travels parallel before passing through the object. If a radiograph is in digital form, the individual elements of the detector are called pixels (picture elements).

CT is a method of recovering a three-dimensional representation of a three-dimensional object. Simply, many projection radiographs (or sometimes just called projections) are taken of the object at different angles. Using computer algorithms, the three-dimensional object can be reconstructed from these radiographs. Although the entire three-dimensional object can be reconstructed from a 2-D radiograph, it is common that only a *slice* or cross-section is reconstructed for viewing. (The slice is in the plane of the radiation travel in contrast to a projection radiograph where the image is perpendicular to the direction of radiation travel.) Unlike a projection radiograph, the slice does not superimpose information but allows the identification of depth information. Often thousands of

Copyright © Marcel Dekker, Inc. All rights reserved.

digital radiographs are required in order to have sufficient data to achieve high fidelity "reconstructions" of an object. The result of CT is a data set with three spatial dimensions where each volume element is called a *voxel* and each voxel is assigned a value of radiation attenuation. (These voxels are the spatial resolution equivalent to pixels in two dimensions.)

Some specialized techniques account for scatters and secondary radiation transport in radiography using Monte Carlo methods, which are very computer intensive. These methods have until recently required large mainframe computers. They have now entered the realm of desktop and laptop computers.

7.2.2 Fundamental Sources of Radiation

All radiation has its origins in the atomic structure of the atom. The radiation may be a result of the energy released as electrons decay from one energy level to another in the process of becoming more stable. Similarly, energy changes in the nucleus can cause radiation to be emitted from the atom. The specific type of energy change will determine what type of radiation is emitted such as α, β, neutron, γ, or X radiation. α and β are particles where the α particle is the helium nucleus (i.e., two protons and two neutrons and no electrons) and the β particle is an electron.

Electrons and Radioactivity

Electrons surround the nucleus, orbiting in distinct shells related to their energy level. Those close to the nucleus are lower energy and more tightly bound than those at distant orbits. In the classical atomic physics definitions most commonly used in NDE, the innermost shell is designated as the K shell. Moving outward from the nucleus, the designations are K, L, M, N, etc. Within each shell there are small energy differences that are denoted by subscripts.

When an orbiting electron is dislodged by incident radiation, i.e., a photon or a particle, the atom exists in an *excited state*. Shortly, the electrons will rearrange themselves to return the atom to the *ground state*. During this transition, energy can be liberated in the form of an X-ray photon with energy equal to the difference between the excited state and the ground state. This X-ray is called a *characteristic X-ray* because its energy is characteristic of the element undergoing the transition.

The energy of a characteristic X-ray depends on the energy difference between the electron shell with the vacancy and the electron shell donating the electron. An example illustrates the notation used for characteristic X-rays. If the vacancy exists (momentarily) in the K shell of an atom and it is filled from the L shell, the characteristic X-ray emitted is called a K_α X-ray. If the donor shell is the M shell, the emitted X-ray is called a K_β X-ray. A similar process then fills the vacancy left in one of the outer shells and another characteristic X-ray is emitted.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

This process occurs until the atom decays into its ground state. The K series X-rays are of most technological importance because they are of the highest energy and most penetrating. The highest energy K series X-ray is created when the vacancy is filled with a free (unbound) electron.

The K series X-rays increase in energy with increasing atomic number. The maximum for Fe is 7 keV while that for W is 70 keV. Tables of characteristic energies are widely available and are easily found on the Internet such as on this web page: http://xdb.lbl.gov/Section1/Sec_1-2.html.

While the emitting atom determines the energies of characteristic X-rays, the relative numbers of each depend on the details of the source of excitation. In addition, there is a competing process, emission of an Auger electron, in which the energy is used to eject an electron from the atom and no characteristic X-ray is emitted.

The Nucleus and Radioactivity

The nucleus of a hydrogen atom (one proton) is stable without a neutron. If we add two neutrons to the stable hydrogen atom, it becomes the unstable radioactive isotope tritium. An isotope is any of two or more atoms of a chemical element with the same atomic number (number of protons) but different in the number of neutrons. If a nucleus has too many or too few neutrons, it is energetically unstable and will decay to a more stable energy level. The radioactive decay process is accompanied by the emission of α, β, neutron, γ, or X-radiation in some combination.

Some *transuranic* nuclides will spontaneously fission at a rate that can produce a useful neutron source. Neutron sources can also be manufactured making use of secondary nuclear reactions. When an α-emitting radioisotope is mixed with Be, neutrons are produced by (α, n) reactions. Neutrons can also be produced by mixing an appropriate (high-energy) γ emitter with Be. These sources are not usually used for NDE because their γ emission is many times their neutron emission.

Since we have already defined X-rays as arising from orbiting electrons, an explanation is in order as to how *nuclear* decay can result in X-rays. There are two basic mechanisms. The nucleus can move toward stability by capturing an electron—called *electron capture*. A proton within the nucleus is thereby converted to a neutron. This results in a transformation of the element to a new element with the same atomic mass number, Z, as before, but a lower atomic number. This new element is missing an electron from its inner shell. As we describe in the previous section, X-rays are then produced as the vacancy is filled. The second mechanism for nuclear decay to produce a missing electron and consequently X-rays, is for α or β radiation from the nucleus to eject an orbiting electron as it moves through the material.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Radioactive decay is a probabilistic phenomenon. The disintegration rate depends linearly on the number of radioactive atoms present. If one starts with a fixed quantity of radioactive nuclei, the number remaining continuously decreases and the decay rate continually decreases. If the number of nuclei at time = 0 is N_0, then the number of nuclei, N_t, remaining at a later time, t, is given by

$$N_t = N_0 e^{-\lambda_r t} \tag{1}$$

where λ_r is the *decay constant*.

The *half-life* τ of a radioisotope is defined as the time required for the number of radioactive nuclei to be reduced by one half. Therefore

$$\tau = \frac{\ln 2}{\lambda_r} \tag{2}$$

The *activity* of a radioisotope source is its rate of decay. From Eq. (1),

$$\Lambda = \left|\frac{dN_t}{dt}\right| = |-\lambda_r N_t| \tag{3}$$

The historical unit of activity is the Curie (Ci), defined as 3.7×10^{10} disintegrations/second. While Ci is still in widespread use, the SI unit of activity is Becquerel (Bq), defined as 1.0 disintegration/second. Useful sources for radiography contain many MBq or even GBq.

Two cautionary notes are in order before leaving the subject of radioactivity. A source does not always produce a given type of radiation on every disintegration. Therefore, the source strength of desired radiation might be only a fraction of the activity. In addition, absorption and scattering takes place within the source itself, so that radiation escaping from the source may only be a fraction of the decay amount.

7.2.3 Photon Radiation in NDE

In the previous section we described the fundamental origins of photon radiation. Here we will present the physics of common photon radiation sources and the interaction of photon radiation with materials.

Physics of Photon Radiation Sources for NDE

Radioisotope sources can emit both γ- and X-rays, often from the same source. We will use ${}^{109}_{48}Cd_{61}$ as an illustrative example. Cadmium-109 decays 100% by electron capture (The nucleus and radioactivity in Section 7.2) and becomes ${}^{109}_{47}Ag_{62}$. Silver-109 is produced in an excited neutron state after electron capture. The excited ${}^{109}_{47}Ag_{62}$ nucleus decays to a ground or stable state by a neutron transition. This neutron transition emits a γ-ray with energy of 88 keV. The electron used in the electron capture leaves the silver atom in an excited electronic

Copyright © Marcel Dekker, Inc. All rights reserved.

state as well. The K shell electron vacancy (usually) is filled by one of the L shell electrons emitting a characteristic X-ray with energy of 22 or 25 keV. These energies are characteristic of Ag. So ^{109}Cd produces three *photon energy lines*—one γ-ray and two X-rays. (The term *line* refers to the emission of radiation at a discrete energy producing a spike, peak or line in a plot of intensity vs. energy—the radiation emission spectrum.) Other radioisotopes are available with γ-ray energies ranging in excess of 1 MeV. Some radioisotope examples are presented in Radioisotopic Sources (Section 7.3.2).

Electrically powered X-ray tubes are the most technologically important photon sources for radiography by a large margin. These tubes function (as described in more detail in Electrically Powered sources in Section 7.3.2) by impinging energetic (high velocity–high kinetic energy) electrons onto a target (also called the *anode*, because of its function within the tube). This process produces many different X-rays some of which are characteristic X-rays with energies determined by the anode element. The impinging electrons eject orbital electrons from the atoms of the anode material. These electron vacancies are promptly filled and X-rays are emitted. The most common anode material is tungsten, W, and its K-series X-rays are at energies of 58, 59, and 67 keV. The incident electrons must have an energy exceeding 70 keV in order to eject a K-shell electron from W, so K shell X-rays (lines) from W appear only when the electrons in a tube exceed 70 keV in kinetic energy.

There is a second process that takes place in X-ray tubes, which usually dominates X-ray production. Energetic electrons passing through matter are subject to deflections caused by their interaction with the electronic structure of the atoms. Each of these results in radiative energy loss in the form of *bremsstrahlung* (German for braking radiation). On a radiation emission spectrum plot bremsstrahlung populates a continuum of energies extending up to the energy of the incident electron. The continuum of bremsstrahlung is superimposed on the characteristic lines to form the output spectrum from X-ray tubes (Figure 7.14). The fraction of the incident electron energy converted into bremsstrahlung energy increases with electron energy and with the atomic number of the anode.

Only a small fraction of the X-ray energy produced in a tube anode reaches the object under examination. Absorption in the anode, the tube window and the flight path all contribute to modifying the spectrum. In addition, X-rays are produced over a small volume with nearly isotropic angular distribution. This spherical radiation pattern results in an X-ray *flux density* ϕ from a source with strength S (photons produced per unit time), that decreases as $1/r^2$:

$$\phi = \frac{S}{4\pi r^2} \tag{4}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

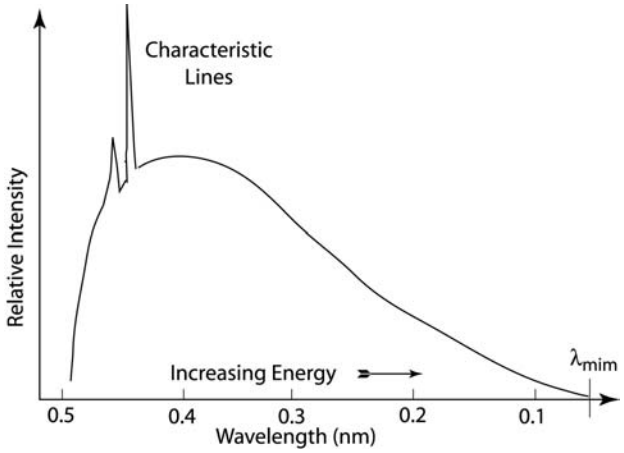


FIGURE 7.14 X-ray tube radiation spectrum—continuous radiation emission and characteristic "line" radiation spikes.

where r is the distance from the source. This $1/r^2$ decrease in radiation flux density with distance from the source is called the *inverse square law.* Tube source strength increases linearly with the electron current (number of electrons per unit time striking the target) and approximately as the square of the energy of the accelerated electrons. Another useful descriptor for NDE photon sources is their *brightness*. Brightness is defined as the available photon flux density divided by the square of the beam divergence. The brighter the source, the more photons are contained in a data set acquired for a fixed time.

The most intense sources of X-rays for NDE are synchrotrons, large machines with very high-energy electrons orbiting in a *storage ring*. When these electrons are deflected (angularly accelerated), either at bending magnets or in a device called a *wiggler*, they emit *broadband* photon radiation, some of which are in the 1–20 keV range. These photons travel down *beamlines* to experimental stations were the test object can be exposed to the high intensity radiation. Synchrotrons deliver a photon *brightness* that is 10^9 times that of an X-ray tube. Their brightness is so high that a monochrometer can be used to filter the broad band beam to pass only a narrow band of energy (equivalently, wavelength) for use in quantitative imaging and there is still enough X-ray brightness remaining to far exceed radioisotopes and X-ray tubes. The relative brightness of the most common radioisotope, X-ray tubes, and the Advanced Light Source (ALS) synchrotron at Lawrence Berkeley Laboratory are shown in Figure 7.15. Cost, scarcity, physical size and modest X-ray energy (typically less than 20 keV) limit the role of synchrotrons in conventional NDE, but the highest spatial resolution and most quantitative work in the field is done with these

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

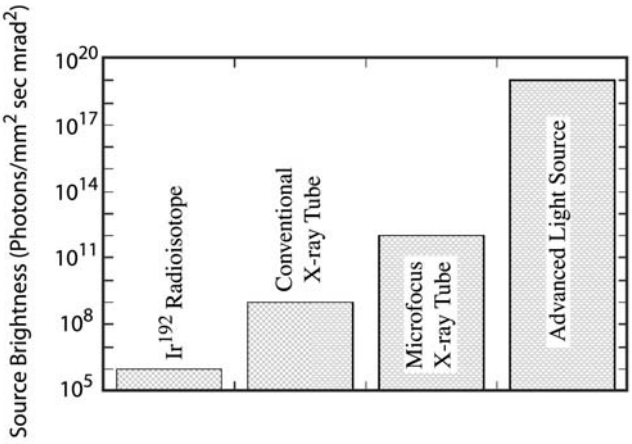


FIGURE 7.15 The relative brightness of four types of X- or γ-ray sources is shown on this semilog plot. Brightness for tubes and radioisotopes is proportional to the photon source strength divided by the square of the source diameter.

machines. The Advanced Light Source at Lawrence Berkeley Laboratory operates with electrons at 1.9 GeV. Its construction cost was ~US$$10^8$ in 1990 dollars. Operating staff is 175 persons. The electron storage ring has a circumference of 200 m.

Plasmas—completely ionized gasses composed of ions and free electrons—also emit radiation in the X-ray range. All objects emit electromagnetic radiation at energies (wavelength) according to their temperature. We are all familiar with the military applications of infrared (IR) imagers for detection and targeting of the enemy. Objects near room temperature emit in the IR region (~0.02 eV). Hotter objects emit in the visible (0.1 eV), thus the expression "red hot." Plasmas that emit in the lower end of the X-ray energy range can now be produced with commercial short-pulse, high-energy lasers. Research systems have produced plasma-based X-rays at several MeV. These sources are likely to be technologically important for NDE in the future. They are already becoming the sources of choice for X-ray lithography. X-ray lithography is a new process that is being developed to produce the small structures needed for small fast computer chips (25).

Photon Radiation Interaction with Materials

We will discuss five different mechanisms through which X- and γ-rays interact with matter:

1. Coherent scattering

Copyright © Marcel Dekker, Inc. All rights reserved.

2. Compton scattering
3. Photoelectric absorption
4. Pair production
5. Photodisintegration

All of these mechanisms act to attenuate the photon radiation emitted at the source and received by the detector. In general these attenuation mechanisms fall into two categories: *scattering* or *absorption*. (Often the attenuation is a combination of both mechanisms with one of the two dominating.) In scattering the direction of the radiation is deflected by the atoms of the material from its ideal straight-line path. Coherent scattered radiation does not lose energy, while Compton scattered radiation loses some of its energy in the process to absorption. In the absorption process, the incident radiation's energy is absorbed by the material. This absorbed energy may or may not be re-emitted as a secondary radiation. Typically the secondary radiation is emitted isotropically or omni-directionally and therefore very little reaches the detector. Note, ideally all scattered and secondary radiation would not reach the detector. The amount that does acts as noise and degrades the image quality.

Coherent scattering, also called Rayleigh scattering, is similar to two billiard balls colliding—elastic collision. Its most important attribute is that the photon survives (or more correctly is reborn) with its energy and phase unaltered. It simply moves in a new direction. This is depicted graphically in Figure 7.16a. For most NDE applications, Rayleigh scattering is either insignificant or a nuisance. It is seldom a large enough contributor to total attenuation to be important in gauging, radiography, or computed tomography. It is potentially a nuisance because the photon embarks on a new direction and can cause a response in the detector that violates our assumption that photons travel along straight lines from source to detector. Such a deflected photon is considered noise by the detection system.

Compton scattering takes place between a photon and an electron in the interacting material. The photon is deflected and transfers some energy to the electron as depicted in Figure 7.16b. When the angle of deflection is small, the energy lost is small. Some energy is always retained by the photon, even when the photon reverses direction. The relationship between energy transfer and scattering angle is derived from conservation of energy and momentum. The most detrimental attribute of Compton scattering for NDE is that the scattering is more likely to occur at small scattering angles, so the resulting photon is shifted slightly lower in energy and deflected through a small angle. This small angle deflection tendency increases with increasing incident photon energy. Compton scattered photons have high probability of reaching the detector in common NDE configurations. They confuse the image. In the least detrimental form, they are fairly uniformly distributed and simply cause an offset (uniform background

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

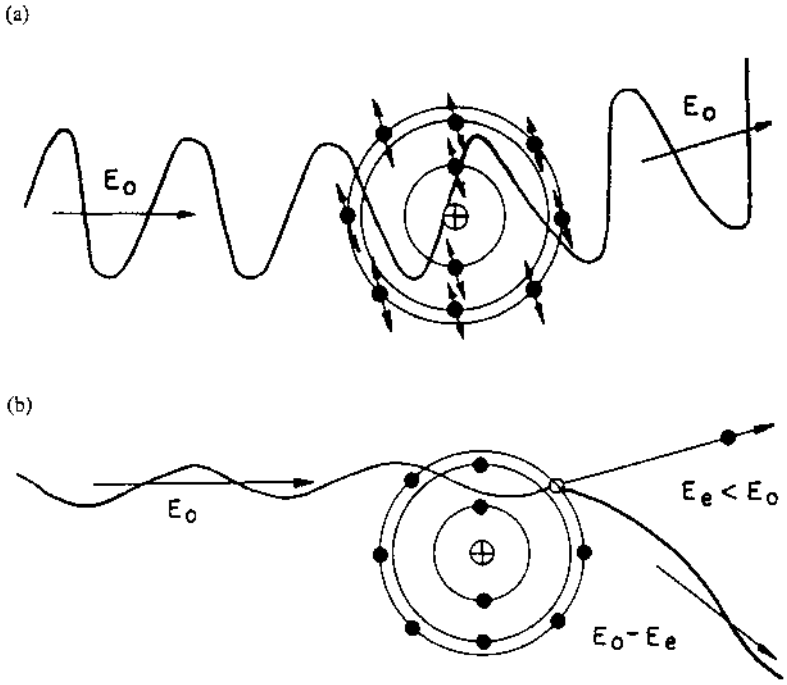


FIGURE 7.16 Pictorial of scattering interactions: (a) Schematic of coherent scattering of a photon in which *without loss of energy* the incident photon interacts with the ensemble of electrons in their atomic shells, producing a new photon of the same energy, but different direction. and (b) schematic of Compton scattering of a photon in which an incident photon ejects an electron (or β particle) and a lower-energy scattered photon moving in a new direction. (© 1986, The American Society for Nondestructive Testing, Inc. Reprinted with permission from the Nondestructive Testing Handbook, 2nd ed.: Vol. 3, Radiography and Radiation Testing.)

noise reducing the dynamic range of the detector). In the worst form, they are correlated with the transmitted photons (those that pass through the object in a straight line) and create both noise and blur (fuzzing of sharp features).

Photoelectric absorption is the process that is most nearly ideal from the perspective of the goal of NDE. The original photon disappears having transferred all of its energy to an electron in some shell of an atom, depicted schematically in Figure 7.17. This energy may move the electron to another shell or it may eject it from the atom. When ejected, the electron energy is the difference between the energy of the incoming photon E_0 and that of the binding

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

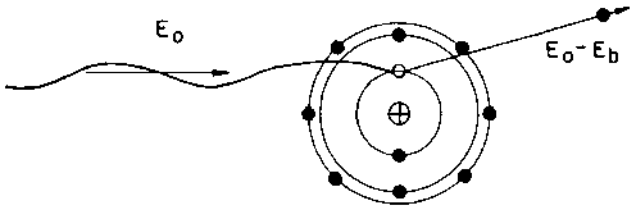


FIGURE 7.17 Schematic of photoelectric interaction between an incident photon and an orbital electron. (© 1986, The American Society for Nondestructive Testing, Inc. Reprinted with permission from the Nondestructive Testing Handbook, 2nd ed.: Vol. 3, Radiography and Radiation Testing.)

energy E_b of the particular electron in the atom. A photon (X- or γ-ray) cannot eject an electron until its energy exceeds the binding energy of the electron to the atom. Outer shell electrons, being more loosely bound, require less energy than inner shell electrons. A photoelectric interaction leaves the atom with a vacancy in one of its shells. This is filled essentially instantaneously with electrons from other shells and produces one or more characteristic X-rays as described in Electrons and Radioactivity Section 7.2.2. These characteristic X-rays can reach the detector, though they are not often significant for NDE. Two attributes help mitigate their negative effect (noise). They are emitted omnidirectionally, so few are directed toward the detector, and they seldom travel far.

If a photon exceeds 1.022 MeV (in energy units—$E = mc^2$—twice the rest mass of an electron) it becomes possible for *pair production* to occur. In pair production the photon produces an electron and a positron, vanishing in the process (Figure 7.18). The positron will subsequently slow down and annihilate

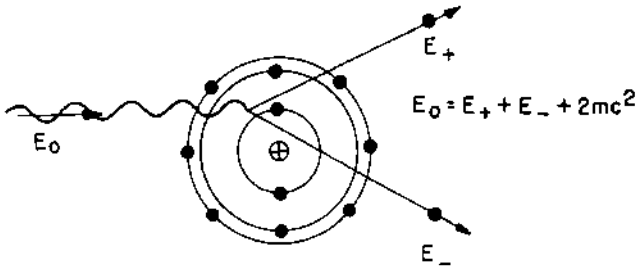


FIGURE 7.18 Schematic of pair production. An electron and positron are created from an incident photon whose energy exceeds 1.02 MeV. (© 1986. The American Society for Nondestructive Testing, Inc. Reprinted with permission from the Nondestructive Testing Handbook, 2nd ed.: Vol. 3, Radiography and Radiation Testing.)

Copyright © Marcel Dekker, Inc. All rights reserved.

with an electron, producing two *annihilation photons* each with an energy of 511 keV. In most instances, pair production is less disruptive to NDE than Compton scattering, because the annihilation photon is emitted omnidirectionally rather than preferentially in the forward direction, and therefore has lower probability of reaching the detector.

If a photon has sufficient energy, it becomes possible to cause nuclear reactions, called *photodisintegration*. These reactions are never significant contributors to attenuation, but may produce secondary neutrons and/or radioactivity. When using bremsstrahlung X-ray sources at *potentials* above 6 MV, the possibility of activating the equipment or object to become radioactive must always be considered.

We avoided offering rules of thumb about *Z*-dependence and energy-dependence of each of these processes because these rules have so many exceptions as to be generally useless. The most useful depiction of the relative importance of these separate processes for NDE is that of Evans (26) shown in Figure 7.19. Rayleigh scattering and photodisintegration are never dominant. Photoelectric dominates at low energy and low *Z*. Compton is dominant over most of the useful NDE energy and Z range. Pair production becomes significant at only the highest energies and the highest *Z*.

Narrow Beam or ''Good Geometry'' and the Attenuation Coefficients

In physics measurements, there exists a concept referred to as *good geometry* where any photon attenuated by the test material is directed away from the detector by geometry and collimation leaving only "straight-line" photons to strike the detector and indicate the degree of attenuation within the material. (A *collimator* is a highly absorbing material that blocks all radiation except where a small hole or slit is placed in the material.) Conceptually, good geometry can be thought of as a pencil (parallel) beam of photons and a small, well-collimated detector. In NDE literature, good geometry is called *narrow beam*. In good geometry, all interacting photons leave the beam and do not reach the detector. The probability of interaction is given by the sum of the probability of each of the potential interaction possibilities. This numerical descriptor of interaction probability is known by different names and is expressed in several units. In NDE, the most common is the *mass attenuation coefficient* μ/ρ often tabulated in cm^2/gm. The term μ is known as the *linear attenuation coefficient* and has units of inverse length, for example cm^{-1}; ρ is density with units of g/cm^3. It is equally correct and more common in physics circles to use *atomic cross section* (σ_{cs}). The most common unit of cross section is the barn (b), equal to 10^{-24} cm^2. So it is simply an exercise using Avogodro's number and gram molecular weights to convert from common NDE units to common physics units. A second concept common

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

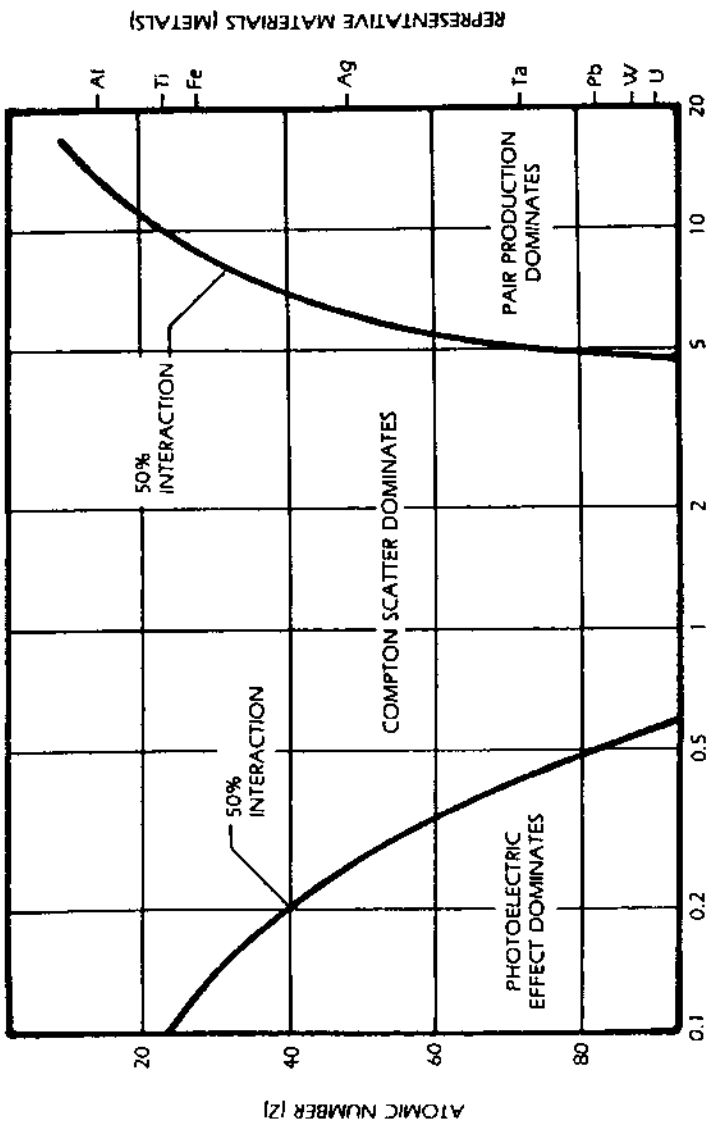


FIGURE 7.19 Photon attenuation between 100 keV and 20 MeV is dominated by one of three processes: photoelectric, Compton scatter, or pair production. The plot identifies the regions in atomic number and photon energy space where each of the three processes dominate (contribute $> 50\%$ of the photon interactions). (© 1986. The American Society for Nondestructive Testing, Inc. Reprinted with permission from the Nondestructive Testing Handbook, 2nd ed.: Vol. 3, Radiography and Radiation Testing.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

in physics, but less so in NDE is that of *mean-free-path*, *mfp*. Mean free path has units of length and is the inverse of linear absorption coefficient, μ.

Consider Pb at 125 keV. From (27),

$$\frac{\mu}{\rho} = 3.124\,\frac{\text{cm}^2}{\text{g}}$$

$$Z = 81$$

$$A = 204.37$$

$$\rho = 11.36\,\frac{\text{g}}{\text{cm}^3}$$

Conversion leads to

$$\mu = \left(\frac{\mu}{\rho}\right)\rho = 35.5\,\text{cm}^{-1}$$

$$\sigma_{cs} = \left(\frac{\mu}{\rho}\right)\left(10^{24}\,\frac{\text{b}}{\text{cm}^2}\right)\left(\frac{204.37\text{g}}{6.02 \times 10^{22}\,\text{atom}}\right) = 10,605\,\frac{\text{b}}{\text{atom}}$$

$$mfp = \frac{1}{\mu} = 0.028\,\text{cm}$$

For narrow beam geometry attenuation of a photon beam, the attenuation is defined by the integral of the linear attenuation coefficients of the material through which the beam passes. *Areal density* is line integral of volume density along a ray. For foil or plate of uniform density, it is simply the product of volume density and thickness. The most common unit of areal density in NDE is g/cm^2.

X- and γ-Ray Attenuation

The attenuation of single energy X- and γ-ray photons as they pass through an absorbing material is governed by an exponential relationship.

$$\frac{I}{I_0} = e^{-(\mu/\rho)\rho l} \tag{5}$$

where

- I is the number of photons reaching the detector in *good geometry* in presence of the test material.
- I_0 is the number of photons that would be detected without the test material.
- μ is the linear attenuation coefficient for a *homogeneous* material.
- ρ is the test material density.
- l is the thickness of the test material and represents the path of attenuation of the radiation—*ray path*.

Copyright © Marcel Dekker, Inc. All rights reserved.

This relationship is known as *Beer's law* for *homogeneous* material and single energy photons, and is often rewritten as

$$R(\text{raysum}) = -\ln\left[\frac{I}{I_0}\right] = \ln\left[\frac{I_0}{I}\right] = \mu l \tag{6}$$

So, if we wish to know the effectiveness of a 1-mm thick Pb sheet for attenuating 125-keV X- or γ-rays, we use the values given in the preceding example.

$$\frac{I}{I_0} = e^{-(3.124)(11.36)(0.10)} = 0.028$$

One millimeter of Pb allows 2.8% of photons with energy of 125 keV to pass through without interaction.

Mass attenuation coefficients are strong and complex functions of composition and energy. This is illustrated by Figure 7.20. Figure 7.20a gives mass attenuation for Pb as a function of energy. Starting at low energy we see high values that drop sharply as energy increases. Then we see a sharp jump upward at 88 keV. At 88 keV the photon energy is suddenly sufficient to eject a K-shell electron and the probability of a photoelectric interaction increases sharply. This is called an *absorption edge*. At 500 keV, Compton scattering is becoming dominant and the mass attenuation coefficient is slowly varying, showing a gentle minimum near 3.5 MeV. Above 5 MeV, pair production becomes significant and the mass attenuation coefficient rises slowly with energy. Figure 7.20b shows the behavior of mass attenuation coefficients as a function of atomic number for photons of 50 keV and 3.5 MeV. For the 50 keV photons, photoelectric absorption is the dominant attenuation mechanism for all but the lightest elements. We see μ/ρ increase with increasing Z, until $Z = 64$ where the value drops sharply. This decrease occurs because at this atomic number, 50 keV is no longer sufficient to eject a K-shell electron. The shape of the curve for 3.5-MeV photons is flatter, because at this energy Compton scattering is the dominant process. This interaction depends mainly on electron density, which varies little with Z.

Photon attenuation depends strongly on which elements are present in a material, but is essentially independent of their chemical state. (Of course, we need to account for a change in density to compute the attenuation.) This means that attenuation coefficients for chemical compounds, metallic alloys, and composite materials are all computed from attenuation coefficients of the elemental constituents in the same manner. If one has available the mass attenuation coefficient of the elements, it is only necessary to determine the mass fraction of each element present to compute the mass attenuation coefficient of a multi-element material. This is the great convenience of using this form for the attenuation coefficients.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

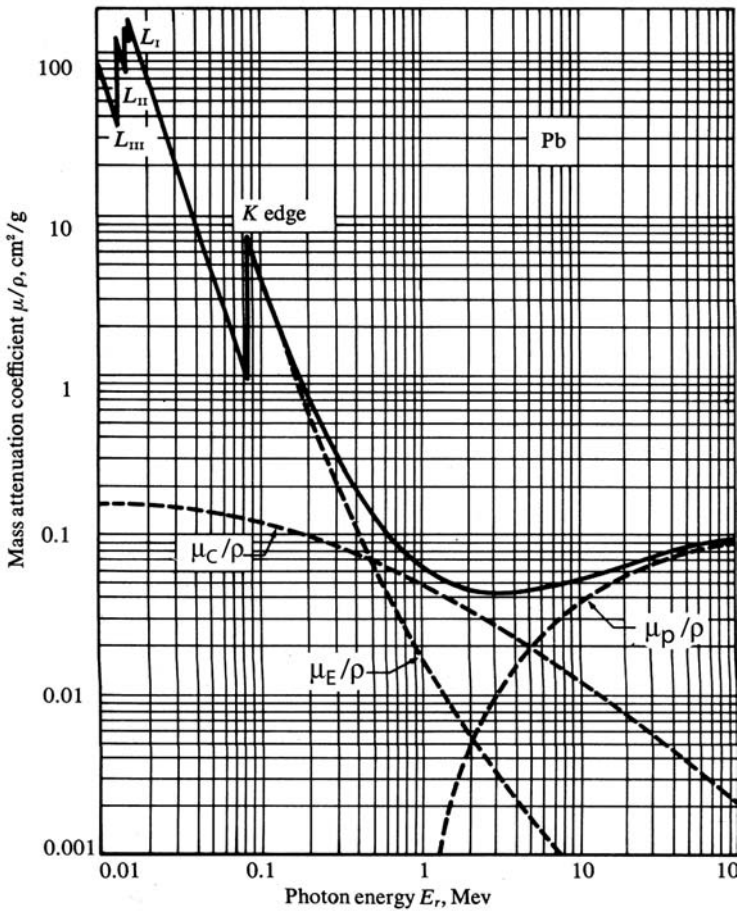


FIGURE 7.20 (a) Plot of the mass attenuation coefficient for Pb from 10 keV to 10 MeV (solid line). Note the sudden increase in absorption (an *absorption edge*) at about 0.088 MeV when the energy level is high enough to activate the photoelectric effect. The dashed lines represent the relative contribution from the different attenuation mechanism of Compton scattering (μ_c/ρ), photoelectric effect (μ_E/ρ), and pair production (μ_p/ρ). The energy dependence of attenuation is complex and simple rules are inadequate. (b) Plot of the mass attenuation coefficient as a function of atomic number for two energies: 50 keV and 3.5 MeV. While it is often stated that attenuation increases with atomic number, neither of these curves is well-described by this description. (Fig. 1.5 p. 716 from WE Meyerhof. © 1955, by Addison-Wesley Publishing Co. Inc. Reprinted by permission of Pearson Education, Inc.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

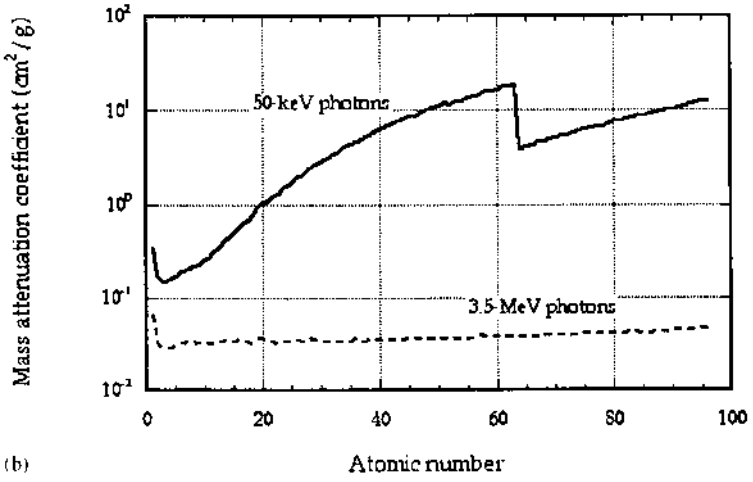


FIGURE 7.20 (*continued*)

It is important to point out that radiographic measurements are measurements of attenuation, which is a function of energy, Z, and density. You cannot separate the elemental composition, Z, and density from the attenuation coefficient from a single energy measurement since there are two unknowns and only one equation. The Z-dependence of attenuation coefficients suggests a method of analysis based on attenuation measurements at different photon energies. The method can be applied when the constituents of a material are known and the relative amounts are to be determined. This method is called *dual energy*, and is most obviously applied when one of the constituents has an absorption edge in a useful energy range, so that attenuation measurements can be performed on each side of the absorption edge. It can be employed to great advantage in less obvious situations, such as that of Martz, et al. (28) in which they imaged, in three dimensions, the density of each constituent in a two-part explosive material made up of explosive and binder. For further details on this method see Process Control of High Explosive Materials using CT in Section 7.5.2.

One important consideration when making NDE measurements employing photons is what energy to employ. This can be analyzed in various ways, but first let us perform a mental experiment. *Realize that to choose energy is to choose attenuation*. As photon attenuation approaches infinity, then no photons are transmitted and no information about the part in question is contained in the "transmitted" radiation. On the other hand, as attenuation approaches zero, then no information about the part is contained in the transmitted radiation. It is obvious that the only interesting case is when some, but not all, radiation is

Copyright © Marcel Dekker, Inc. All rights reserved.

transmitted. It has been shown by Grodzins that for ideal conditions, the optimum energy should be chosen so that $\mu l = 2$, or transmission is about 13%. These results are for a vacuum and most often DR and CT imaging is not done in a vacuum and for these cases the transmission can range from ~10–30%.

7.2.4 Non-Photons Radiation Used for NDE

Three radiation forms other than photons are widely used for NDE: β particles, neutrons and protons. α particles are rarely used in NDE or medical imaging.

β particles are energetic *electrons*, *positive (positrons) or negative*, emitted from nuclei during decay. The electrons are emitted with a continuous energy spectrum. Electrons slow down and stop in material much differently from photons, with many interactions and tortuous paths. The combination of the continuous spectrum and attenuation behavior often yields a transmitted energy that is approximately exponential with absorber thickness, for relatively thin absorbers ($< 100\,\mathrm{mg/cm^2}$). This leads to a technique called *β gauging* where a sheet of absorber is placed between source and detector. The resulting signal can be a sensitive indicator of the areal density of the absorber.

Neutron attenuation (usually used for projection radiography) is a very important specialty in NDE. It is largely complementary to X-radiography—answering different questions about an object. Neutron radiography with *thermal* neutrons dominates the field, though recent interest in using ~10 MeV-neutrons (29) is spurring development at these high energies. Attenuation of neutrons is governed by exponential attenuation, as with X- and γ-rays. However, the dependence on atomic number is, fortunately, quite different from photons. Many of the applications for thermal neutron radiography are based on the high mass attenuation coefficient for hydrogen, which is three orders of magnitude higher than lead. Thus hydrogenous materials are displayed dramatically with thermal neutron radiography, while with X-ray photons they are often scarcely visible at all. For instance, thermal neutron radiography is widely used to image hydrogenous corrosion products that are invisible to X-rays, and to visualize explosives inside of metal casings. The application niche for ~10-MeV neutrons is for objects of high opacity to X-rays, especially nuclear weapons. Thermal neutron sources are either nuclear reactors or fissioning isotopes. Neither is simple nor cheap. Sources of 10-MeV neutrons require an accelerator-based system.

Protons are charged hydrogen nuclei produced in an accelerator. They attenuate very differently from other forms of radiation. As it moves through a material, the proton loses energy to electrons because of the positive and negative electronic charge interaction. The energy of the proton is far in excess of electron binding energies, so the slowing down of the proton depends only on the electron density along the path, and not on the atomic number of the material.

Copyright © Marcel Dekker, Inc. All rights reserved.

The energy loss that occurs is fundamentally different from the photon-by-photon removal that takes place with X-rays and neutrons. With *protons*, every proton interacts with many millions of electrons, so measurement of the energy loss of a single proton is a significant measurement. It is not necessary to count a large number of protons. Contrast this to X-ray measurements where collecting of a single photon is a binary event. The result of using a low count of protons is subject to Poisson statistics. *One hundred incident protons can yield a result accurate to 0.1%.* One hundred incident photons used optimally result in an accuracy of 30%.

The proton measurement is inherently a single *pencil* (parallel) *beam* and single element detector. The beam is swept and the part is rotated/translated. Spatial resolution is about 1–10 μm depending on the areal density of the sample. Electron density resolution of 1 part in 1000 is typical. Sample size is limited by penetration power to about 1 mm of solid material.

7.2.5 Radiography

We have all seen the results of a medical radiograph where a two-dimensional image is created of a three-dimensional object. While the lateral dimensions of the object—those perpendicular to the direction of radiation travel—are retained in the radiograph, the depth dimension, parallel to the direction of radiation travel, is lost. The intensity of the image yields the absorption qualities of the object along the depth but superimposes any variations along that dimension thus loosing any depth information of a feature. Alternatively, a radiograph can be thought of as the superposition of many discrete images, each one representing a single plane within the object. This can be confusing and causes ambiguity. *Different objects can produce identical radiographs*.

Most radiographs record the spatial dependence of absorption using a two dimensional detector such as film. Other systems scan the object using linear array or point detectors by either moving the object or by moving the source-detector combination. (Point detection is very slow and only used for highly quantitative work.) In any case, the theory governing radiation transport is the same as we described earlier for single ray transmission. Each point in the image is the result of radiation transmission along a single ray from the source through the object to the detector.

When radiography is performed with area detectors, scattered radiation from the object can be an important cause of image degradation. This is the *opposite* of good or narrow beam geometry and is often referred to as *broad beam geometry*. In severe cases, scattered radiation can exceed transmitted radiation by a factor of 10 or more.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

7.2.6 Computed Tomography (CT)

It is perhaps easiest to understand computed tomography by describing the end result of the process. The product of CT is a block of data representing the object in three spatial dimensions. Numerical values assigned to each voxel are linear attenuation coefficients. Without cutting the object or damaging it in any manner, CT enables one to view any cross-sectional plane within an object with all other planes eliminated.

The process to create a CT image of a three-dimensional object begins with the acquisition of multiple projection radiographs at different angles. Computer algorithms take these (digitized) radiographs and reconstruct an image that contains attenuation information in all three dimensions. Figure 7.21 illustrates how individual radiographs (that lack depth information) can be combined to extract information in all three dimensions. In this simple example, we wish to determine the attenuation coefficient for each of the four voxels. Imagine that we take two projection radiographs in orthogonal directions and then digitize each radiograph into two pixels. The results of the two projection radiographs yield the integral for the attenuation values of the ray in their respective directions. The attenuation integral in the top row is 13, the bottom row is 6, the left column is 14, and the right column is 5. One possible reconstruction algorithm to determine

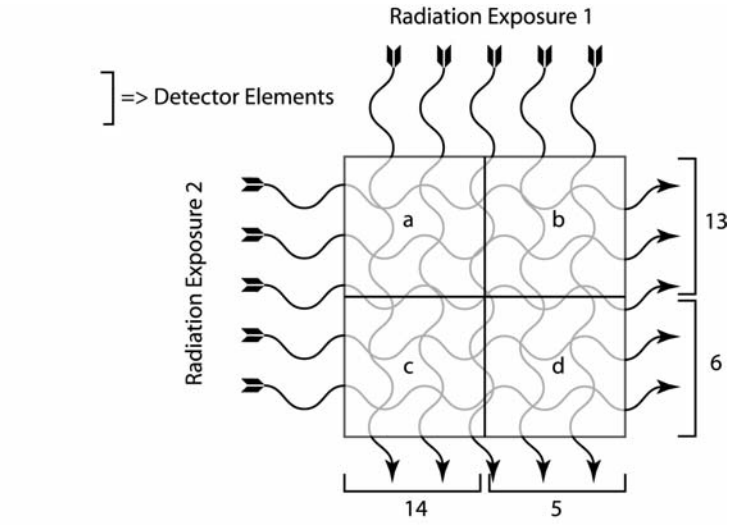


FIGURE 7.21 Model indicating how the attenuation of a volume element (voxel)—or in two dimensions a pixel—might be calculated from multiple projection radiographs.

Copyright © Marcel Dekker, Inc. All rights reserved.

the specific attenuation in each of the voxels a through d simply solves the four simultaneous equations:

$$a + b = 13$$
$$c + d = 6$$
$$a + c = 14$$
$$b + d = 5$$

$a = 10$, $b = 3$, $c = 4$, and $d = 2$.

Figure 7.22 shows the simplest configuration for recording CT data. It consists of a source and single detector, which are usually fixed in NDE systems and the object is translated and rotated to acquire the necessary projections. All data acquired in this scanning scenario represent a single plane parallel to the direction of radiation travel through the object with no confusion from adjacent planes. As a consequence, the reconstruction will be of this CT plane. The object is moved vertically and another scan performed to acquire the data for other CT planes.

The input to the reconstruction process is multiple attenuation measurements along specific known rays (Figure 7.21). Reconstruction, then, is a process of determining what attenuation values for each voxel best represent the entire

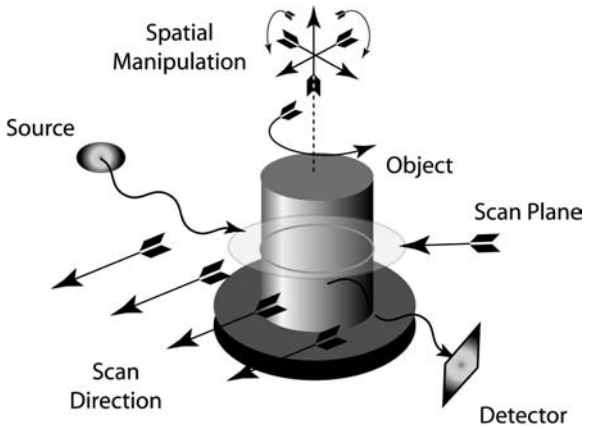


FIGURE 7.22 The simplest configuration for acquiring CT data is shown here, a single source, a single detector, translation of the object in the direction X_f, and rotation of the object, ϕ. This system is sometimes referred to as a pencil-beam CT scanner. This data can be reconstructed into single CT slice. The object must be shifted vertically and another data set acquired in order to reconstruct other CT slice planes to get all three dimensions.

Copyright © Marcel Dekker, Inc. All rights reserved.

data ensemble. It has been described as being analogous to reconstructing a steer from hamburger. The mathematics of reconstruction is too complex to address in detail here.

We adapt the description of Barrett and Swindell (30) to define the problem. We define two coplanar reference frames: the x–y frame, which rotates with the object, and the x_r–y_r frame in which y_r is in the direction of the X-ray beam. Both frames have their origin at the center of rotation provided by a rotation stage. The angle ϕ defines the angle between the two frames. The linear absorption coefficient is *spatially variant*. Therefore, we assign notation $\mu(\vec{r})$ for μ at position defined by vector $\vec{r}$. Eq. (5) must now be modified to account for *nonhomogeneity* in μ along the absorption path l (i.e., along the path traveled by the radiation or raypath):

$$I(\phi, x_r) = I_0 e^{-\int_t \mu(\vec{r})dy_r} \tag{7}$$

where $I(\phi, x_r)$ is expressed in terms of parameters ϕ and x_r. Taking the natural logarithm of both sides of the equation gives rise to a new function $R(\phi, x_r)$ which is the line integral or *raysum* over the *raypath* of $\mu(r)$:

$$R(\phi, x_r) = -\ln\left(\frac{I(\phi, x_r)}{I_0}\right) = \int_t \mu(r)dy_r \tag{8}$$

The quantity $R(\phi, x_r)$ is called the *Radon transform*. CT reconstruction is the process of inverting this equation.

This derivation treats the Radon transform as if it were a continuous two-dimensional function of ϕ and x_r. In practice sampling effects are important. If the projection data, $R(\phi, x_r)$, are plotted as a function of ϕ and x_r, in Cartesian coordinates with ϕ on the ordinate, the result is known as a *sinogram*. A single absorber in object space not on the center axis of rotation maps as a sine function in a sinogram.

The sinogram is an extremely useful method to display CT data. To the novice, it appears pretty but nonfunctional. However, it makes two types of CT artifacts immediately apparent, and partially correctable. If there are motion irregularities, especially abrupt source spatial instability, it appears as a discontinuity in the ϕ-direction. In multiple detector systems, detector response calibration errors appear as vertical streaks, since each vertical line in the image is the result of a single detector.

Many methods for inverting the Radon transform have been devised. All rely on numerical computation and depend on modern computing capacity. Reconstruction times are generally shorter than data acquisition times except for complex (cone beam) geometries. This is also an area of current research. These new methods tend to be more computationally intensive as greater computing power becomes available.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

7.2.7 Radiation Transport Modeling—Monte Carlo Method

For five decades, the Monte Carlo method has been applied to radiation transport problems of high importance. It is helpful to think of this method as a computational analog of the actual physical process. For the sake of discussion, we will talk about computation involving photons, though early applications were more focused on neutrons. A source photon is created with energy and direction as specified. The dice are computationally rolled to determine where this photon first interacts in the specified geometry, then again to determine which interaction process takes place and what secondary radiation is produced etc. This is a powerful technique resting on compilations of probabilities (attenuation coefficients). It is the best method of treating scattered and secondary radiation. When a photon has either escaped the problem or been absorbed, the calculation returns to start another source photon.

The Monte Carlo technique suffers from the same statistical limitations imposed in nature. To get a precise answer, you have to run a lot of source photons. This makes Monte Carlo calculations computationally intensive. This is no longer prohibitive on modern computers. Enormously powerful general purpose Monte Carlo radiation transport codes are available from Los Alamos National Laboratory (31) and Lawrence Livermore National Laboratory (32). These run on affordable computers.

Simulations open a window for understanding that cannot be approached experimentally. Once a problem is defined, it is easy to change parameters and run again. Photons can be made perfectly monoenergetic and perfectly on-axis. Scatter can be tallied separately from direct radiation. Scatter contributions from different objects, for example air and collimators, can be separated.

7.2.8 Dosimetry and Health Physics

An area of radiation measurements that is frequently misunderstood is that of *exposure*, *dose*, and *dose equivalent*. These are vital issues in personnel protection. The concept of exposure was introduced early in the history of working with γ- and X-radiation. The unit of *exposure* is defined in terms of the charge produced by secondary electrons formed in *air*. The historical unit of exposure is the Röntgen (R). It is defined as the radiation that results in 2.08×10^9 ion pairs per 0.001293 g (1 cm^3 at Standard Temperature and Pressure) of air. Once again, exposure is defined in terms of ionization produced in air. There are many experimental nuances involved in exposure measurements that are beyond this text, but one of the most useful characteristics of exposure is the ease with which it is measured, most commonly with an air-filled ionization chamber (33). Radiation exposure is not simply related to radiation dose. It is an unfortunate coincidence that using the historical units, the numerical values of dose in soft tissue and exposure are numerically similar. Adding to the confusion is that the

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

historical unit for dose is *rad* and the abbreviation, R for Röntgen, is often *incorrectly* taken to mean rad. In SI units, exposure would be expressed in coulomb/kg. *However, no special name has been assigned to this unit and it is almost never used*.

Radiation dose, more precisely called *absorbed dose*, is defined in terms of energy absorbed from any type of radiation per unit mass of any absorber. As mentioned, the historical unit of dose is the rad, D_r, which is defined as 100 ergs/g of absorber material. The SI unit of dose is the *gray* (Gy) defined as 1 J/kg. This leads to a simple conversion of 1 Gy = 100 rad. Absorbed dose should be an approximate indicator of cellular damage and therefore a reasonable unit to use for radiation effects in biological specimens. Unfortunately, it is generally difficult to measure, so indirect methods based on ionization are most common. *Be aware that a radiation dose stated in rad is not complete without a definition of the absorber.* The dose delivered to soft tissue may differ considerably from the dose to bone or the brain from the same radiation exposure!

The term *dose equivalent* is used as a measure of potential biological effects from a radiation field. An organism exposed to equal dose equivalents in different radiation fields would be expected to incur equal effects. When considering the effects of radiation on living organisms, the absorption of equal amounts of energy per unit mass (absorbed dose) under different radiation conditions generally does not result in equal biological effects. It has been shown that radiation with high energy density along the particle track is more damaging than radiation in which the same energy is more dilute. This may arise in part from less effective damage repair in highly damaged regions. The measure of energy density used to characterize this effect is the *linear energy transfer* (LET). Its units are energy/length. A term known as *quality factor*, Q, relates absorbed dose, D_r, to dose equivalent:

$$H = QD_r \tag{9}$$

H = dose equivalent. The historical unit for dose equivalent is rem. Multiplying the absorbed dose in units of rad times the quality factor yields dose equivalent in rem. If absorbed dose in the SI unit of Gy is multiplied by the quality factor one gets dose equivalent in the SI unit of *Sievert* (Sv). Quality factors for X- and γ-ray photons are 1.0. Q for charged particles are higher (as high as 20) and Q for neutrons, which deposit energy by creating charged particles in situ, are also higher and vary significantly with neutron energy.

Health physics specialists carry this concept even further and have defined tissue weighting factors (34). This is beyond our scope here, but be aware that they use a term *effective dose*, which has units of Sv. The effective dose is designed to express a radiation dose in terms of the whole body equivalent. A uniform dose equivalent of 1 Sv delivered to the entire body would result in an

Copyright © Marcel Dekker, Inc. All rights reserved.

effective dose of 1 Sv. A dose equivalent of 1 Sv delivered to one hand would result in a much smaller effective dose.

Radiation protection standards are specified in terms of dose equivalents, with different limits set for whole body and extremities. The maximum permitted occupational dose is 50 mSv to the whole body. The maximum annual dose for personnel in unrestricted areas is 5 mSv, which is comparable to dose received from natural and medical radiation. Minors and fetus have special limits. In addition, the dose should always be made as low as reasonably achievable. Personnel who could be reasonably expected to receive occupational dose *must* wear personal dosimeters—radiation dose meters. Thermoluminescent dosimeters (TLDs), the most common dosimeter used, are monitored and changed on a time basis, e.g., monthly. The dosimeters may be badges for whole body exposure or rings for measurement of dose to fingers for localized radiation where the hand may be exposed. There are various electronic dose meters available that provide prompt readout, and even audible indication of dose rate.

7.3 EQUIPMENT FOR RADIATION TESTING

7.3.1 Introduction

Radiological NDE systems contain the same basic components similar to many other NDE methods. These components can be separated into the two categories: (1) the NDE measurement system and (2) analysis and interpretation (Figure 7.23). In this section, we will concentrate on the measurement system that includes the source, object, and detector. The source represents the origin of the interrogating energy—X-, γ-ray, or neutron radiation. Radiation sources can be

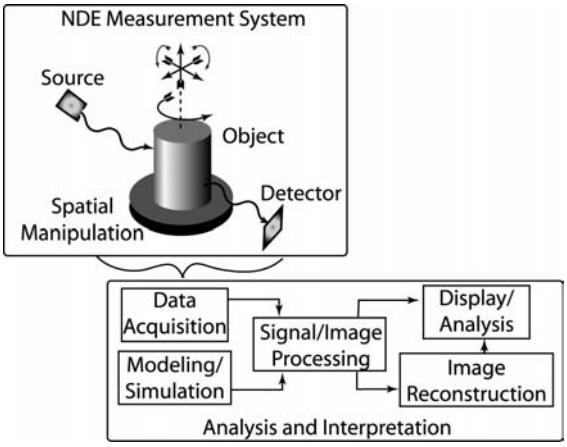


FIGURE 7.23 Generic NDE system.

Copyright © Marcel Dekker, Inc. All rights reserved.

configured to act as either a point, a line, or an area source. The object is the interrogated material that must interact with the radiation (absorption) to reveal material characteristics and internal features. The detector monitors the changes in the radiation from the source due to radiation–object interaction. While X- and γ-ray systems and equipment vary greatly from simple film radiography to complex CT imaging systems, they all posses these same basic elements. Generally, the equipment used in these systems is not unique to the particular system, e.g., the same linear detectors may be used for projection radiography of a baggage scanning system as those used for a CT machine provided the detector has the necessary properties, such as scanning rate and sensitivity, for the particular application.

In this section, we will present the most common radiation sources and detectors used in radiological NDE. Additionally, we will introduce spatial manipulation equipment used in scanning and CT systems.

7.3.2 Radiation Sources

We have divided our discussion of radiation sources into two categories:

Electrically powered—a stream of energetic electrons is either slowed within a target or turned by magnetic fields, liberating photons with energy that range from a few keV up to several MeV

Radioisotropic—radioactive materials that spontaneously emit X-ray or gamma(γ)-ray photons as these unstable atoms decay

Electrically Powered Sources

X-ray Machines. The most important source of X-rays for NDE employs the physical principle that led to Röntgen's discovery, i.e., when energetic electrons strike an object, X-rays are produced. X-ray machines consist of an X-ray generating tube, power supply and cables. X-ray tubes create radiation by the collision of accelerated electrons with a target material releasing photons with energy that range from a few keV up to several MeV. The emitted radiation spectra consist of two distinct components: (a) a broad continuous spectrum; and (b) narrow band (discrete or characteristic lines) of radiation. X-ray tubes offer a significant degree of control over the output radiation:

By adjusting the voltage that accelerates the electrons, we can control, to a large degree, the output energy spectrum, i.e., the energy of the photons.

By inserting intentional X-ray filtration, we can control, to some degree, the output energy spectrum, i.e., the energy of the photons.

By choice of target material we can select the energy of the discrete bands (characteristic lines) of radiation produced.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

By adjusting the filament current, which controls the number of accelerated electrons (*tube current*), we can control the output intensity, i.e., the number of photons.

By choosing the focal size, we can strongly influence the spatial resolution and number of photons.

A basic X-ray tube consists of a filament, target, and a vacuum enclosure (Figure 7.24). A vacuum is needed inside of the tube in order to generate and transport electrons. The supporting system includes a power supply for the filament current and a high voltage power supply for the biasing voltage to accelerate the electrons (Figure 7.25).

In operation, current flows through the resistive *filament*, also called the cathode, liberating electrons that are accelerated, by a biasing voltage, towards the *target*, also called the anode. X-rays are produced by the conversion of electron kinetic energy into electromagnetic radiation. On average, most of the energy of electrons produces unwanted heat. This limits the output intensity and lifetime of

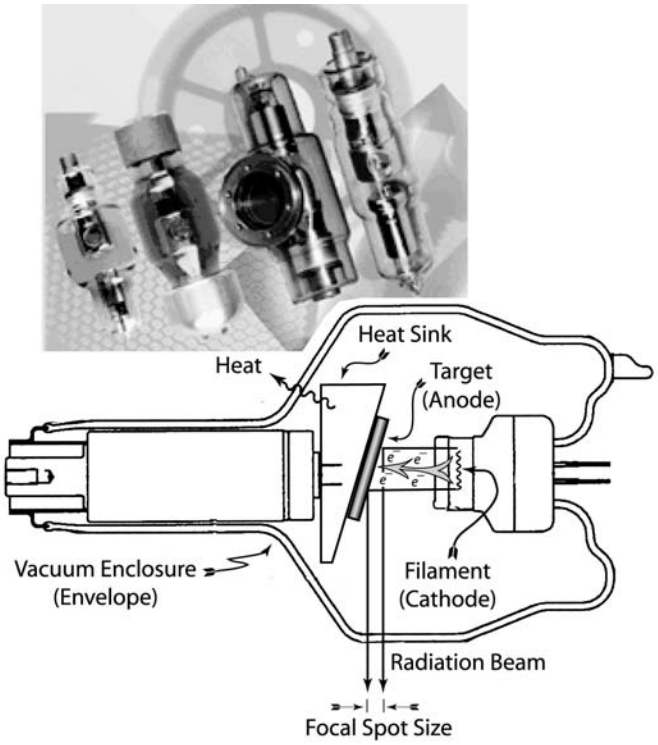


FIGURE 7.24 X-ray tubes. (Adapted: Photograph courtesy Varian.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

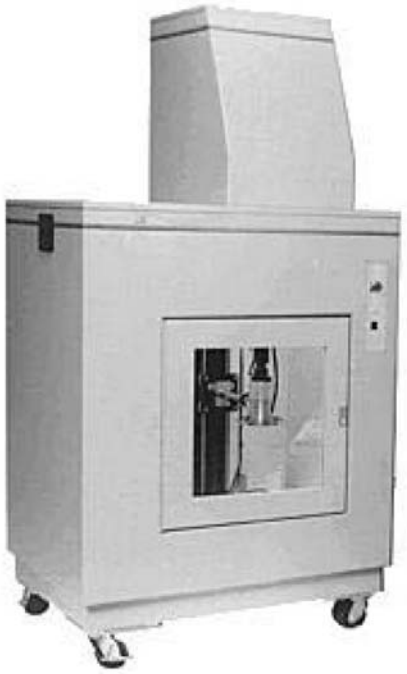


FIGURE 7.25 Industrial X-ray machine.

X-ray tubes. Once in awhile, an electron comes into close proximity to the large electric field caused by the positively-charged nucleus and a significant energy loss occurs, creating a useful photon. X-rays produced in this manner are called *bremmstrahlung* radiation, a German term meaning *braking* radiation (Figure 7.26). A direct impact of an electron with the anode nucleus results in the loss of all of the electron's kinetic energy and therefore conversion of all of this energy into an X-ray. This is an exceedingly rare event. The intensity of bremmstrahlung radiation continuously increases as one goes to lower energy, but few of the lowest energy (< 5 keV) X-rays produced escape the anode, pass through the tube window and are transported to the object being examined. The most probable energy of bremmstrahlung X-rays reaching the object (without intentional filtration) is 0.3–0.5 of the incident electron energy. The maximum possible X-ray energy (keV) is equal to the electron energy (keV) which is numerically equal to the accelerating potential (kV).

When the energy of the incident electron exceeds the binding energy of an orbital electron, it can eject the orbital electron leaving the atom with an unfilled shell. An electron from another shell will fill this vacancy and emit a *characteristic* photon (Figure 7.26). These photons are characterisitc of the anode element.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

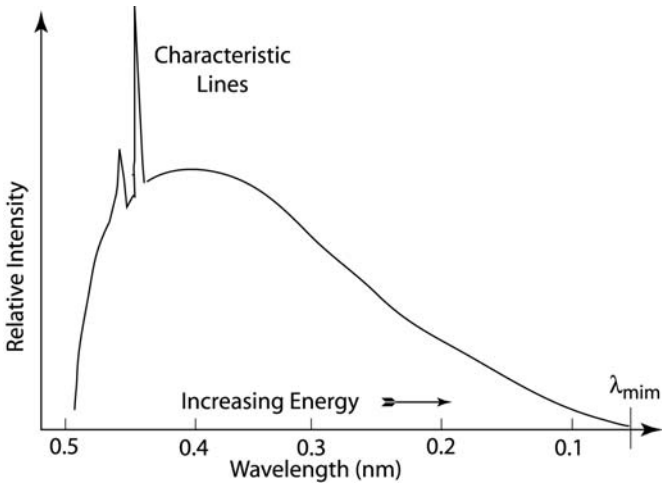


FIGURE 7.26 Radiation spectrum from an X-ray machine showing *bremsstrahlung* (continuous) radiation and the characteristic peaks.

The dominant characteristic X-ray energy emitted from a tungsten anode is 59.0 keV when the electron energy exceeds the binding energy of an inner-shell electron, 69.5 keV.

In practice, the most often used control of the spectrum of emitted X-rays is variation of accelerating voltage that gives the electrons their kinetic energy $\frac{1}{2}m_e v^2$. The minimum wavelength (maximum energy) of emitted energy (maximum photon energy) is related to the kinetic energy by

$$\mathrm{KE}_e = \frac{1}{2}m_e v^2 = eV = \frac{hc}{\lambda_{\min}} \tag{10}$$

where

KE_e = kinetic energy of the moving electron.
m_e = mass of the electron.
v = velocity of the accelerated electron.
V = bias voltage.
e = electronic charge on an electron (1.602×10^{-19}C).
h = Plank's constant (6.626×10^{-34} Js).
c = speed of light in a vacuum (2.998×10^{8} m/s).
$\lambda_{\min}$ = wavelength corresponding to maximum energy radiated.

Rearranging this equation, we have

$$\lambda_{\min} = \frac{hc}{eV} = \frac{1.24}{V}\,\mathrm{nm} \tag{11}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

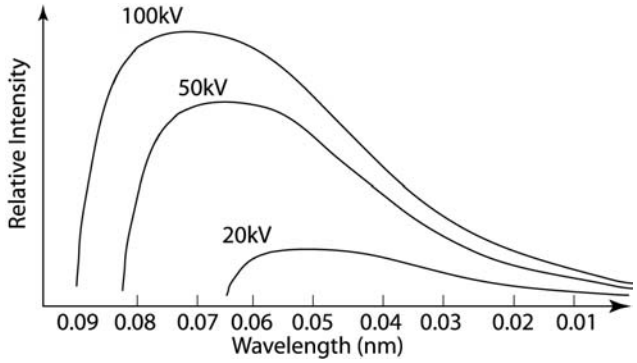


FIGURE 7.27 Effect of varying the biasing voltage on the energy spectrum for an X-ray machine.

Figure 7.27 shows that as the biasing voltage is increased the continuous radiation spectrum shifts to shorter wavelengths (higher energies)—for clarity the characteristic lines are not shown.

The *intensity* of the radiation—the number of photons emitted per unit time—is controlled by the number of electrons that strike the target—the *tube current*. For a fixed bias voltage the tube current is controlled by the *filament current*. Typical values of tube current range from hundreds of microamperes for microfocus (~10's μm spot size) X-ray systems to tens of milliamperes for high intensity industrial units with larger (~1–2 mm) source size.

Often the usable tube current is limited by the power (heat) dissipation in the target (anode). Only about 1 to 10% of the bombarding electron energy is converted to X-rays; the remaining energy is converted to heat. This power limit (given in watts, W) is the product of the tube current (amps, A) and the bias voltage (volts, V). A tube with a 50-W limit is capable of dissipating 1.0 mA of current when operating at a voltage of 50 kV but only 0.5 mA when operating at 100 kV. To improve cooling, targets are usually backed with a high thermal conductivity material such as copper that acts as a heat sink and conductor. These heat sinks may be actively cooled. Another method to dissipate target heat is to employ a rotating anode where the tube current strikes only a small portion of the anode disk near the perimeter as the disk rotates (Figure 7.28).

The effective (focal) spot size of the X-ray beam is another important attribute of the radiation source. The component of image blur called *source unsharpness* is directly related to the focal spot size of the X-rays leaving the target. To explain this concept, compare a shadow of an object when illuminated by a point light source vs. an extended light source (Figure 7.29). The figure shows ray traces from the source of the light to the image. Note the point source produces a well-defined shadow (sharp edges) whereas the extended source has

Copyright © Marcel Dekker, Inc. All rights reserved.

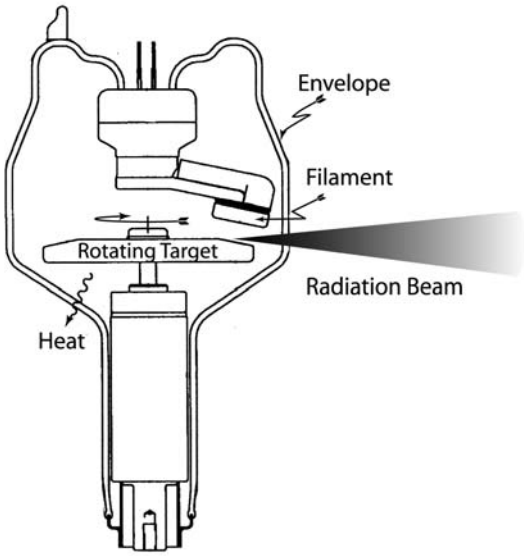


FIGURE 7.28 Rotating X-ray target.

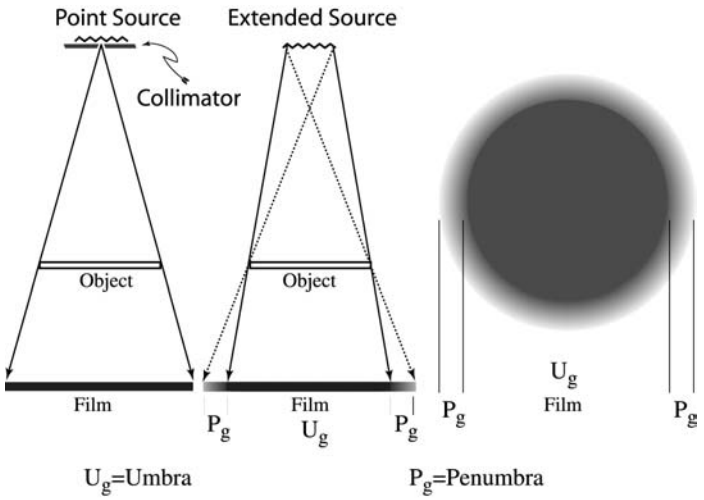


FIGURE 7.29 Unsharpness (penumbra) due to a point and an extend source.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

an undefined (blurred) shadow edge—the *penumbra*. (The *umbra* is the entire shadow minus the penumbra.) The size of the penumbra is a measure of the sharpness (or *unsharpness*) of the image and can be calculated by

$$P_g = \frac{fd}{D_o} \tag{12}$$

where

P_g = width of the penumbra.
f = effective focal spot size.
d = distance from the object to the image plane.
D_o = distance from the source target to the object.

Note that the size of the penumbra decreases with decreasing effective focal size or as the distance between the source target and the object increases just as an automobile head light appears as a point source at large distances. On the other hand, the penumbra increases with increasing distance between the object and image plane or increasing effective focal size.

The relationship between the effective spot size and the actual focal spot size on the source target is determined by the angle of the target to the incident tube current (Figure 7.24). Typically, the target is angled such that the useable X-ray beam leaves at an angle approximately 90° to the tube current. Thus the effective spot size is the angular projection and not the actual size of the target struck by the tube current. To aid in controlling the focal spot size, X-ray tubes incorporate a focusing cup at the filament to control the diameter of the electron beam. Many incorporate additional electrostatic or electromagnetic elements that shape and guide the electron beam.

Now that we know how to improve image sharpness, what is the cost? We have explained that heat dissipation typically limits tube current. Similarly, heat per unit surface area (power density) on the anode imposes a limit. Therefore, tube current must be reduced if the spot size is made smaller. For high-spatial resolution (minimal blur) microfocus systems with focal spot sizes on the order of a few micrometers, the radiation intensity is typically very small. A 10-μm focal size operates at a power level of $\sim$10 W.

Industrial tubes are usually designed for continuous operation. Exposure times of minutes are common, especially when film is the image detector. For industrial CT, tubes are routinely left in continuous operation for several hours or even days. This is in sharp contrast to medical X-ray tubes, which are designed for short intense bursts of radiation. Subsecond exposure times are necessary in medicine to minimize motion problems from heartbeat and respiration. Medical tubes are usually of rotating anode design and therefore capable of higher output than industrial tubes for short exposure times. These tubes store heat in the anode which is then dissipated slowly while the patient is repositioned or changed for a new one. They cannot be operated continuously.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

In general, electrically-powered X-ray sources and associated equipment are not highly portable. They can have high source intensity (sometimes referred to as *strength*) and small source focal spot size. They have the important characteristic that they stop emitting radiation when electrical power is removed. This facilitates safe operation and makes it simple to work around the source when not in operation. They emit complex spectra with both bremsstrahlung and characteristic radiation. The spectrum can be adjusted in a limited way by control of the applied voltage and by filtration. The lowest energy X-rays available from conventional X-ray tubes is about 5 keV. This is limited by the vacuum window, usually ~0.03-mm of beryllium, and by the escape of X-rays from the anode. Above 450-kV applied potential, electrostatic acceleration (as in tubes) give way to linear electron accelerators called *linacs*. The practical upper limit for machines that accelerate electrons in order to generate X-rays is 25-MeV electron energy.

Synchrotrons. Synchrotrons provide a near ideal radiation source: high-intensity, tuneable monochromatic energy, and a highly collimated source of X-rays that can be used for quantitative X-ray imaging (35, 36). Synchrotron radiation is produced as charged particles are bent or angularly accelerated. This is usually done with a large electron accelerator with a circular orbit. The electrons are turned to maintain the orbit or wiggled—angularly accelerated—and emit a broad energy spectrum of high-intensity radiation ranging from infrared to ~50 keV X-rays. These are of such high intensity that monochrometers can then be used to select the desired energy. However, due to their physical size and cost, there are very few synchrotron sources and access is limited. Nonetheless, synchrotron sources are increasing in popularity, as the imagery produced with these sources can be exceptional.

Radioisotopic Sources

Many natural and manmade materials spontaneously emit photons, either γ- or X-rays. X-rays emanate from electronic orbital transitions whereas γ-rays arise from proton or neutron orbital transitions (protons and neutrons have energy orbital analogues to electrons). X-rays are produced from radioactive materials when a nucleus decays by electron capture or when primary decay radiation ejects an orbital electron, leaving an excited atom. The subsequent filling of an inner shell electron orbit results in the emission of characteristic X-rays (Electrons and Radioactivity in Section 7.2.2). Similarly, the filling of an inner proton or neutron orbit results in the emission of characteristic γ-rays. In general, radioactive isotopes that decay by emission of an α-particle, also emit characteristic X-rays.

For industrial applications these radiation sources offer the advantages of portability, no power requirements, small physical size (sometimes as small as marbles or BBs) as compared to size of an X-ray machine (tube head, cables and

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

power supply), and radioactive emission at specific photon energies (monochromatic radiation). Without radioisotopic sources many field applications such as pipeline and petroleum storage tanks could not be inspected using radiography. The small source size allows inspection of interior spaces and complex geometries such as piping and ship weld inspection. The monochromatic radiation emitted from radioisotopic sources enables quantitative applications. The primary disadvantage of radioisotopic sources is their low source strength per unit source volume. In order to get useful source strength, (number of photons emitted per unit time), large focal spot size (actual size of the source) must be employed. Secondary disadvantages of radioisotopic sources are that they cannot be turned off and they present a contamination hazard if containment is accidentally breached.

Radioactive material is energetically unstable and emits radiation at specific energy levels when the nucleus *decays* to a new state. For neutron or proton transitions the emitted photon radiation is called a γ-ray, while electronic transitions photons are called X-rays. This event is called a *disintegration*. The new state (*daughter* or *daughter product*) may be a new radioactive isotope or it may be stable. A given radioisotope may emit at multiple energies, e.g., Iridum-192 emits γ-rays at three distinct energies: 0.31, 0.47, and 0.60 MeV. γ-rays generally have higher energy than X-rays and thus more penetrating power. (However, this is not always true, for example, a 4 MV LINAC produces higher energy X-rays than the γ-ray emitting Ir-192 radioisotope.) Although many materials are radioactive, only a few are commonly used in industrial NDE. Table 7.3 list four common radioactive isotopes used as sources for gamma radiography.

Once a radioactive nucleus decays, it is no longer possible for it to emit the same radiation. Therefore, the intensity of radioactive γ- and X-ray emitting sources decreases with time. The radioactive *half-life* (τ)—the time for half of the originally radioactive atoms to decay into daughter atoms—is commonly used to describe the average decay rate of radioisotopic sources. The actual probabilistic decay rate λ_r (disintegrations/second) is related to the half-life by

$$\lambda_r = \frac{\ln 2}{\tau} = 0.693\frac{1}{\tau} \tag{13}$$

This *disintegration rate* (dis/s) has a unit of *Curies* (Ci) or in SI units *Becquerel* (Bq): $1\text{Ci} = 3.7 \times 10^{10}$ dis/s and 1Bq = 1dis/s. Normalizing the decay rate to the mass of the radioactive material, we have the *specific activity*

$$\text{Specific activity} = \Lambda\frac{N_A}{A_w} = \lambda_r N_t \frac{N_A}{A_w}\left(\frac{\text{Ci}}{\text{g}}\right) \tag{14}$$

where

N_A = number of atoms per mole of material (*Avogadro's number*)
A_w = atomic weight of the specific atoms (grams per mole)
$\Lambda = \lambda_r N_t$ = (disintegration rate)(radioactive nuclei at time t)—see Eq. (3).

Copyright © Marcel Dekker, Inc. All rights reserved.

TABLE 7.3 Important Characteristics of Four Radioisotopic Sources

Isotope	^{60}Co	^{137}Cs	^{192}Ir	^{170}Tm
Half life	5.3 yr	30 yr	74 day	129 day
Chemical form	Co	CsCl	Ir	Tm_2O_3
γ-ray(s) energy (MeV)	1.17 and 1.33	0.66	0.31, 0.47, and 0.60*	0.084 and 0.052
γ-rays/disintegration (for each γ-ray listed)	1.0 and 1.0	0.92	1.47, 0.67 and 0.27	0.03 and 0.05
Practical source activity (Ci)	20	75	100	50
Specific activity (g/Ci)	50	925	350	
Practical source diameter (mm)	3	10	3	3
Al half-value thickness (mm) (for each γ-ray listed)	42 and 48	34	25, 29 and 33	13 and 7.5
Al tenth-value thickness (mm) (for each γ-ray listed)	140 and 160	115	89, 97 and 109	44 and 25
Fe half-value thickness (mm) (for each γ-ray listed)	15 and 17	12	8.1, 10.2 and 11.4	1.6 and 0.51
Fe tenth-value thickness (mm) (for each γ-ray listed)	51 and 57	40	27, 34 and 38	5.5 and 1.7

* ^{192}Ir has 24 γ-rays only a few of the high activity γ-rays are listed.

The number of disintegrations per unit time is observed to decrease exponentially as

$$n = n_0 e^{-(\lambda_r t)} = n_0 e^{-\left(\frac{0.693}{\tau} t\right)} \tag{15}$$

where

n = number of disintegrations occurring at time t
n_0 = number of disintegrations occurring at an earlier time $t = 0$—when the source strength was first measured

Copyright © Marcel Dekker, Inc. All rights reserved.

While n is the number of disintegrations, the source (strength) *intensity* (I) describes the number of photons emitted from a source per unit time. Depending on the particular radioisotopic source, there may be multiple photons emitted per disintegration. For example a 2.3-mm source size, Cobalt-60 is commercially available with activity of 10 Ci. Since one Ci is 3.7×10^{10} dis/s and this radioactive isotope emits two γ-ray photons per disintegration, this represents source strength of 7.4×10^{10} photons/s. For comparison, commercial 9-MV electron linacs produce 1000 times more photons with a smaller (2 mm) spot size. Source intensity takes the same form as the number of disintegrations per seconds:

$$I = I_0 e^{-\left(\frac{0.693}{\tau}t\right)} \tag{16}$$

where I_0 is the measured intensity at $t = 0$.

Representative industrial radioisotopic sources listed in Table 7.3 span a range of energies that implies a range of penetrating ability. Of those listed ^{60}Co has the highest energy of 1.33 MeV and ^{170}Tm has the lowest at 0.052 MeV (52 keV). If the energy is too low for inspection of a given object, the photons will be absorbed within the object and will not expose the detector. On the other hand if the energy is too high, very few of the photons will be absorbed within the object and there will be no discrimination (*contrast*) of the density or material variations within the object; the object would look nearly transparent to the radiation. Therefore, the radiation energy must be matched to the material (elemental) composition, density and the thickness of the object being inspected. Many radioisotopes emit at multiple energy levels as indicated in the table. The table gives the thickness of Al and Fe that is required to attenuate each listed γ-ray to one-half and one tenth of the unattenuated signal.

While the energy of the emitted photons must match the elemental composition, density and thickness of the object such that a reasonable fraction ($\sim$80%) of the energy is absorbed, we also want the length of time required to expose the detector to be as short as possible. For a given source energy, the detector exposure is controlled by the number of photons (source strength or intensity) emitted by the source and the attenuation caused by the object. With respect to Table 7.3, note that ^{60}Co emits photons at 1.17 and 1.33 MeV, but the number of photons emitted per gram of ^{60}Co, *specific activity*, is quite low.

In addition to the source energy and intensity, the size of the source, as previously discussed, influences the image sharpness (unsharpness) via the penumbra effect. For radioactive sources, the spot size is defined by the lateral dimensions of radioactive source or if a collimator is used by the dimensions of its opening (or aperture). Because of the generally low-source strength of radioisotopic materials, the source size is typically quite large (Table 7.3). A useful measure of activity and size is given by the ratio of activity to mass and is

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

called specific activity. Thus radioisotopic source spot sizes are typically much larger than those produced by X-ray tubes.

Because the radiation from a radioisotopic source cannot be turned off, special handling and operating equipment is required to maintain safety. A typical source containment device called a *projector* is shown in Figure 7.30. When not in operation the projector surrounds the source with a radiation shield commonly made from depleted uranium. In operation the system consists of the projector, a control unit, source guide tube, and possibly a collimator. The control unit contains a flexible cable on a reel that moves the source out of the projector and through the source guide tube to the desired location to be interrogated. Both the reel and source guide are detachable. To improve resolution and reduce safety risks, a *collimator*—a highly absorbing container with an aperture that limits radiation to a specific direction—may be added to the end of the source tube. The projector equipped with a snake like source guide allows the operator to negotiate the source through bends in confined internal spaces. Sometimes the source can get stuck inside the snake and it needs to be unstuck. This can lead to radiation exposure to the radiographer who has to manually fix the problem. Some systems use projectors that have a local shutter that can be easily opened and closed without the hassles of using a snake to remotely position the source.

Not only is care required in using radioisotopes, they must be recycled or disposed of under strict regulatory and environmental guidelines. When control is lost, serious incidents can result. For example, in 1987, an abandoned Cesium-137 source was removed by junk dealers from a radiotherapy machine in Goiânia, Brazil (37). The source was taken home and ruptured. Four of those exposed ultimately died and 28 people suffered radiation burns. Several buildings had to be demolished and about 3500 m^3 of radioactive waste were generated during decontamination.

7.3.3 Radiation Detectors

Radiation detectors are transducers that convert the incident radiation into an optical or electrical format that can then be converted to a visual image of the radiation pattern. The wide range of radiation detectors can be classified as analog or digital, by their spatial extent (point, line, or area), or by the specific detection mechanism. Film is the most common analog detector. Analog detectors have a continuous amplitude (contrast) response to the incident radiation. Digital detectors convert the incident radiation to a discrete digital electrical signal with typical data (contrast) resolution of 8 to 16 bits. (Data resolution or dynamic range is a function of digitization of the data and should not be confused with spatial resolution.) There are a large number of different types of detectors operating on different energy conversion principles. The most common detectors are

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

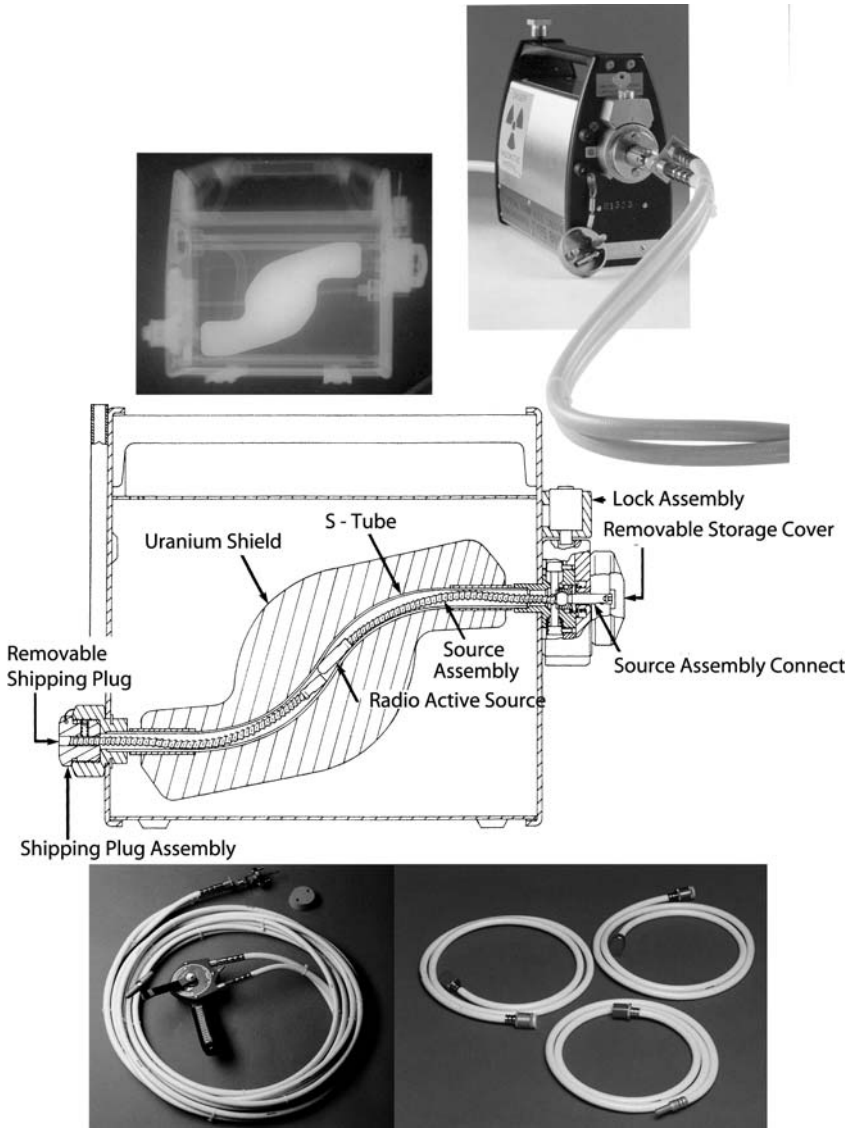


FIGURE 7.30 Isotropic source containment and manipulation equipment. (Adapted: Courtesy AEA Technology.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Film—typically used as an area detector for projection radiographs
Photomultiplier tubes—usually used as point detectors or small area detectors where high gain is required, they are used for projection radiography and tomography
Linear arrays—usually used as a linear detector for both projection radiographs and tomographic imaging
Charge-coupled devices (CCD)—usually used as an area detector for projection radiographs and tomographic imaging
Photostimulable (storage) phosphors—typically used in medicine as an area detector for projection radiographs, with some NDE use as a film alternative
Thin film transistors (TFT) or flat panel—usually used as an area detector for projection radiographs and tomographic imaging

The last five detectors have digital outputs. Note: as with many digital sensors the initial sensing is often analog which is then converted to digital in the output stage.

Before we begin the discussion of these detectors, we will introduce screens and scintillators, which are commonly integrated into the detector system to improve detector efficiency.

Screens and Scintillators

Many radiation detectors incorporate a device that converts the primary radiation into a secondary energy of either lower energy electrons—a *screen*, or light—a *scintillator**. The primary function of these devices is to convert the primary radiation into an energy that more readily interacts with the detector whether it is analog (film) or digital. For film, increased interaction implies a significant reduction of exposure time. Many digital detectors, on the other hand, are specifically visible light detectors and require a scintillator to convert the radiation to visible light.

Screens. Screens are a thin foil (~0.1–0.2 mm) of material (usually lead) used as an intensifier of the radiation that interacts with the detector. Screens perform two functions: (a) intensify the radiation interaction with the detector; and (b) reduce the affects of scattered radiation.† When the primary radiation strikes the screen, part of the absorbed energy liberates lower energy electrons. These lower energy electrons readily interact with the detector. For example, in film radiography, usually less than 1% of the ~80-keV primary radiation energy

*Scintillators are also called screens or fluorescent screens. For clarity, we will use the term scintillator for the light conversion device and screen for the electron conversion device.

†Screens used for many medical applications often place a single, thick screen on the backside of the film. For these cases, the scintillator primarily acts as a barrier for further radiation exposure to the patient as well as a intensifier.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

interacts within the active portion of the film. With the addition of a screen the lower energy electrons increase the exposure by a factor of 2 to 4—*intensification factor*. This intensification factor implies a reduction of the detector (and object) exposure time by one half to one quarter the time without the screen. More importantly in many cases, the lead screens selectively absorb radiation that has been scattered through its interaction with the object. This scattered radiation can significantly degrade image quality.

The advantage of a screen's reduction in exposure time comes at the cost of spatial resolution. As the primary beam interacts with the screen liberating low energy electrons, the screen itself also produces scattered radiation that violates our assumption of radiation moving in a straight line. This scattered radiation moving in random directions also liberates electrons, which slightly blurs the image. In addition, the electrons themselves move in directions other than the incident radiation creating them, adding another source of blur.

Screens are placed directly against the detector to minimize the effect of screen-produced blur. For digital detectors, the screen would be typically incorporated into the front face of the detector and acts as both an intensifying screen as well as a radiation shield for the sensitive elements of the detector. In the case of film detectors, a thin screen is placed in front of the film and a thicker foil is placed behind the film. The thicker layer increases the exposure of the film and also allows for additional absorption (containment) of the radiation. Care of the screens is important as a simple scratch or piece of dust will produce nonrelavent markings on the detector.

Note the thickness of the front screen must be matched to the energy of the incident radiation. If the primary radiation does not have enough energy to penetrate the lead foil and create electrons exiting the foil, the screen will in fact have a negative intensification factor, i.e., cause an increase in exposure time. For this reason, front screens are not used for medical work that is always performed at modest X-ray energy.

Scintillators. Scintillators—often called *fluorescent screens*—like screens are used to increase the effective interaction of the radiation with the detector. However, unlike screens that convert some of the primary radiation into electrons, scintillators convert some of the primary radiation into light. When used with film, scintillators have intensification factors between 2 and 60; they reduce the exposure time by one half to a sixtieth over the primary radiation alone. As might be expected, gain must come at a cost of equal magnitude. When the primary radiation is converted to light, the light propagates isotropically, i.e., in all directions, as if it where emanating from point sources—the resultant blurring on the detector is called *blooming*. Thus, the large advantage of a high intensification factor offered by a scintillator is offset by a significant reduction in spatial resolution.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

In use with film the scintillator is placed in direct contact to reduce blooming effects. For digital detectors applications, where the detector or detector electronics might need to be protected from exposure of the primary radiation, the light is optically steered out of the path of the primary radiation to the detectors (Figure 7.31). The figure shows two turning mirrors that redirect the light to the detector lying outside the path of the primary beam. To minimize the reduction in resolution caused by beam spreading of the light, sophisticated optical methods are employed such as lenses, mirrors, and bundled optical fibers and tapers.

Though more complex than passive scintillators, we include image intensifiers here because functionally they act like a scintillator. An image intensifier is an evacuated electronic device in which the X-ray image is converted to light, then to electrons, which are subsequently accelerated and focused. These electrons impinge on a scintillator, producing the output image. Because of the energy added to the system during electron acceleration, the output image is much brighter than that formed at the input. The output image is typically viewed with a CCD camera. The desired attribute of image intensifiers is that a useful image can be formed with relatively few incident X-rays. This enables them to operate at real time (30 frames/s) rates.

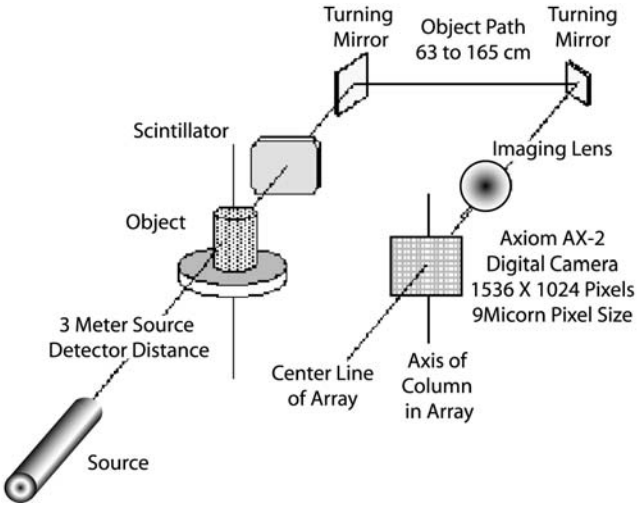


FIGURE 7.31 System using a scintillator to convert X-rays to visible light before being detected by a CCD array. This detector is used for 9 MV LINAC imaging and two mirrors are used to get the CCD camera out of the X-ray beam path and to ensure that the camera is well shielded.

Copyright © Marcel Dekker, Inc. All rights reserved.

Analog Detector (Film)

Film recording of images, representing the differential absorption of radiation in an object, has been the dominant detector since Röntgen's discovery of X-rays. It remains the most common method today. (However, digital detectors, at their current pace of advancements, are expected to be the dominate detector within a few years.) The high spatial resolution combined with large available area of film gives it a distinct advantage over modern digital detectors. The flexibility of film is convenient for conforming to curved objects. Film serves as both data recorder and display. Display is done with simple equipment (e.g., a light box). Finally, film often enjoys a cost advantage.

The disadvantages of film are that extreme measures are required to make images quantitative. Film response is nonlinear and requires expensive developing equipment. Contrast within a film image depends strongly on local exposure so varies with position. It is not possible (unless the image is digitized) to adjust contrast and brightness or to apply image-processing techniques. Film is easily marred by handling and chemical development processing defects. If high quality duplicate images are required, two "originals" must be taken. Film images are awkward to incorporate into reports and publications. Repeated shots are often required to adjust technique after film is processed and dried. Film is bulky (expensive) to store and requires good temperature and humidity control for long-term storage. Film is often lost or misfiled. It is impractical to do CT using film images. Film generates hazardous waste and consumes silver.

The basic structure of film consists of a flexible cellulose or polyester base coated with an emulsion containing the radiation-sensitive material. The emulsion consists of silver halide crystals, typically chloride or bromide, suspended in a gelatin on the order of 0.012-mm thick. When exposed to incident radiation, small portions of the silver halide crystals become sensitized through interaction of the radiation with electrons of the silver halide molecules creating an invisible *latent image* of the projection of the object. The amount of converted material is directly proportional to its *exposure*—the radiation intensity times the duration—to incident radiation. The film containing the latent image is then chemically developed. The developer converts the silver halide crystals to black silver. This chemical action preferentially converts the sensitized crystals to silver. Thus those crystals with high exposures are converted more rapidly than those with low exposure. To halt the developing action from converting all of the crystals and creating a uniformly dark film (*overdeveloped*), the film is submersed in a stop bath that chemically removes the undeveloped silver-halide crystals. The resultant stable, visible image is proportional to the incident radiation with the darker area indicating greater exposure to the radiation than the lighter areas (e.g., see Figs. 7.3 and 7.11).

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

While radiation does expose the film, the degree of interaction of the radiation with the halide crystals—*film efficiency*—is generally very low; for radiation energies of ~80 keV, less than 1% of the incident radiation interacts with the film. To improve the film efficiency, both sides of the film base may be coated with the photosensitive emulsion, the size of the halide crystals can be increased, or the incident radiation is converted to low-energy electrons using a screen or visible light using a scintillator. The film sensitivity is proportional to the size of the halide crystals—the larger the crystal the higher the sensitivity. However, as the crystal size increases, the image resolution decreases due to the *graininess* of the image. Film is very sensitive to visible light created by scintillators and moderately sensitive to low-energy electrons created by screens. As usual, this increase in exposure efficiency comes at a cost of image resolution. All of these methods increase image blur approximately proportional to the increase in efficiency.

Film Characteristics. After the film has been exposed and developed, we inspect the image with visible light. *Photographic density* (D), is a measure of the optical opacity of the film caused by its exposure to radiation. A *densitometer* quantitatively compares the intensity of the light incident onto the film to the light transmitted through the darkened film. In practice the light-beam intensity is measured by a photodetector with and without the presence of the film, i_{Film} and i_0, respectively. Photographic density is defined as

$$D = \log \frac{i_0}{i_{\text{Film}}} \tag{17}$$

Thus $D = 1$ states that 10% if the incident light is transmitted through (90% is blocked by) the darkened image where as $D = 2$ states that 0.1% of the light is transmitted. New unexposed film has a *base density* of $D \approx 0.06$ inherent in the film materials. Additionally, even with the best packaging, the film is exposed over time to cosmic radiation that produces a *fog density* that increases with storage time. Film with a value of $D_{\text{Base+Fog}} > 0.2$ is considered unusable. Both base and fog density decrease the dynamic range of the film.

The difference between the photographic densities of neighboring points is called *radiographic contrast*:

$$C_s = D_1 - D_2 = \log \frac{i_0}{i_{\text{Film}_1}} - \log \frac{i_0}{i_{\text{Film}_2}} = \log \frac{i_{\text{Film}_2}}{i_{\text{Film}_1}} \tag{18}$$

While the radiographic contrast does indicate a relative change in attenuation in the test object, expressed simply as the ratios of optical transmission, it does not

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

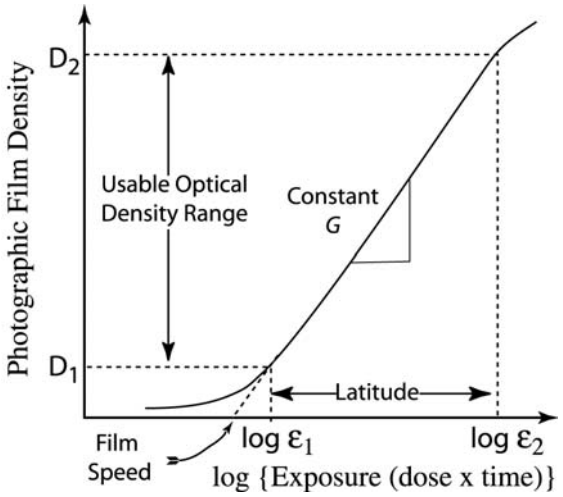


FIGURE 7.32 Characteristic density curve for film radiography.

indicate the dependence on the response of the film. The *film contrast* or *dynamic range** represents the film response—how much it darkens—for a given exposure. Specifically, film contrast is the slope of the photographic density vs. logarithm of the exposure (Figure 7.32):

$$G = \frac{D_1 - D_2}{\log \varepsilon_2 - \log \varepsilon_1} \tag{19}$$

From the definition of exposure ($\varepsilon = It$, intensity $\times$ time), the film contrast can be expressed as

$$G = \frac{D_1 - D_2}{\log \dfrac{I_1}{I_2}} \tag{20}$$

Thus the radiation contrast can written as

$$C_s = D_1 - D_2 = G \log \frac{I_1}{I_2} \tag{21}$$

Exposures levels are taken within the linear region of the film contrast. *X-ray film latitude* describes the range of exposure levels where the film contrast is constant; this is also called *span* or *range* of the film. Thus a film that accommodates a wide range of intensity levels has a large latitude. Note that latitude (span) and

*The dynamic range is specifically defined as the ratio of the largest to the smallest dynamic inputs along the linear (usable) portion of the film response curve.

Copyright © Marcel Dekker, Inc. All rights reserved.

contrast (film or radiation) are inverses: a high span accommodates a large exposure range but at a cost of contrast. *Film speed* is defined as the abscissa intercept of the constant film contrast (Figure 7.32). Film speed is only meaningful for films with equal constant film contrast.

Digital Detectors

Digital detectors, in recent years, have become increasingly common in nearly all applications of radiation detection. For extremely small fields of view, digital detectors have achieved submicrometer spatial resolution, better than commercial film. Over large area, 35.6 cm × 43.2 cm (14 in. × 17 in.), bare (without screen or scintillator) film enjoys a considerable spatial resolution advantage. Digital detectors, with the most significant drawback of higher initial cost, have advantages over film of fast frame rates—the rate at which successive images can be taken—digital storage of the image, data is readily transferred over the Internet, copies of the images are made with no loss of resolution, no environmental concerns from waste products, digital enhancement of the image, minimal concerns for long term storage, and the image is immediately available for viewing.

Digital equipment and conversion from an analog to digital signal requires the introduction of a variety of new definitions and distinctions from analog systems:

Pixels. Unlike analog sensors, the sensing area of a digital array is discretized into separate elements called *pixels*. The pixel includes the active sensing area (*active element*) plus any dead space between active elements. An individual pixel produces a single value for the entire active element, i.e., if the radiation level varies over the active area the output of the element is the average radiation level.

Data (contrast) resolution or sensitivity. When an analog signal is converted to digital format, the resolution of the data is express in terms of the number of binary digits or bits. In digital systems the bits are referenced to base 2, e.g. 8 bit resolution implies that the data is uniformly discretized into 2^8 or 256 bits between high- and low-input limits (dynamic range) of the sensor. A 16-bit signal is broken into 65,536 bits. For analog signals the resolution scale is continuous. The contrast (data) resolution is controlled by the analog to digital converter.

Active element. The active sensing portion of a pixel.

Fill factor. The ratio of the area of the active element to that of the pixel.

Spatial resolution. The minimum resolvable spatial component. In digital systems, spatial resolution is ultimately limited by the size of the pixels and film to be limited by the size of the silver halide crystals. The distinction is that the film resolution is typically below the noise level

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

and therefore practitioners often consider the practical resolution of film to be noise limited.

Frame rate. The number of images per unit time that the detector can acquire.

As in analog detectors, most digital detectors do not discriminate between X-rays of different energies but simply integrate over the source energy spectrum—*energy* (*or current*) *integrating detectors*. However, for special cases, there are energy discrimination detectors that are capable of performing spectrum analysis of the source energy spectrum—*energy discriminating detectors*. In this section, we present energy integrating detectors followed by a specific section of energy discriminating detectors.

Photomultiplier Tubes. Photmultiplier tubes are devices to convert visible and ultraviolet light into an electrical signal with exceeding large gain (as much as a billion). They are used in gauging measurements and in energy-discriminating detectors. Some commercial tubes are available with small arrays ($\sim 8 \times 8$ pixels). They are used primarily in medical applications. The Auger camera used in medicine for single photon emission CT employs multiple photomultiplier tubes as detectors.

Linear Arrays. Of all digital detectors, linear arrays are the most commonly employed for both projection radiography and tomography. The basic technology is depicted in Figure 7.33. The array consists of two rows,

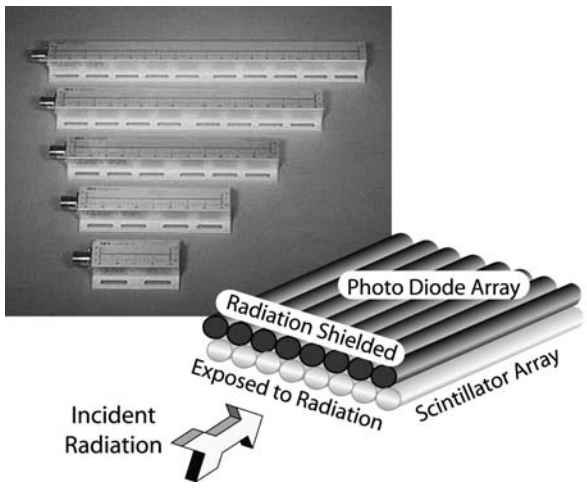


FIGURE 7.33 Diagram and photograph of a photodiode coupled linear array. (Adapted: Photograph courtesy Trixell.)

Copyright © Marcel Dekker, Inc. All rights reserved.

one of scintillator material that converts the radiation to visible light and one row of photodetectors that collect the light. Because the element pair lies side-by-side and parallel to the incident radiation, the radiation-to-light conversion can be increased by simply increasing the length of the element pair. However, while the side-by-side photodiode–scintillator pair increases the detector sensitivity (high energy conversion), the configuration essentially limits the device to a linear array. A similarly constructed two-dimensional array would have too much dead space between the rows where the photodiodes reside. To protect the photodiodes and the detector electronics, the end of the photodiodes are covered with a radiation absorbing material. Slit aperture collimation is typically placed in front of the scintillator and sometimes between each scintillator (septa). Consequently, this collimation on the linear detectors reduces X-ray scatter noise and can be used with relatively high radiation levels.

In typical operation the source and linear array or, more commonly, the object is translated across the detector to create an area scan for a projection radiograph or rotated and reconstructed to create a CT slice or tomograph. The array scan rate is primarily dependent on the detector efficiency, which depends upon several factors including pixel size, source energy, the data (contrast) resolution and object being inspected. A comparison of two commercially available arrays is given in Table 7.4.

Charge Coupled Devices (CCD). Charge coupled devices (CCD) are semiconductor detectors with sensitivity over a broad range of the electromagnetic spectrum—sensitivity from infrared to X-rays. CCD technology is commonly used in everyday camcorder devices. As a radiation detector these devices are configured as single element or, more commonly, in a linear or area array of detector elements. The elements can easily be saturated or damaged by the incident radiation; consequently they are rarely used as direct radiation detectors. Typically, a scintillator is employed to convert the radiation to visible light that is then detected by the CCD camera. For low energies the camera may be placed directly in-line with the incident radiation. To protect the camera

TABLE 7.4 Comparison of Specifications of Two Commercially Available Linear Array Digital Detectors

Sensitive Length in./mm	Sensitive Element Pitch (μm)	# of Sensitive Elements	Minimum Integration Time (ms)	Typical Scan Rates (cm/s)	Data (Contrast) Resolution (parallel)
8/220	27	8192	1	5.5	12 bit
9/230	450	512	0.3	150	8 bit

Copyright © Marcel Dekker, Inc. All rights reserved.

detector elements from the radiation, lead glass is used in the optical elements of the camera. As previously mentioned, the camera can be placed completely out of the line of the primary radiation by optically turning the light produced from the scintillator out of the path of the primary radiation beam. Figure 7.31 depicts such a setup for CT with a 9-MeV electron linac using two turning mirrors to protect the 1536×1024 elements with a 9-μm pixel-size array CCD camera.

Photostimulable (Storage) Phosphors. Photostimulable (or storage) phosphors are sometimes employed as area detectors for NDE work, although their use is more widespread in medicine. In medical applications, these systems are often called "computed radiography" or CR. The operating principle is that sheets of material are manufactured containing compounds that store energy deposited by radiation. This stored energy can be released by exciting the material with visible light, as from a laser. In use, the material is exposed to X-rays thus recording an image. It is then taken to a reader where it is scanned by a focused laser beam that is of a different color than the light released by the storage phosphor. This released light is measured by a light detector, usually a photomultiplier tube that is appropriately filtered to reject scattered light from the scanning laser. The resulting image is digital. The storage phosphor material is then fully restored by exposure to bright light and then reused (38).

Storage phosphors are rarely useful for CT for the same reasons as film, i.e., very slow frame rates. For industrial radiography, storage phosphors offer significantly lower spatial resolution than most other digital detectors and film.

Flat Panel Detectors. Flat panel digital area-array radiation detectors offer rapid detection for both radiographic and tomographic applications. The development of these panels is an extension of the technology used for active-matrix, liquid-crystal flat panel displays that are commonly used in laptop computers. Using *amorphous* thin-film semiconductor technology, high density active-matrix arrays are created with high spatial resolution.

Figure 7.34 shows the basic setup and operation of a flat panel detector. The fundamental component of the array is the detector matrix pixel. The pixel size includes both the active area plus the dead space between elements and dictates the spatial resolution of the detector—typical pixel spacing is 80–400 μm. Each pixel consists of a radiation-to-current conversion element, an electronic charge storage element, and a transistor switch. Initially the thin-film transistor switch is in the off position—no current can flow out of the transistor. After the panel is exposed to the radiation for a period of time, the transistors are turned on allowing the stored charge to be converted to a digital signal that represents the amount of radiation incident onto the individual pixel. The individual pixels are accessed sequentially. Some panels must have the incident radiation turned off for a reading cycle, others read continuously, reading each pixel at the set frame rate. Frame rates depend primarily on the number of pixels and the data (contrast)

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

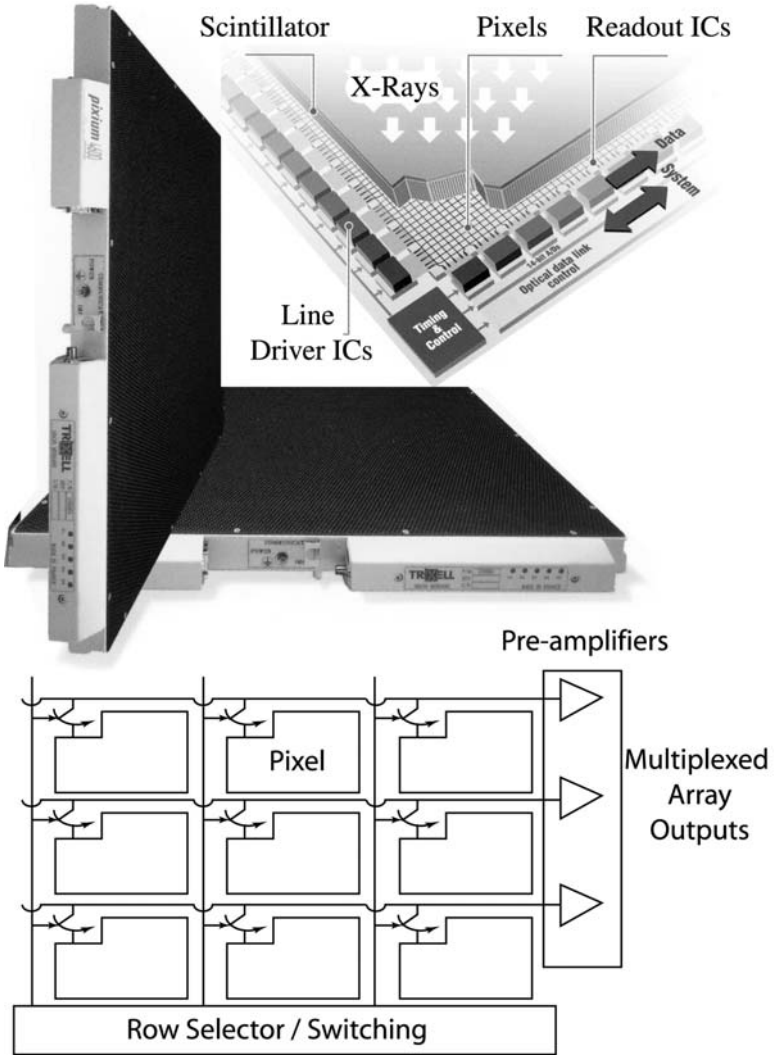


FIGURE 7.34 Flat panel array. (Adapted: Photograph courtesy Trixell.)

resolution, e.g., for a 2240 × 3200 pixel (29 cm × 40 cm) array with 12-bit data resolution, a typical frame rate is ~2.5 seconds.

There are two common methods to convert the radiation into electronic charge used in flat panel arrays: photoconductor and photodiode methods

Copyright © Marcel Dekker, Inc. All rights reserved.

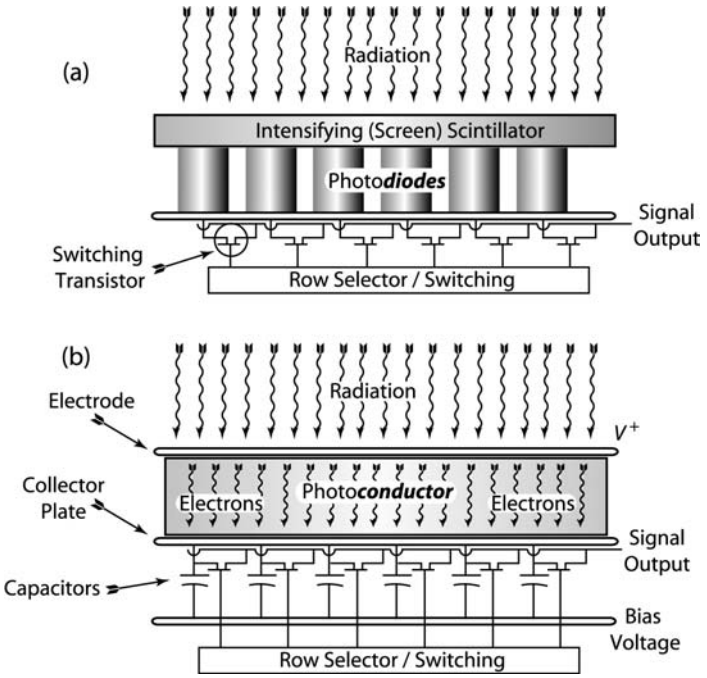


FIGURE 7.35 Flat panel array using the (a) photodiode method and (b) photoconductive method.

(Figure 7.35).* The photoconductor method uses a photoconductive layer such as amorphous selenium (a-Se) attached to a capacitor (Figure 7.35b). The electrical charge generated by the photon interaction with the photoconductive layer is stored in the capacitor until the transistor switch is activated allowing the charge to be digitized. The second and more common method utilizes an intensifying scintillator that converts the radiation photons to visible light that is then captured by amorphous silicon (a-Si) photodiodes (Figure 7.35a). Again, the thin-film transistor controls the pixel read out. A description of an a-Si detector is given by Weisfield, et al. (39). A summary of the performance for an a-Si and a-Se detectors is given elsewhere by Dolan et al. (40) and Logan et al. (41), respectively, for industrial applications.

The transistors in either type of flat panel are exposed to radiation. While the electronics are susceptible to radiation damage, the amorphous Si is

*A third method—*intrinsic*—directly converts the absorbed radiation energy to electron-hole pairs. Currently the technology is limited and generally only available in arrays with one or two lines using crystalline silicon. These detectors are very costly.

Copyright © Marcel Dekker, Inc. All rights reserved.

unaffected by radiation and the thin film structures have been tested to millions of Röntgens without measurable degradation. These thin structures also have essentially no direct response to X- or γ-radiation or electrons.

These large array digital capture detectors with millions of pixels can place significant demands on the data acquisition and storage system—analog-to-digital capture board and the computer speed and storage capability. Having previously stated that the frame rate depends on the number of pixels and the data resolution, ultimately the data storage rate will limit the frame rate and the number of frames. The latter is particularly important for tomographic applications where thousands of frames may be required.

Energy Discriminating Detectors. The transmission of X- and γ-rays as they pass through matter is energy dependent. Since X- and γ-ray sources are polychromatic except for a few radioisotopes and synchrotrons, it is necessary for the most quantitative work to measure the incident and transmitted energy spectra. This requires the detector to have the ability to discriminate between radiation of different energies. Radiation detectors for this type of work usually operate by the collection of electrical charge or by detection of visible light produced by the interaction of an X- and γ-ray with the detector material. Such detectors operate by counting individual detector events (individual photons striking the detector). The energy of the individual photon is converted by the detector to a charge or a voltage indicating its specific energy. The recording of the incident radiation as individual photons, each related to a specific energy, allows the detector electronics to determine the energy spectrum of the incident radiation. This form of X- and γ-ray spectroscopy is frequently called *pulse-height spectroscopy.*

While the pulse height spectrum in a detector is related to the X-ray energy spectrum incident on the detector, there are several factors that complicate the transformation. Even for monoenergetic radiation, there will be variation in the detector response. A complete discussion of these effects is beyond the scope of this text. Knolls (42) offers such a discussion for many detector types.

The most technologically important class of energy-discriminating X-ray detectors is semiconductors—Si, Ge, HgI and CdTe are most often used. When radiation interacts with a semiconductor, electron-hole pairs are produced. The energy expended by the primary radiation to produce an electron-hole pair is found to be nearly independent of the energy and type of radiation producing the effect. This is convenient, for it means that the number of electron-hole pairs produced is proportional to the energy deposited in the active volume of the detector. For semiconductors, the energy required to produce an electron-hole pair (sometimes called the ionization energy) is 3–5 eV. This means that a 100-keV X- or γ-ray stopping in a semiconductor detector will create tens of thousands of electron-hole pairs. This large number results in small relative

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

statistical fluctuation for identical events (leading to good energy resolution) and yields manageable signals for amplification.

Energy-discriminating detector systems require some time to collect and process the radiation-induced charge. Obviously, they must operate in a mode where there is minimal opportunity for two X-ray interactions to occur so close together in time that they are confused. Detector systems may either include the energy from subsequent events with that from the current event, or they may be designed to ignore the second event entirely. Neither effect is desirable. The minimum time separation between two pulses for correct recording is often called the dead time. One effect of dead time is to limit the counting rate that can be effectively employed. For this reason, CT systems based on energy-discriminating detectors tend to be slow, especially if they employ a single detector.

Comparison of Film and Digital Detectors

While using film for image capture and display is most economical in small quantities with hand processing, digital detectors have the advantage for some industrial applications. An example illustrates the issues.

Consider the situation of a manufacturer of Be window assemblies for mammographic X-ray tubes. The central 1.27-cm diameter of the Be must be radiographed with 25-micrometer spatial resolution for inclusions and other flaws that would cast a shadow in a mammogram. The Be window is brazed into a 7.62-cm outer diameter stainless steel ring. The manufacturer desires to inspect after brazing so as to detect flaws introduced by brazing. The windows are produced in lots of 500 on demand. Inspection of a 500-piece lot must be completed within 24 hours. The image of each unit is to be both supplied to the customer and archived for 5 years by the manufacturer.

If done by film, six units can be fit on one 20.32 cm × 25.4 cm (8 in. by 10 in.) film. With double exposures and 10% repeat shots, this results in 200 films per lot. An industrial film processor and dark room is required. Since this will be done at 20-kV, tube potential bare (no screen or scintillator) film is appropriate. If done digitally, a CCD-based system with a 2-cm field of view and a million pixels (1k × 1k) would be the system of choice. This CCD system has a pixel size of 20 micrometers and a resolution of 50 micrometers at the object. This leads to the cost comparison of Table 7.5. These results reveal a ∼50% cost savings for the digital detector as compared to the film detector system. We assume appropriate space is available for either system and ignore permits and other similar issues (sunk costs) that are similar for either approach.

7.3.4 Collimators

Most systems employ some form of a collimator that controls the direction of travel of radiation at either the source or the detector or both. Generically, a

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

TABLE 7.5 Cost Comparison of Film and CCD Radiographic Systems

Capital cost comparison in $ (thousands)		
Item	CCD system	Film system
Dark room	–	15
Film processor	–	30
Cassettes, viewer, etc.	–	10
Source	25	25
Radiation safe enclosure	10	10
CCD/lens/scintillator	25	–
Computer	5	–
Total	65	90
Consumable cost comparison (per window assembly) in $		
Item	CCD system	Film system
Film	–	1
Chemicals	–	0.2
Waste disposal	–	0.1
Labor	2 (1 operator/min)	4 (2 operators/min)
Storage for 5 years	0.5	1
Media to customer	0.5	Included
Total	4/unit	6.3/unit

collimator is a radiation absorbing material with an aperture that allows radiation traveling only in specific directions to pass. On the source, a collimator can be used to create point, line, or area radiation patterns designed to match the detector geometry. On the detector, collimators are generally used to reduce the effect of scattered radiation and to shield electronics from the incident radiation. By reducing the scattered radiation, the measurement sensitivity increases as we approach the approximation that radiation travels in a straight line from the source to the detector. There are two basic types of collimators used to enhance the detector resolution. They are distinguished by their function of reducing out-of-plane radiation or in-plane radiation—the later are called *septa*. Here the reference plane refers to the plane created by the path of the radiation from the source through the object to the detector.

Figure 7.36 depicts a generic collimator (out-of-plane) placed on a detector. The aspect ratio of aperture length to hole diameter dictates how much of the scattered radiation reaches the detector—the larger the ratio, the less noise due to scattered radiation. However, this increase in resolution comes at a cost of a reduced number of photons reaching the detector.

Copyright © Marcel Dekker, Inc. All rights reserved.

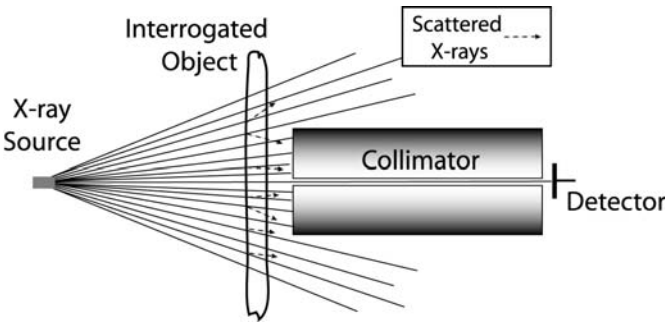


FIGURE 7.36 Collimator used at the detector to reduce out-of-plane scatter.

Septa collimators are used on linear array detectors primarily to reduce in-plane scatter thereby reducing cross-talk between neighboring elements of the array. Septa collimators are typically the shape of the perimeters of an individual element and protrude out of the plane of the *detector*, thus creating a wall of absorbing material between elements.

7.3.5. Stages and Manipulators

Much of radiation inspection requires either the object or the source-detector combination to be translated or rotated. In industrial applications typically the object is moved to preserve the sensitive alignment of the source and detector. However in medical imaging, the object is typically held stationary and the detector-source combination are moved.

For many systems, particularly projection radiography, the object need only be translated in a single dimension. However, for tomographic systems, the object may have to be manipulated linearly in multiple dimensions plus rotations about one or more axes. While for small lightweight objects, precision manipulation is relatively straightforward, with commercially available systems, manipulation of large, odd shaped, objects can be a significant engineering problem. To reduce the amount of manipulation many systems incorporate multiple source-detector units or resort to source-detector manipulation.

7.3.6 Radiological Systems

As we have indicated, the choice of which source-detector and manipulator depends on the requirements of the specific application. More often than not, the choices involve engineering tradeoffs such as spatial resolution vs. speed vs. cost. Generally, the choices for a system can be thought of as a kind of mix-and-match of the components. For example, a CT system will have scan rate as an important

Copyright © Marcel Dekker, Inc. All rights reserved.

parameter. Therefore one might choose a flat panel area array that have very high scan rates and also reduce the required spatial manipulation of the object. However, this choice is at a sacrifice of contract resolution that is offered by a slower scan rate digital linear arrays. In the next section on techniques, we will present common configurations of radiological systems.

7.4 RADIOGRAPHIC TECHNIQUES

Radiation source, object, and detector configurations and settings are driven by the specific application. As indicated in the discussion of the equipment, the radiation system is configured in a sort of mix-and-match method utilizing whichever source, detector, and manipulator most appropriately fits the specific application, i.e., the object to be inspected or characterized. Among the choices, we must first decide the spatial format of the inspection: 1-D gauging, 2-D projection radiography, 2-D CT slice, or 3-D CT imaging. Once this decision has been made, there are numerous controllable variables that affect the quality of the inspection, such as radiation energy level, source and detector spatial resolution, object or source-detector manipulation, and dynamic range of the detector.

For example, if you were to characterize the structure of a walnut you would select a system that is of low-energy and high-spatial resolution. The energy is chosen to produce high contrast in the various materials and components that make up the object. You do not want to select an energy that makes the different materials (e.g., shell and nut) in the walnut all have the same X-ray attenuation since you would not be able to tell the difference between them in the radiographic or tomographic imaging data. You would also want to select an X-ray energy resulting in a reasonable overall sensitivity to differences (contrast) within the walnut. Second you would not select a DR/CT system that has 25-mm spatial resolution to image a walnut that is about 25-mm outer diameter. If you did you would just have one radiographic pixel or CT voxel—volume element that defines the 3-D cube sampled in the object by CT-for the walnut. One pixel or voxel would not be enough to image any structure within the walnut. In other words, you would not see that the walnut has walls a few millimeters thick and there is a nut inside as well as air and membranes.

The appropriate energy to use for DR/CT is determined by the general rule of thumb that the attenuation times path length μl should be about two (43). This results in the selection of the X-ray source energy so that the radiation transmission,

$$T = \frac{I}{I_0} \tag{22}$$

for the maximum attenuation length in the object is about 13%. (I and I_0 represent the detected radiation with and without the object, respectively.) Experience has shown that T in the range of about 10–30% produces high-

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

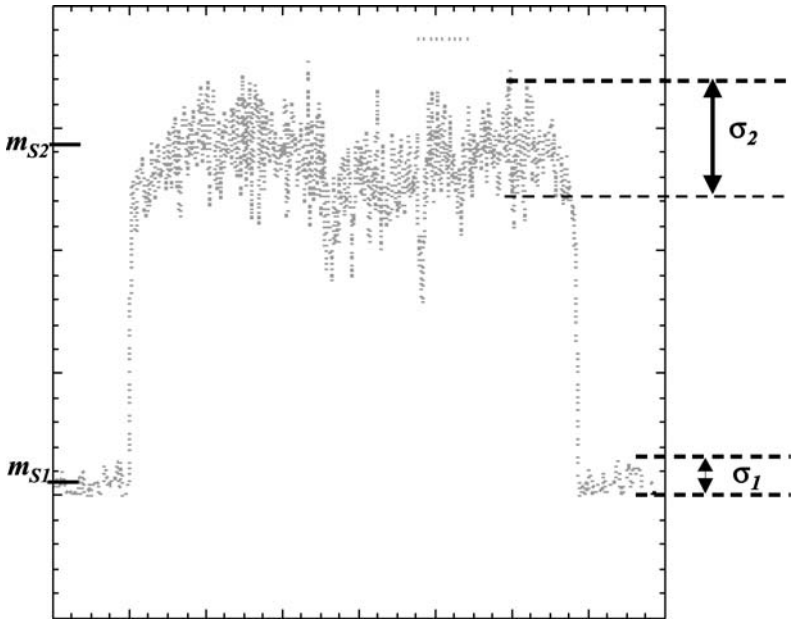


FIGURE 7.37 Schematic showing how the contrast or signal to noise ratio is calculated from a 1-D line-out or profile from CT attenuation data.

contrast images. The X-ray energy is selected to yield the best contrast and signal-to-noise ratio (SNR):

$$SNR = \frac{\Delta S}{\sigma_{\Delta S}} = \frac{|m_{S_2} - m_{S_1}|}{\sqrt{\sigma_1^2 + \sigma_2^2}} \tag{23}$$

where ΔS is the difference between two signals, S_1 and S_2, and m and σ are the mean values and the standard deviations of the signals, respectively (44). Figure 7.37 shows how *SNR* is determined.

A simple method to determine the spatial resolution necessary to image a specific object is based on the Nyquist criterion. The Nyquist criterion is a mathematical theorem that states that a digitally recorded analog signal can be completely reproduced provided the sampling of the analog signal (the incident radiation on the detector) is two times the maximum frequency of the analog signal.* This states that the resolution or element spacing of the digital detector

*For example to completely reproduce music to its original analog form (a requirement to listen to it) from a digitally recorded audio CD, the original (analog) music must be sampled at a frequency of 40 kHz (or greater)—two times the nominal maximum frequency detectable by humans.

Copyright © Marcel Dekker, Inc. All rights reserved.

should be one-half (Nyquist) or better than the spatial size of the structure you are trying to image within the object. Thus to resolve the wall (~1 mm thick) of the walnut a system with less than or equal to 0.5-mm spatial resolution would be necessary.

The ultimate limit of the DR system spatial resolution is determined by the source and detector resolution, the object-to-detector spacing, and whether the beam reaching the detector is a diverging or parallel beam. Thus, to assure the system has the necessary effective resolution to resolve the features of the object—the walnut shell thickness—we need to measure the actual resolution of the entire system. The modulation transfer function (MTF) provides a quantitative method to *measure* the spatial resolution of a given system. The MTF can be determined for either radiographic or tomographic data.

The MTF can be calculated from the point-, line-, or edge-spread functions (45). In practice a radiographicly opaque material (test piece) with either a pin hole, a thin slit, or an edge is used to represent these spread functions. The edge-spread function is most often measured since it is the easiest data to obtain. The edge-spread function method requires a 1-D profile (or line-out) across the edge, the differential is taken of this edge and then the Fourier transform is taken of the differentiated edge. The transform results are normalized to unity and only the positive frequencies are plotted. This is the radiographic MTF and is a convenient experimental procedure for characterizing the spatial response of a radiographic imaging system.

To be an accurate measurement of the detector, the edge test piece should be positioned close to the detector to effectively eliminate source size contributions. It should also be opaque to radiation and not be a source of scattered radiation. It must also be thin or of a large radius to facilitate alignment. While convenient, this method is prone to error and uncertainty regarding the low-frequency portion of the response. This information is carried in the subtle roll-off in signal at large distance from the edge. It is difficult to measure in the presence of noise and effects from nonuniform radiation field. When deriving an MTF in this manner, it is imperative that the line-out and its derivative extend over sufficient distance to capture the low-frequency effects. The results will depend on the lateral extent of the data. If the data do not sufficiently capture the low frequency content, then the accuracy of the derived MTF for higher frequencies will be reduced accordingly.

The ASNT standard tomographic MTF is determined from the edge of a uniform disk (46). Several different radial line-outs are obtained from the outside edge of the reconstructed image of the disk. These are averaged together and the tomographic MTF is determined using the same steps as the radiographic MTF method, i.e., the edge is differentiated and then the Fourier transform is taken of the differentiated edge. The transform results are normalized to unity and only the positive frequencies are plotted. This is useful but it is only representative of the

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

spatial response on the outside of the object. It is more useful to scan a hollow disk and calculate the MTF for the inner edge of the hollow disk. This MTF includes scatter and is more representative of the spatial response of internal features within the object.

7.4.1 Practical Considerations

It is useful to point out some practical tips for radiographic (analog and digital) and tomographic imaging. Analog film radiographs are dark where there are a lot of photons, e.g., for I_0 where there is no object, and light or clear for little to no photons, they are *negative* images (see Figure 7.3 and 7.11). Typically images from digital detectors are displayed as *positive* images, i.e., they are bright where there are a lot of photons, e.g., for I_0, and dark or black for little to no photons (see Figure 7.10). Sometimes this can be confusing when you go from film or analog films that have been digitized to digital detectors.

For all imaging it is important that the object being inspected not move accidentally during the time the image is acquired. If the object moves the resultant image will be blurred. Usually this is not an issue for radiographic imaging with area detectors since the object is not usually moved. Sometimes it can be an issue for radiography if you have to reposition (rotate or elevate) the object to acquire several images, and it is often an issue for tomography, since the object is at a minimum rotated and sometimes elevated to acquire several angular projections and slice planes, respectively. For the latter two cases it is important to have a method to hold the object being inspected without damaging the object.

In order to obtain proper CT angular sampling at the outer edges of the object under inspection you need to have a sufficient number of projection angles. This has been determined to be 1.5 (or $\pi/2$) times the number of ray sums required to image the object (47). In practice, for a digital system the number of ray sums equals the number of detector elements used. For example, for an object that requires 1000 ray sums you should acquire 1500 projections. Usually it is good to be close to this rule, i.e., the range of projections can be from one to two times the number of the ray sums. Fewer may suffice for an object with little detail in the outer portions. The user must be aware that under sampling can result in missing important features.

The radiographic dynamic range should be sufficient to measure all the ray paths within the object being inspected. (The dynamic range for film is determined by the latitude of the film and for digital detectors it is specified by the number of bits in the detector.) In the case of a radiograph where the dynamic range of the detector is not sufficient to cover the range of attenuation within an object (Figure 7.38), the object can be imaged at several different energies (or exposure times) to capture all intensities within the object. For CT imaging this won't work, it is essential to be able to measure all of the ray paths that pass

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

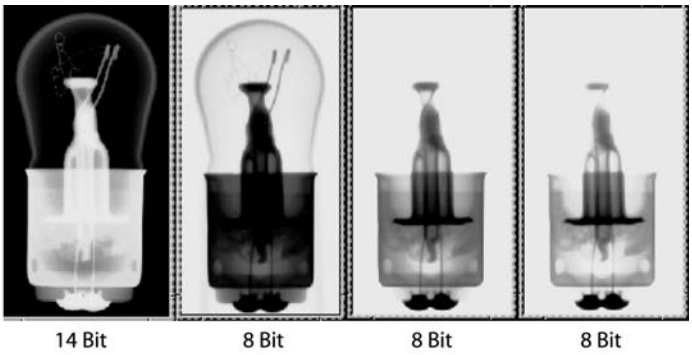


FIGURE 7.38 Comparison of four digital radiogaphs. Two different cameras were uesd to acquire the data. One CCD camera had an 8-bit, the other a 14-bit dynamic range. Three different exposures were required using the 8-bit camera to be able to produce the dynamic range of the 14-bit camera. This shows the usefulness of a large dynamic range camera. The entire dynamic range of the 14-bit image can only be covered using multiple 8-bit exposures.

through the object, I, and those that propagate uninterrupted from the source to the detector, I_0. This is important since before CT image reconstruction you *must* normalize the data to I_0, i.e., I/I_0. You cannot do CT without a quality measure of both I and I_0. An inaccurate measure of either can result in severe CT imaging artifacts. In contrast, radiography does not require specific knowledge of I_0. It is common for the background to be completely saturated (dark in the case of film) as can be seen in the common medical X-ray.

Test objects (or phantoms)* are a vital component of performing quality DR/CT. These *image quality standards* contain known features to evaluate system performance. We cannot underestimate the importance of system characterization through the use of radiographic or CT phantoms (well-known test objects). This is the main means for distinguishing artifacts of the measurement system from actual features within the object.

7.4.2 General Measurement (Detector) Limitations

Before we discuss systems in more detail and present some representative applications of industrial DR and CT, it is useful to point out some general

*Test objects in the medical industry are commonly referred to as *phantoms*.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

detector characteristics that lead to fundamental limitations of the measurement equipment used to capture the projection image in space.

While the entire system must be taken into account to determine the effective limits, the detector is where these measurement limits manifest themselves. We can separate the majority of the sources of limitations into those contributed by source-detector geometry, scattered photons, and the specific characteristics of the detector itself. These limitations vary from system to system, creating performance/system-flexibility/cost trade-offs that must be considered for a particular application. Table 7.6 lists some common trade-offs affecting performance—cost is not included.

Sketches of the three general geometric types of source-detector configurations are shown in Figure 7.39. These configurations illustrate the concept of *source utilization*—the amount of the radiation generated by the source that is received by the detector in the absence of the object. The single- (or point-) detector based systems use a parallel-beam geometry while the linear- and area-array based systems typically use a fan- and cone-beam geometry, respectively. Single-detector and linear-array detectors produce 1-D or line projections that can generate 2-D projection radiographs or 2-D CT slices of an object. Area-array detectors directly acquire 2-D projections to produce the common 2-D projection

TABLE 7.6 Trade-offs that Result from Different Detectors

Detector	N_P/N_S	Detector blur	N_S	Quantum efficiency	Source efficiency
Area array:					
Film	Low	Small	Low	Low	Medium
Flat panel	Low	Small/ medium	Medium	High	High
Camera/scintillator with small cone	Low	Medium/ high	High	Low	High
Camera/image intensifier with large cone	Medium	High	High	High	High
Linear array:					
Slit collimated with septa	High	Small	Low	High	Low
Slit without septa	Medium	Medium/ small	Low	Medium	Low
Single detector:					
Single-detector with spectroscopy	Highest	Smallest	Smallest	Medium	Lowest

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

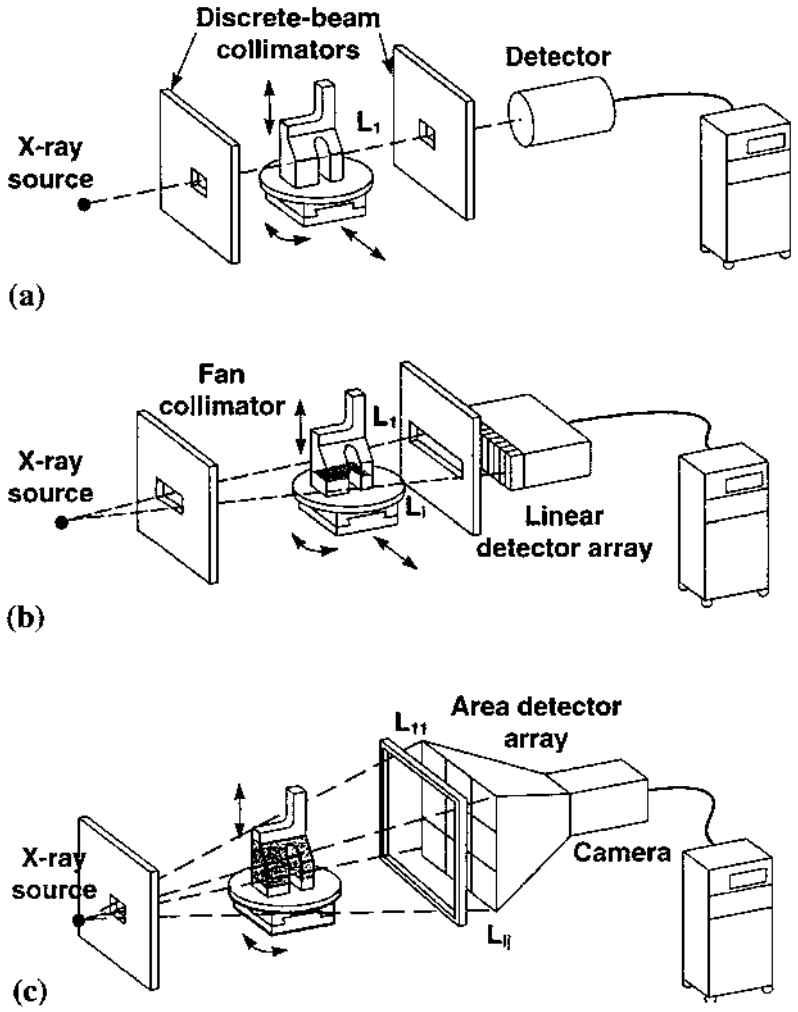


FIGURE 7.39 Data acquisition geometries used in industrial digital radiography and computed tomography imaging. (a) First-generation, discrete-beam translate/rotate scanner configuration using a single detector. (b) Second-generation, well-collimated fan-beam configuration using a linear detector array. (c) Third-generation, cone-beam configuration using an area-array detector. Second- and third-generation scanners alike can acquire data using either fan or cone beams. They differ in that third-generation scanners do not require horizontal translations, acquiring data by rotating the object or source/detector only.

Copyright © Marcel Dekker, Inc. All rights reserved.

radiograph and can be directly reconstructed into 3-D CT data sets. We note that in order to obtain the required beam profiles, much of the total radiation from the source is lost—low source utilization. This is particularly true for the single-point and linear-array detectors (see source efficiency column in Table 7.6). Additionally, for all parallel-beam systems whether point, line, or area, most of the source energy (due to $1/r^2$ loss) does not reach the detector. (Parallel-beam DR has the advantages of ease of image analysis and reduced post processing.) This loss of source energy implies a longer exposure time for a given detector.

For ideal imaging, all rays begin at a *single point* source and are propagated in a straight line to the detector. In doing so the ray interacts with the object and is attenuated according to the absorptive properties of the object. Deviation from this idealization causes a degradation of the image called *blurring*. There are three common sources of image blurring: (a) extended source, (b) scattered photons, and (c) the detector itself. In X-ray Machines in Section 7.3 we discussed the blurring effect caused by a *nonideal* (but realistic) extended source verses a point source. From an extended source multiple rays from different points on the source reach the object and then the detector. Because these rays interact with the object at different angles, a shadowing or blurring of the image occurs. This effect is called unsharpness or penumbra. To reduce this effect, the source size can be decreased and the object should be placed as close as possible to the detector.

Whenever radiation interacts with an object, some of the primary photons are scattered. The process of scattering causes these photons to deviate from the ideal straight-line path from the source to the detector. Because we assume a straight-line path when deciphering the image, these scattered photons effectively carry no information and simply cause a blurring of the image as they are superimposed onto those photons that are not scattered—the primary photons.

Additionally, all detectors have some level of inherent blurring that occurs within the detector itself. This is typically caused by scattered photons that are generated as the incident photons interact with the detector. Systems with low detector blurring incorporate collimators either separate from the detector or septa that are directly incorporated into some digital detectors (Table 7.6) Note that film is the only area detector with low detector blur.

The *radiographic contrast*—the difference in intensities of neighboring pixels on the detector for a given exposure—of an image is strongly affected by the ratio of the number of primary photons to the number of scattered photons (N_P/N_S) that expose the detector. The ratio N_P/N_S can be considered a signal (primary photons)-to-noise (scattered photons) ratio. Thus simply an increase of the radiation from the source will *not* improve image quality. To decrease the image blurring, the number of scattered photons must be reduced relative to the number of primary photons that reach the detector. Therefore, systems involving collimators on the detector side or energy discriminating detectors tend to have higher values of N_P/N_S (Table 7.6).

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

The greater range of choices for industrial systems provides opportunities and tradeoffs for a particular application. For data quality, the best choice is the energy-discriminating or spectroscopy DR/CT based systems (48). This is the experimental analog of imaging with the "primary" photons. Sources of scatter are removed by collimation or from the *energy* resolution in the detector. However, the impact on system speed is dramatic (slower by factors of a thousand). Slit collimated linear array detector (LAD) systems offer the next best alternative for acquiring DR/CT data with the greatest proportion of primary photons (49). The difficulty is when the focus of the application is on obtaining higher-spatial resolution and/or full 3-D inspection data. In this case the use of area-array detectors can improve system speed by a factor of 20. However, all types of area-array detectors include varying amounts of scatter blur that reduces the useful dynamic range of the system (50).

The physical configuration of the detector fundamentally affects its performance. For example, to reduce internal detector scatter some digital systems contain absorbing walls (septa) separating individual detector elements. The fundamental transduction process of the specific detector influences both the image quality as well as its overall usefulness as a detector for a given application. For any detector the speed of data acquisition and the efficiency with which the incident radiation is converted (transduced) to the output signal (quantum efficiency) are important practical detector parameters. Digital detectors have significantly higher quantum efficiencies than film (Table 7.6) Additionally, speed of data acquisition is affected by the source utilization—how much of the actual radiation emanating from the source reaches the detector (Figure 7.39).

The maximum photographic contrast or dynamic range implies an additional limitation on system performance. The *dynamic range*—the ratio of the maximum to the minimum detectable signal—of the system is not simply dictated by the number of bits in the detector (or the latitude of the film), or the number of bits subtracting the read-out noise. Rather, the *effective* dynamic range in a particular area of the image has the limit imposed by the number of bits in the detector (latitude of the film) and then the dynamic range is degraded by readout (detector) noise and background scatter signal in the system. For certain medium energy systems the proportion of background scatter signal can be as high as 80% (for 9MV, it can be 200% or more and it's more than 100% in 28-kV mammography) of the detected fluence, reducing the effective dynamic range greatly. These kinds of considerations are important for system selection and for interpreting images of objects with high scatter fractions.

Depending on the system, different procedures have been developed to correct for some of the effects in the measured signals resulting from the particular system configuration. The goal of these procedures is to process the transmission data into the "ray-path" model, which is the basis for DR imaging and for CT image reconstruction. Industrial DR and CT systems can be organized

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

TABLE 7.7 Preprocessing Steps for Various CT Scanner Configurations

	Scanner Configuration		
Step	Single	Linear	Area
Restoring ray-path geometry			X[a]
Correcting glare and blur			X[a]
Correcting for dark current		X	X
Correcting for detector nonlinearity			X[a]
Balancing detectors		X	X
Normalizing I_0		X	X
Calculating ray sums	X	X	X

[a]This is required to correct for curved area detector arrays.

into just how many processing steps are applied to transmission data prior to reconstruction (Table 7.7). Note that as the detector becomes more complicated, i.e. as it goes from single to linear to area, the extent of the processing steps required to yield good images increases. Thus faster data acquisition comes at the cost of image quality.

7.4.3 Film

The usual geometry used with film imaging is to place the object close to the film, often upon the film. Cabinet radiography units with the source at top shining down are common. This simplifies or eliminates fixturing since gravity holds the object in place. The practitioner must always keep in mind that a film image is a negative image—a void will create a dark spot and the object creates a light spot on the image. Film response to variations in exposure is continuous, with a useful dynamic range (the range where film contrast is high) of about 300. All films eventually saturate (reach maximum film density) at high exposure.

Film photographic density response to variations in exposure is often expressed as a *characteristic curve* (also called *H-D curve*). An example of an H-D curve is shown in Figure 7.40. The slope of the characteristic curve is the film contrast and is usually maximum near a photographic density of 2.0 (Film Characteristic in Section 7.3.3). Thus an optimally exposed film will have a photographic density of 2.0 in the most contrast-critical region. At a minimum the photographic density in the background regions should be 2.0. There are many references that provide exposure charts as a guide to correlate the parameters of the X-ray tube (kV, mA, and exposure time) to radiograph a particular material. These are usually specific to the particular film being used as well as the source,

Copyright © Marcel Dekker, Inc. All rights reserved.

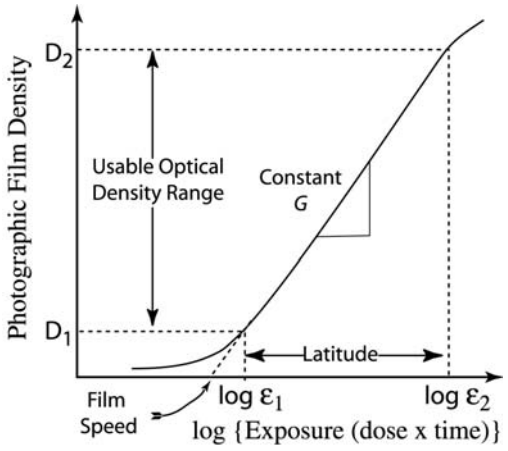


FIGURE 7.40 Characteristic density or H-D curve for film radiography.

source to film distance, and materials within the object. Steel and aluminum are the most common materials for published charts.

Figure 7.41 shows a typical exposure chart for steel. This chart is for film type AKD4303 processed in standard chemistry and exposed with 125-μm thick Pb intensifying screens on both sides, film density of 2.0 with the film at a distance of 1.2 m from the source. The flight path is air. The tube has a 1-mm thick Be window. A 0.3-mm thick Al filter is used. Exposure is stated in terms of

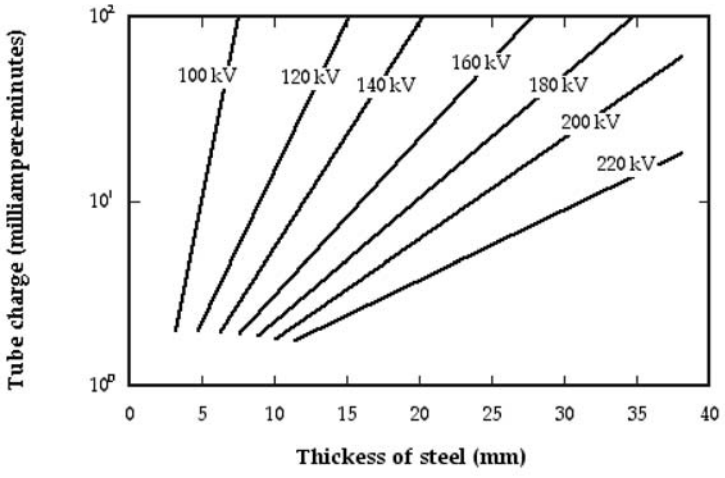


FIGURE 7.41 Typical exposure chart for steel.

Copyright © Marcel Dekker, Inc. All rights reserved.

the charge delivered to the anode in units of milliamperes times minutes. If any different conditions are used, a correction factor must be applied.

Exposure charts are useful starting points, but it is common to adjust technique and take a second or third radiograph. Most experienced radiographers become familiar with their equipment and estimate from experience what initial exposure conditions to use. As in radiography with any detector, it is important to select the source energy so as to achieve reasonable object contrast (film density of ~2).

Test objects, usually called *penetrameters* or *image quality indicators* (IQI), are vital to quality radiography with film. A typical IQI used in the United States is made of the same material as the test object, contains three through holes of diameters 1, 2, and 4 times the thickness of the IQI. The penetrameter thickness d_p is usually 2% of the thickness of the test object. *Equivalent penetrameter sensitivities* (EPS) are given as

$$\text{EPS} = \frac{100}{l}\left(\frac{d_p h}{2}\right)^{1/2} \tag{24}$$

where l and h are the specimen thickness and the diameter of the smallest detectable hole in the penetrameter. Common European IQIs include the *wire-type* and the *step-hole-type*. The wire-type consists of a series of parallel lying wires of varying diameters. The wire diameters range from 0.050 to 3.20 mm with a wire-to-wire spacing of 5 mm. The step-hole-type consists of a series of steps machined in a plate of the same material as the test specimen and each step contains one or more holes of a diameter equal to the thickness of the step. Good practice demands that IQIs be included in every radiograph. Pb characters and symbols are exposed along with the object to produce permanent identification on films. These can often simply be placed on the film.

Processing of film must be done with care and with good quality chemicals. Most film artifacts are introduced during processing. Particulate contamination can produce specks that mimic flaws. Automatic processing nearly always produces linear streaks, visible to the skilled radiographer, arising from mechanical contact of rollers used for transport. Unprocessed film is vulnerable to damage from sparks, especially if used in a dry environment. The resultant artifacts can mask features or be mistaken as actual defects.

Unprocessed film is highly sensitive to storage conditions, especially radiation. A convenient cabinet near the working area is probably not a good storage location. Film must be protected from light. It can be purchased in light tight plastic packages or alternatively loaded into cassettes in a dark room. If screen(s) or scintillator(s) are used, some method of assuring tight physical contact is essential. Any gap amplifies the blurring inherent in use of screens and

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

scintillators. One very effective method is to place film and screen/scintillator in a polymer, which is then evacuated. A quality radiographic facility will have a vacuum system to evacuate cassettes.

An often neglected source of image degradation is backscatter from material behind the film. The best material is none at all. A thin sheet of a low-attenuation material is usually the next best approach. For higher energy, a superior approach may be to use Pb or other high-atomic material to minimize escaping backscatter.

Exposed film should be viewed on a viewer made for the purpose of reading film and is usually done in a dimly lit room. Remember that a 2.0 density film transmits only 1% of the incident light. The light from the a ceiling light or a transparency projector are *not* the right tools for the job.

If exposed film is to be stored for archival purposes, it must first of all be properly processed and washed (there are tests for residual developer chemicals), then stored away from sources of reactive chemicals in a place with stable temperature, darkness and low humidity.

Probably the largest category of film radiographs is welds (Section 7.5.1). Welds can be described as castings made under the worst of conditions, and their integrity is especially suspect. They normally contain voids, inclusions and cracks. Film can be placed inside of pipes and pressure vessels where digital detectors cannot go and it is frequently used in the
field with radioisotopic sources where tube sources and digital detectors may not fit.

7.4.4 Digital Radiography and Computed Tomography

DR and CT are revolutionizing the radiographic imaging business. Radiographic (2-D) analog film imaging over the past 20 years has been extended to digital radiographic 2-D imaging. In digital radiographic imaging the analog film detector is replaced by a digital detector. As previously discussed in detail (Radiation Detectors Section 7.3.3), digital detectors offer the advantage of more direct conversion of X-rays into digital electronic data that can be manipulated by a computer. This has many advantages over analog film imaging since the digital data or image can be enhanced in a computer by many different methods such as contrast enhancement, and low- and high-pass filtering. The disadvantage of DR over analog film is that for large areas (greater than 35 cm by 45 cm) the resultant digital images are typically not as sharp and they can be noisier than those obtained with film. However, new detector technologies such as flat-panel, detector arrays are approaching the spatial resolution of film. Probably the most important aspect of DR is their fast frame rates compared to film. DR images can be viewed in about 1 second, while film typically takes a thousand times longer including processing. Because of this many digital projection images

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

can easily be acquired over several angles about an object, stored in a computer, evaluated or even processed to produce 3-D CT images.

DR and CT imaging system components, e.g., sources and detectors, can be configured in many different ways. Systems are most often distinguished by the type of detector(s) used to obtain the DR data or images. Detectors can be categorized by their spatial extent—point, line, and area—the number of elements or their ability to discriminate or not discriminate the energy of the incoming photons. Energy discriminating detector systems are sometimes referred to as *gamma-ray spectroscopy* or *single photon counting* systems. Non–energy discriminating systems are sometimes referred to as *current integrating* systems. Current integrating detectors are most commonly used in DR and CT imaging. Energy discriminating systems are typically used in research. Both categories are further classified by the number and configuration of the detector elements used to acquire the projection data as shown in Figure 7.39. Examples of energy and spatial resolutions for some industrial and medical DR/CT systems are shown in Figure 7.42.

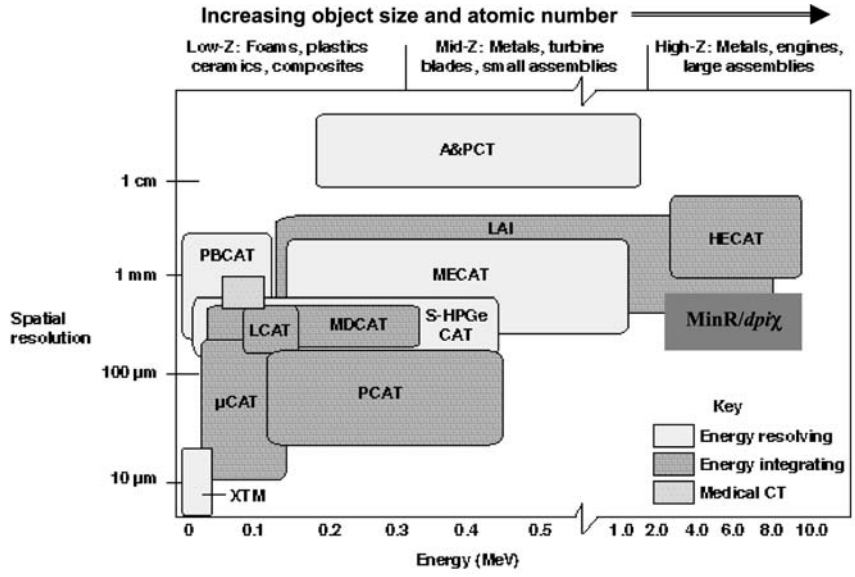


FIGURE 7.42 Schematic drawing showing the many different CT systems as a function of spatial resolution vs. energy. Since industrial imaging applications are varied (see top for of the cart for examples) a wide range of systems is required for industrial imaging compared with medical CT systems.

Copyright © Marcel Dekker, Inc. All rights reserved.

Non-Energy Discriminating (Current Integrating) Detector Systems

Non–energy discriminating gauging, digital radiographic and tomographic imaging systems are based on single, linear-, and area-array detectors. (These detectors are commonly called *current integrating* because they integrate the charge that is generated through the transduction process at a given detector element regardless of the energy of the incident photon.) Even though the best quality (by this we mean quantitative data with little or no artifacts) DR and CT systems are single-detector based, they are slow and rarely used industrially or medically. Therefore, our description of DR/CT systems will be limited to current integrating linear- and area-array detector based systems. X-ray gauging measurements typically are single-detector systems as described next.

Point Detector Systems (Gauging). Gauging is a radiation attenuation measurement along a single line or ray. The configuration for gauging is always source–object–detector and always includes collimation (Figure 7.43). Two applications dominate–quantitative analysis (51), and continuous measurement of sheet products. The radiation source can vary, usually a beta-particle or X- or γ-ray emitting radioisotope, but X-ray tubes are also used.

For complex materials both X-rays and beta-particles are sometimes used. Attenuation of beta-particles depends mainly on the areal density (g/cm^2), Narrow Beam or "Good Geometry" and the Attenuation Coefficient in Section 7.2.3, of the material in the radiation path while attenuation of X- or γ-rays increases strongly with atomic number Z. The information derived from each can

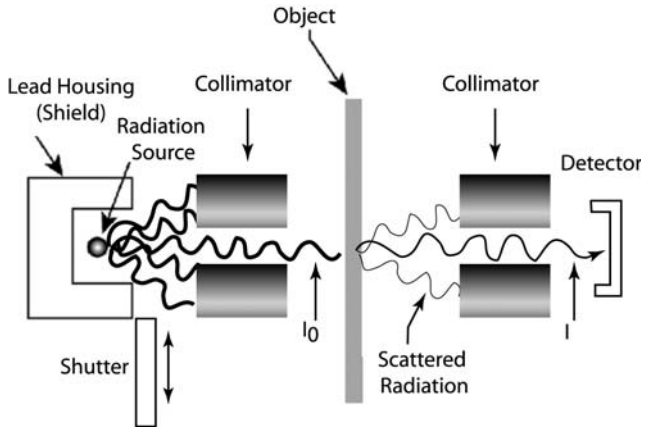


FIGURE 7.43 Schematic of a radiation gauge set up. This method is used for example, to measure sheet steel thickness. I_0 refers to the incident intensity of the beam while I is the transmitted intensity.

Copyright © Marcel Dekker, Inc. All rights reserved.

be complimentary. An example from our experience illustrates an analysis application.

In order to inspect thin sheets of carbon-epoxy composite, a low-energy (10-kV tube potential) radiography capability is used. The flight path for X-rays must be composed of $> 95\%$ He to avoid excessive attenuation and scattering in air. A method was needed to sense the He concentration in air at atmospheric pressure. The solution implemented was an X-ray gauge using the 5.9-keV X-rays from Fe-55. These X-rays are heavily attenuated by air, but relatively unaffected by He. A simple thin-window ion chamber radiation detector was employed and calibrated to read in %He.

Thin sheet products can be measured quickly for thickness uniformity and process control as they pass through the manufacturing process. Low atomic number materials such as plastics and papers generally employ beta sources.

Linear-Array Detector Systems. For data quality, the best choice of current integrating DR/CT systems is based on linear detector arrays. Collimators with a slit aperture on both the source and detector are used to remove out-of-plane scattered photons. Sometimes septa or thin walled collimators are placed between the individual detectors to remove or reduce in-plane scattered photons. Because of this high quality data (high N_P/N_S ratio), most of the medical and industrial based DR and CT systems are linear-array based.

A schematic of a linear-array detector DR/CT system is shown in Figure 7.39. A photograph of a representative linear-array system is shown in Figure 7.44. Most of these systems use an X-ray machine (tube) source. A 160-kV peak potential source and $\sim$1024 element linear array detector systems are most common for both industrial and medical DR and CT imaging. For example carry-on luggage X-ray-screening devices used in airports employ such a source and linear-array detector. There are linear-array systems that use lower and higher energy sources as well. Small spot (or microfocus) size X-ray tube sources are very useful in extending the effective spatial resolution of linear array detectors by the use of X-ray geometric magnification. Sometimes a radioisotope can be used as the linear-array source, but this is very rare mainly because these sources are low in intensity and have safety issue since they cannot be turned off.

The object or source/detector manipulators are usually off-the-shelf equipment. Linear-array radiographic systems need at least a translation stage to move the object between a fixed source/detector to obtain a DR image. For CT in addition to the translation stage, a rotational stage is required to rotate the object or the source/detector to obtain the projections as a function of angle. The former is most common in industrial applications since it is typically easier to rotate the object, while the latter is common for just about every medical CT system since you do not want to rotate the patient. The source, detector, and object or source/detector manipulators are positioned and aligned by fixturing hardware.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

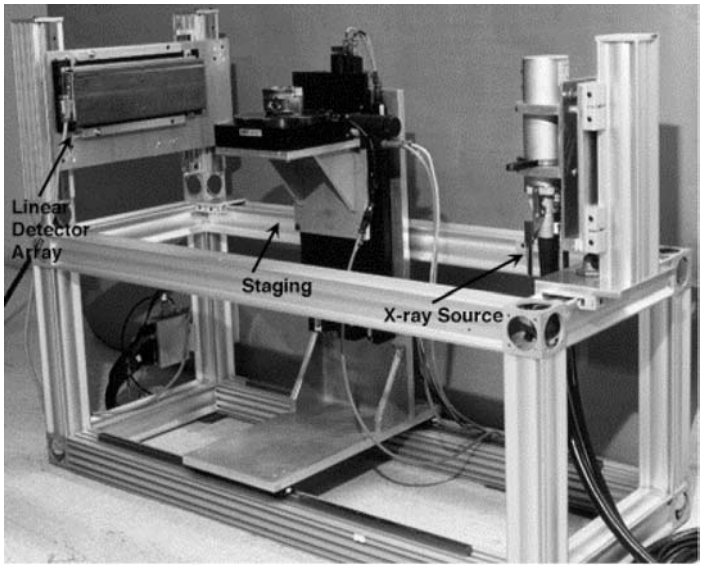


FIGURE 7.44 Photograph of a typical linear-array detector DR/CT system. From left to right are the detector, staging (elevation and rotation), and source. The fixturing structure is used to align the various components.

This alignment is very important to get the best DR and CT data. Mis-alignments or motion error result in skewed DRs and can result in artifacts in the final CT images.

Linear-array detector (LAD) CT systems operate typically in two different modes or what is sometimes referred to as *generations*—second and third generation data acquisition modes. For either LAD CT generation mode, the object or source/detector is typically fixed at one slice-plane and the object or source and detector is rotated to obtain the many 1-D projections needed to reconstruct into a 2-D CT slice. Then the system is positioned to capture data in the next slice plane. In the *second-generation mode* of operation, the object is usually larger than the field of view defined by the X-ray source and detector. In this mode the object is translated across the source/detector and then rotated by the angle that is defined by the source/detector geometry. Thus this mode of operation requires both translation and rotation to obtain the projection data. Typically the second-generation projection data are rebinned* to obtain a simpler

*Rebinned refers to the mathematical conversion from fan-beam second generation data to parallel-beam third generation data.

Copyright © Marcel Dekker, Inc. All rights reserved.

beam geometry data set before image reconstruction. In the *third generation mode* of data acquisition the object fits entirely within the field of view defined by the X-ray source and detector. In this mode, there is no translation (within the slice-plane) required to obtain the projection data for a given slice. The object or source/detector are rotated to obtain the projection data at the required different angles to make up the CT slice. Since LAD uses a *linear array*, both generation modes require the object or source/detector to be translated from one slice-plane to the next to interrogate the entire object.

Typically PCs are used to interface to the X-ray machine source, the object manipulating stages, and detector for radiography projection data acquisition. For DR only a few projections are acquired while for CT hundreds to thousands of projections are acquired. The projection data are processed, and for CT, the radiographs are reconstructed using a PC or workstation computer. Often the DR images are processed to remove imaging artifacts and normalized to the incident intensity in the absence of the object—*field flattening*. It is typical for LAD systems to involve some detector balancing, or detector linearity corrections (Table 7.7). Corrections are made to the projection data to eliminate or reduce these artifacts; however, sometimes they result in other artifacts. Therefore it is preferred to obtain the best projection data and minimize the data processing required to get the best DR and CT images. Normalization to the incident intensity results in attenuation times path length (μl) DR images instead of just intensity images. This is nice since if you know the path length, l, you can determine the attenuation, μ, or vice versa. This is also useful since this normalized data is required for CT image reconstruction.

For CT using a *parallel-beam* LAD system geometry, the projection data needs to be acquired over 180° and a filtered backprojection algorithm is used to reconstruct the projections to yield a 2-D CT slice or image. For a *fan-beam* LAD system geometry, the projection data needs to be obtained over 180° plus the fan angle or is often acquired over 360°. For a 1024 element array, typically at least 1536 (1.5 times the number of ray sums or pixels) projections are acquired. The fan-beam data set is then reconstructed by a convolution backprojection algorithm to yield a 2-D CT slice or image. The object is moved to the next plane within the object to be imaged and a new projection data set is acquired and reconstructed. These additional slices can either be contiguous or spaced at large distances to do a course sampling of the object. The contiguous slices are spaced at the slice plane thickness defined by the detector slit collimator aperture opening. After all the contiguous slices are reconstructed this data can be merged to obtain a 3-D data set of the object being studied.

We describe two different LAD systems—a DR system and a combined DR and CT system—both used to combat smuggler and terrorist activities. These systems are designed to detect drugs, explosives, and firearms. To be successful,

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

these noninvasive systems must have a high-inspection rate and a high-detection probability with low false detection probability.

A mobile DR system used to detect drugs or explosives in cargo containers to subvert smugglers and terrorists activity is shown in Figure 7.45. To combat such threats, it is essential to have an inspection system that is highly flexible and that can provide an element of surprise. For manual inspection, inspectors have to

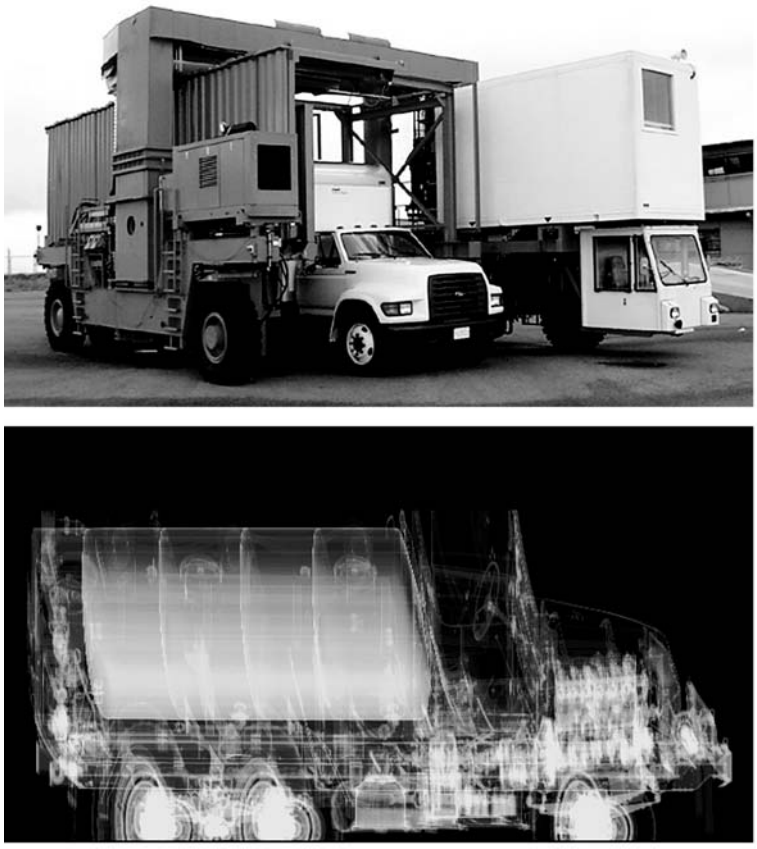


FIGURE 7.45 Top: Photograph of a mobile, fan-beam digital radiography system that has been integrated with a shielding system and is inspecting a truck. Bottom: The DR results of the truck reveal internal details and when analyzed manually by an inspector contraband items can be identified. (Courtesy of ARACOR, Sunnyvale, CA.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

carefully select which cargo to inspect since the unloading, inspection, and reloading of the cargo can take one-to-two person days for large ship containers. Companies have developed self-contained, mobile X-ray inspection systems to meet this need. For example, ARACOR* has developed a fan-beam, linear array, digital radiography system that has been integrated with a shielding system. This system, called Eagle is used to inspect sea and cargo containers, trucks, and rail cars to verify manifests and to detect contraband (Figure 7.45). For inspections, the Eagle straddles the container or vehicle and moves past it illuminating the object with a collimated LINAC beam of 3 or 6 MV X-rays. A linear detector array on the side opposite the X-ray source detects the transmitted X-rays and sends electronic signals to the computer system for the production of an image of the contents as shown in Figure 7.45. At its nominal inspection speed, the Eagle can inspect a 20-ft object in 27 sec, although somewhat faster speeds are possible. The DR images are manually analyzed by a trained operator. There is a concern that stowaways will be irradiated if they are with a cargo ship during X-ray inspection. The "dose to stowaways" for the Eagle has been measured to be about 5 mR if in a container that was inspected by the Eagle. This should be acceptable since it is less than a chest X-ray and only 5% of the allowed human annual radiation dose.

The second system, which combines digital radiography and computed tomography, is routinely used at airports to inspect checked luggage for weapons and explosive devices. These systems need to have a high probability of detection, low false alarm rate, and high throughput to meet the high demand at airports. Several companies have built linear-array-X-ray imaging equipment to "see" inside luggage to find both weapons and bombs (52–54).

A representative DR/CT system built by InVision Technologies† is shown in Figure 7.46. Luggage, inspected one at a time, is first scanned as it passes on a conveyer through a linear-DR system that creates a projection area image and is then scanned by a CT system. The DR system uses a 140-kV and 0.9 mA X-ray source and a 768-element, 12-bit, linear-array detector that produces a DR image with $\sim$1mm $\times$ 1mm pixels of the luggage contents as shown in Figure 7.47. After analysis of the DR image the luggage is scanned by a second linear-array system but this time using CT. This CT system uses a 180kV and 3mA X-ray source and a 480-element, 16-bit, linear-array detector that produces a CT image with a voxel size of $\sim$1.5 mm $\times$ 1.5 mm $\times$ 1.5 mm. The machine automatically sounds an alarm if it finds a bomb. Figure 7.47 shows a CT slice of a bomb highlighted in the image by the dotted box.

*Advanced Research and Applications Corporation, 425 Lakeside Dr., Sunnyvale, CA 94085, http://www.ARACOR.com.

†InVision Technologies, Inc., 7151 Gateway Boulevard, Newark, CA 94560; http://www.invision-tech.com.

Copyright © Marcel Dekker, Inc. All rights reserved.

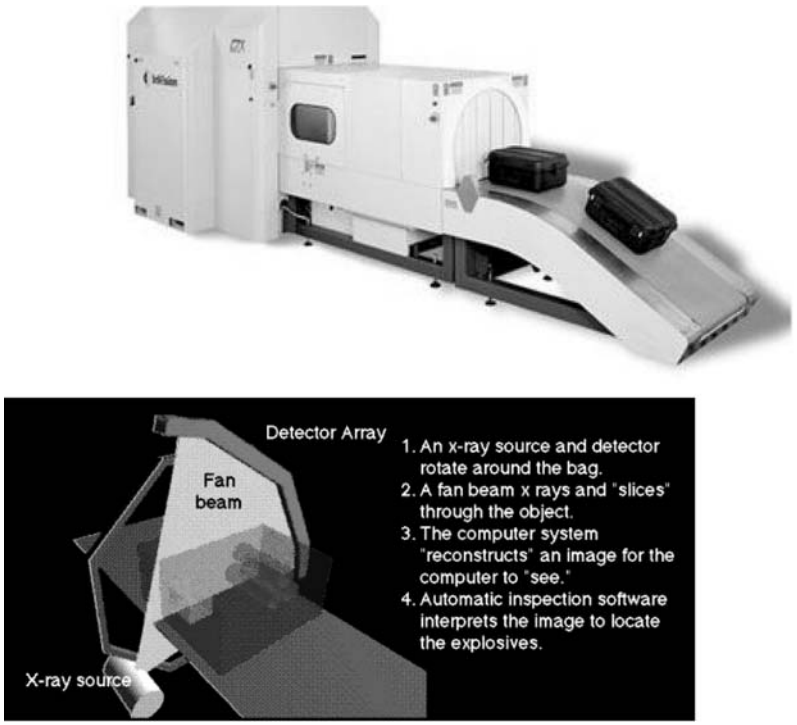
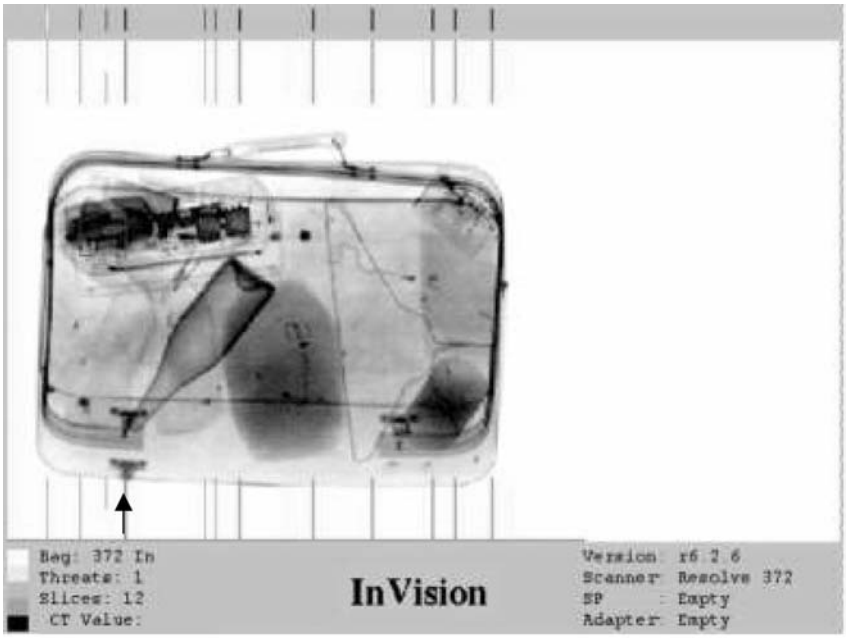


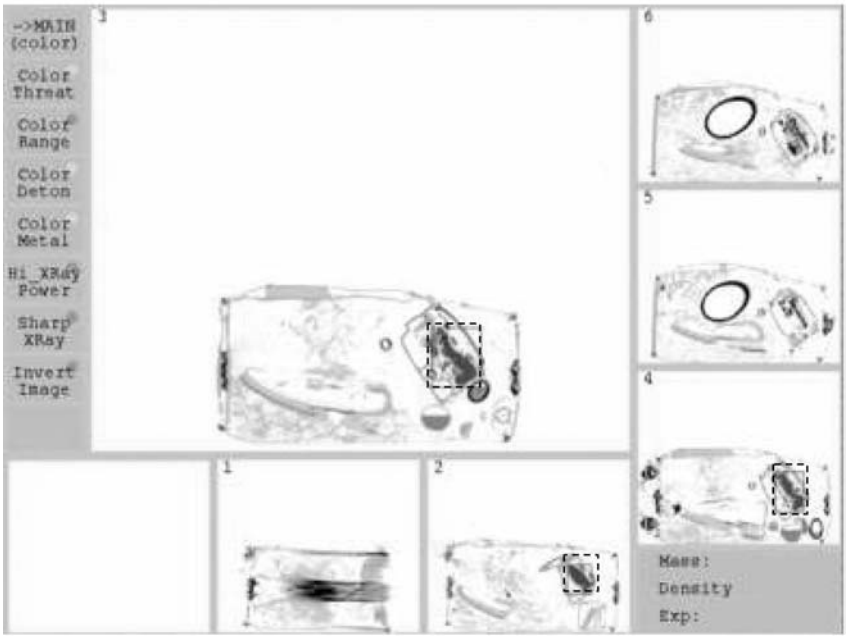
FIGURE 7.46 Top: Photograph of the InVision CTX system used to inspect check airline passenger luggage. Bottom: Schematic showing how the CTX system acquires the projection images. (Courtesy of InVision Technologies, Inc, Newark, CA.)

Area-Array Detector Systems. Area-array based DR and CT systems result in the fastest 2-D radiographs. The high speed is accomplished though high source utilization where an area detector collects data from a large cone of X-rays (Fig. 7.39). It is common to have area arrays with 1000×1000 detector elements or one million detector elements. This is the most efficient use of the X-ray source

FIGURE 7.47 Top: DR image of a test luggage that contains a bomb. From the radiograph it is very difficult to find the bomb. Bottom: Some CT images for the set of lines shown under the luggage in the top DR image. The line with the arrow is the location of the large CT image. The other images along the side are from the lines as labeled. The location of the bomb is automatically highlighted by the dotted box shown in three of the six CT images. (Courtesy of InVision Technologies, Inc., Newark, CA.)

Copyright © Marcel Dekker, Inc. All rights reserved.





Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

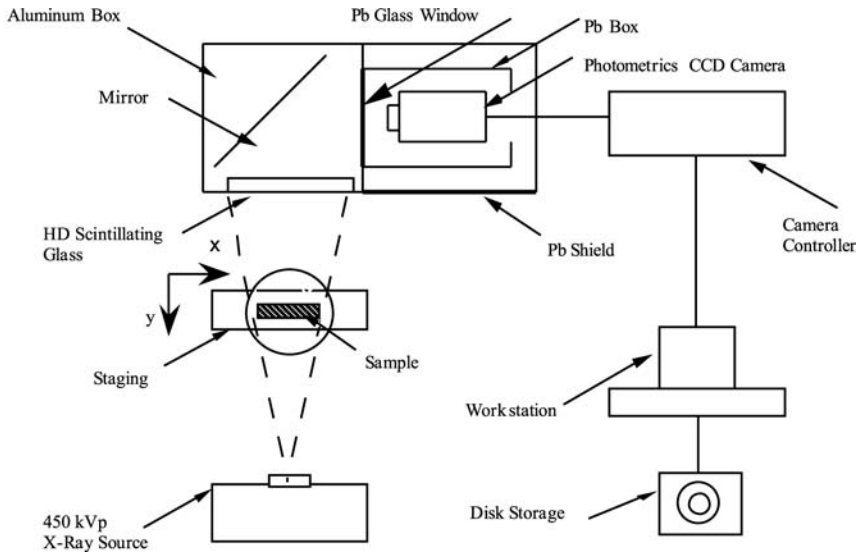


FIGURE 7.48 Top view drawing of a medium energy DR/CT system called PCAT.

photons. Each row of an area-array detector is like a linear-detector array. Thus for a 1000 row area detector it is like 1000 linear detectors. These systems are not as common as the linear-array systems since they have a significant amount of both in- and out-of-plane X-ray scatter. Unlike linear-array systems the in- and out-of-plane scatter cannot be reduced by septa or slit-aperture collimators, respectively. Area-array systems are typically used when digital radiographs and 3-D CT data must be acquired quickly. These systems are used to acquire large data sets (~gigabytes) within 2–4 hours.

Typical area-array (2-D) detector-based systems consist of an X-ray machine source, a detector and staging to manipulate the object. Typical CT systems use a third generation (rotate only) geometry for data acquisition. Area-array detector systems for DR require no manipulation except for initial position. One common area-array detector uses a high-density scintillating glass lens coupled to a thermoelectrically-cooled, astronomy-grade charge coupled device (CCD) camera. A schematic and photograph of such a system, called PCAT,* is shown in Figures 7.48 and 7.49, respectively. Typical staging provides two to four degrees of freedom: rotational, and *x*- and/or *y*-translations and *z*-elevation. (The multiple degrees of freedom are used for initial positioning—only the rotation is

*PCAT or Photometrics Computed Axial Tomography.

Copyright © Marcel Dekker, Inc. All rights reserved.

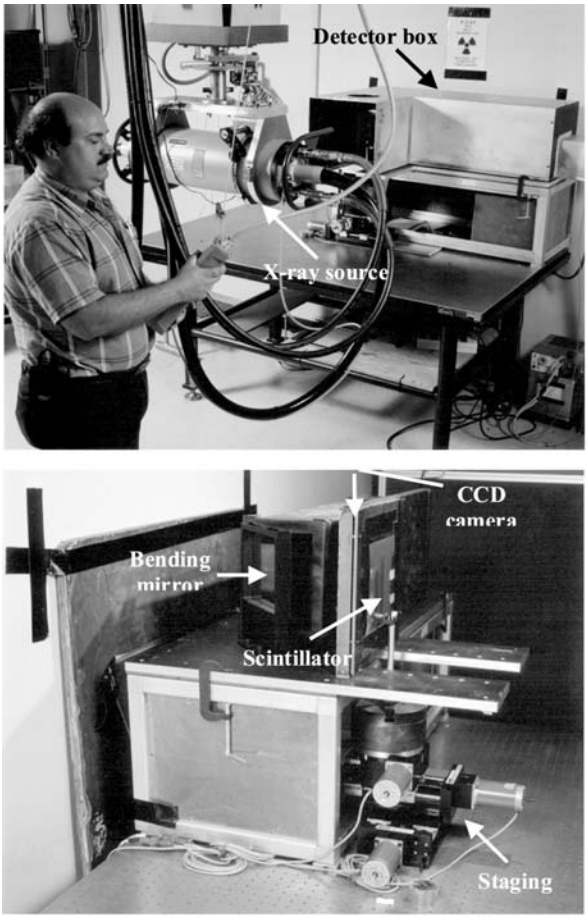


FIGURE 7.49 Top: Photograph of the PCAT system showing the 450-kV X-ray source and detector box. Bottom: Photograph of some of the components within the detector box.

required for CT data acquisition.) These stages are commercially available and are controlled by the same computer used for data acquisition. The source, detector, and object or source/detector manipulators are positioned and aligned by fixturing hardware. This alignment is very important to get the best DR and CT data. Misalignments cause artifacts in the images. Data preprocessing, image reconstruction and analysis are typically done on PCs or workstations.

Area-array detectors involve at least the application of some detector balancing correction and in some cases involve a correction for spatial distortions

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

(Table 7.7). These scanning artifacts (ring artifacts or spatial distortions) can mask features in the object, depending on the location of the feature. Corrections are made to the projection data to eliminate or reduce these artifacts; however, sometimes they cause other artifacts. There is no substitute for good data, free from distortion and accurately aligned.

As an example of the scan data acquired and processing steps needed to create CCD-based tomographic images, we scanned a 2-mm outside diameter, 20-μm thick beryllium spherical shell. These systems require that an image be acquired in the *absence* of incident radiation from the source and in a darkened environment—CCD detectors are sensitive to a very broad spectrum of radiation including visible light. This so-called *dark current* image is a measure of the *thermoelectric noise* created within the camera itself. Typically a dark current image and an *incident* intensity (image acquired in the absence of the object) are obtained before (and possibly after) *transmission* radiographic images (image acquired with the object present) are taken. Figure 7.50 illustrates the response to the dark current D_C and the incident intensity I_0 and their use to improve the

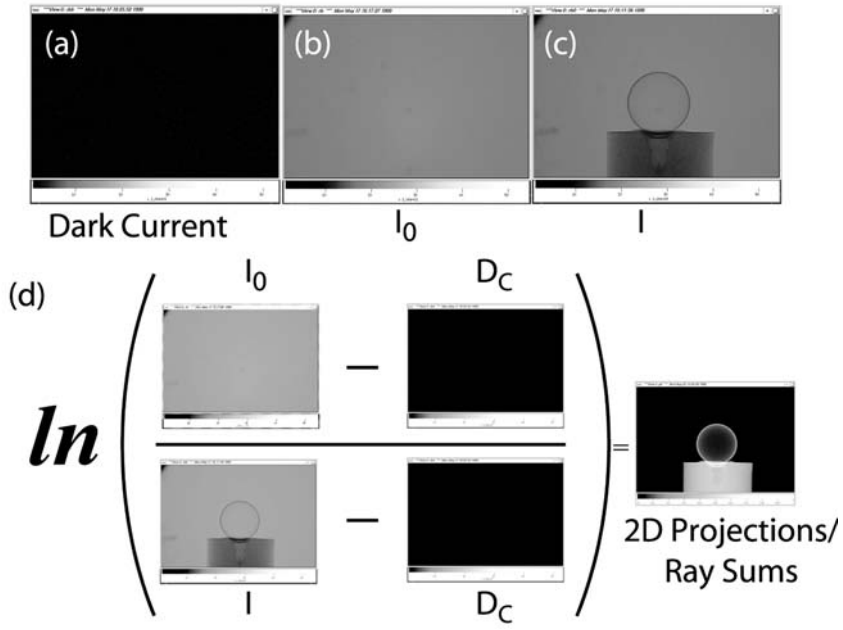


FIGURE 7.50 Representative area array DR images for (a) Dark current, (b) incident intensity, and (c) transmission through a 2-mm o.d. Be hollow sphere of 20-μm wall thickness. (d) These DR images are used to calculate the ray sums used for CT image reconstruction as shown in Figure 7.51.

Copyright © Marcel Dekker, Inc. All rights reserved.

image quality of the transmission radiograph I. Usually only one dark current and one incident intensity image are required while several hundred if not thousands of transmission radiographs are required as a function of rotational angle about the Be sphere. Before the data is reconstructed it is important to remove the dark current, D_C, from all of the I_0 and I projection images or the radiographs as follows:

$$R(\phi, x_r) = -\ln\left[\frac{I - D_c}{I_0 - D_c}\right] = \ln\left[\frac{I_0 - D_c}{I - D_c}\right] = \mu_{Be} l \tag{25}$$

I_0 and I are the digital levels recorded in the detector for the incident (no object) and transmitted intensity images, respectively. (Note for detectors without thermoelectric noise, we simple normalized the data as I/I_0.) This data is processed to create ray sums as in X- and γ-Ray Attenuation in Section 7.2.3 and Computed Tomography (CT) in Section 7.2.6 according to Eq. (6). This subtraction of the dark current and normalization needs to be done on a pixel by pixel basis and is illustrated in Figure 7.50. For a *parallel-beam* area-array system the normalized radiographs are converted into sinograms (image of ray sums verses angle of rotation—Computed Tomography (CT) in Section 7.2.6). Each sinogram is reconstructed using the filtered backprojection algorithm to yield a 2-D CT slice or image (Figure 7.51). The projection data needs to be acquired only over 180° (parallel beam). In this simple example, we did not include all the addition processing steps required to remove bad pixels, beam hardening, rings, and other imaging artifacts.

Cone-beam CT geometry systems have X-ray-source cone angles that are greater than about 5°. For cone-beam area-array system geometry the projection data need to be obtained over 360°. The radiographs are *not* converted into sinograms, they are input directly into the reconstruction process. The cone-beam data set is reconstructed by one of several cone-beam algorithms to yield a 3-D CT image. An overview of cone-beam CT algorithms is provided in a paper written by Smith (55). These images are usually processed to create 2-D slices along any of the three orthogonal planes. Sometimes they are processed along an oblique plane to reveal internal features. Often these 2-D slice data sets are rendered and displayed in 3-D.

Next it is useful to describe in some detail three different, yet common, X-ray source, area-array DR/CT systems. One uses a large (∼2 mm) spot size X-ray machine source and operates in a parallel-beam geometry. The second employs a small (∼10s of μms) or microfocus spot; the third uses an image intensifier to acquire real-time (1/30 sec frame rate) DR images. (The latter two commonly employ cone-beam geometries.) We end this section with a neutron DR/CT system application.

Area-Array Parallel-Beam DR/CT Systems. A photograph of a representative medium-energy, area-array DR/CT system (PCAT) is shown in Figure

Copyright © Marcel Dekker, Inc. All rights reserved.

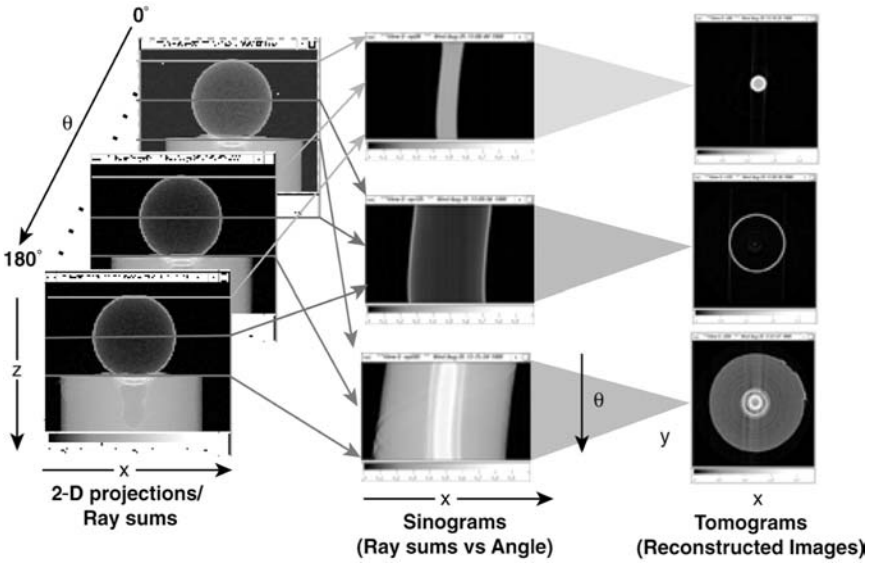


FIGURE 7.51 Three representative angles (0°, 90° and 180°) of several hundred 2-D radiographic projections or ray sums are shown to the left. These 2-D radiographic ray sums are sometimes (mainly for the parallel-beam case) converted to sinograms (ray sums vs. angle for a particular CT slice) as shown in the middle. These sinograms are processed to remove artifacts (e.g., beam hardening and rings) and then reconstructed to provide the CT slice results shown to the right.

7.49. This system employs a 450-kV, 2-mA, 2-mm spot size X-ray machine source. Two stages are used to translate and rotate the object. The detector consists of a high-density, scintillating glass lens coupled to a CCD camera (Figure 7.48). The camera has 1024×1024 detector elements, 14-bits with a 15-μm pitch. Because of the large spot size X-ray source used, this system typically uses a parallel-beam geometry and the data is reconstructed by a filtered backprojection algorithm. Representative DR and CT images are shown in Figures 7.52 and 7.53.

Microfocus DR/CT Systems. The defining characteristic of microfocus systems is its small X-ray spot size, 3–100 micrometers. This small spot size of the source enables the user to position the object some distance away from the detector plane without incurring large source unsharpness, i.e., low blurring. Two benefits result from this geometry; object scatter is reduced and the image is

Copyright © Marcel Dekker, Inc. All rights reserved.

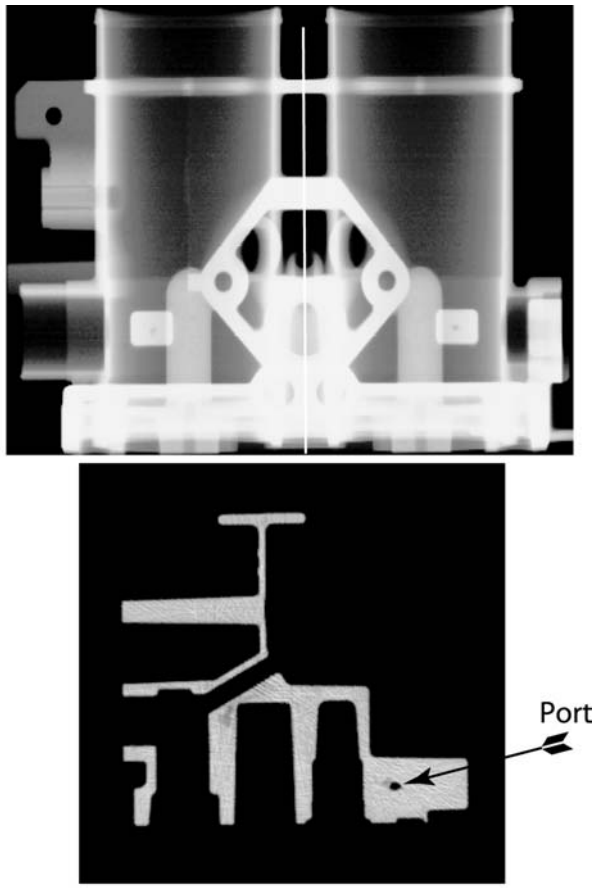


FIGURE 7.52 Top: Digital radiographic image of an automotive water pump. Bottom: CT image along the vertical center of the DR image above. Note that the pore shown in the CT image is not revealed in the DR image due to the complexity of the water pump in this region and superposition of the 3-D image into 2-D.

magnified. Magnification reduces blur induced by the detector. The penalty for small spot size is low source intensity, as described in X-ray Machines in Section 7.3.2.

One representative microfocus system called KCAT* is used for DR/CT of small objects, from 1- to 15-mm maximum chord length (Figure 7.54 and 7.55).

*KCAT or Knolls Computed Axial Tomography.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

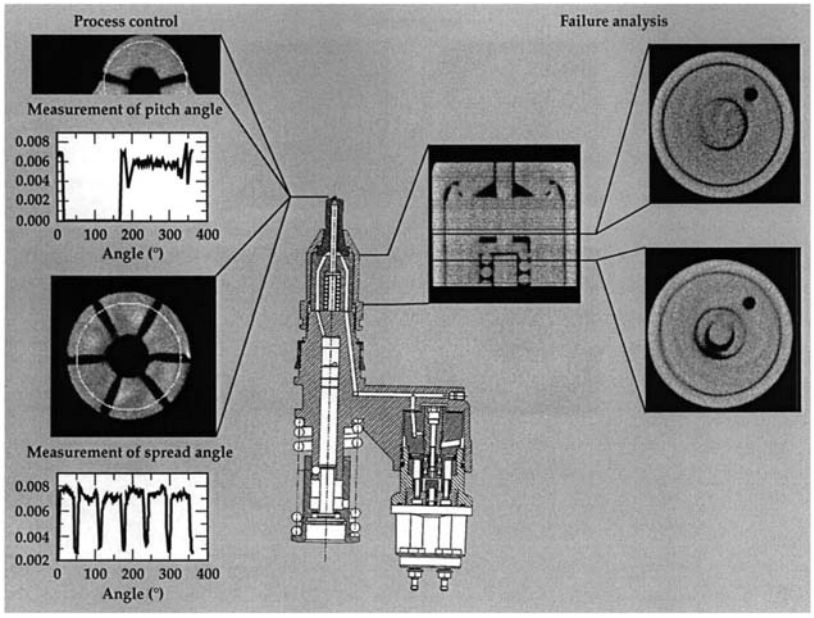


FIGURE 7.53 A fuel injector consists of many components as shown in the schematic drawing. CT was used to measure pitch and spread angle (left) in the nozzle and it was used to determine that this injector failed because the shaft had no clearance as shown in the CT images to the far right.

Usually, the source-to-detector distance is about 100 mm. The source can be operated with an 8-μm source size and the detector has about 9 μm of detector blur. In this situation, maximum spatial resolution is achieved with the object midway between source and detector. This results in an *effective* detector size of 4.5 μm and is also the position for minimum object scatter. The detector for KCAT is very similar to that for PCAT described earlier, except KCAT is an in-line optical system—there is no turning mirror. Operation is administratively limited to 125-kV tube potential to keep CCD damage tolerable. Representative DR and CT images from KCAT are shown in Figures 7.50 and 7.51.

Real-Time DR/CT Systems. The defining characteristic of real-time systems is the ability to acquire radiographs at a rate of 30 frames/s. This provides two significant advantages: (a) the ability to perform fast CT and (b) the ability to detect motion in DR images. For the past several decades these systems have used image intensifiers. Very recently, flat panels area array detectors have become available with comparable frame rates. These digital detectors will likely replace image intensifiers in future systems. An example of real-time DR is a waste drum inspection system (Figure 7.56). For this type of waste, liquids are

Copyright © Marcel Dekker, Inc. All rights reserved.

FIGURE 7.54 Schematic diagram showing the details of the KCAT system used to acquire CT data using a microfocus X-ray source and in-line (no bending mirrors) scintillator lens coupled CCD detector.

Copyright © Marcel Dekker, Inc. All rights reserved.

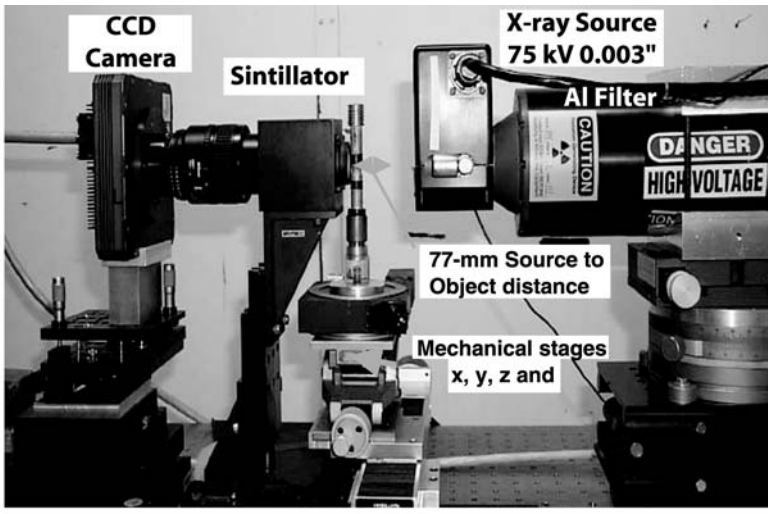


FIGURE 7.55 Photograph of the microfocus KCAT DR/CT system.

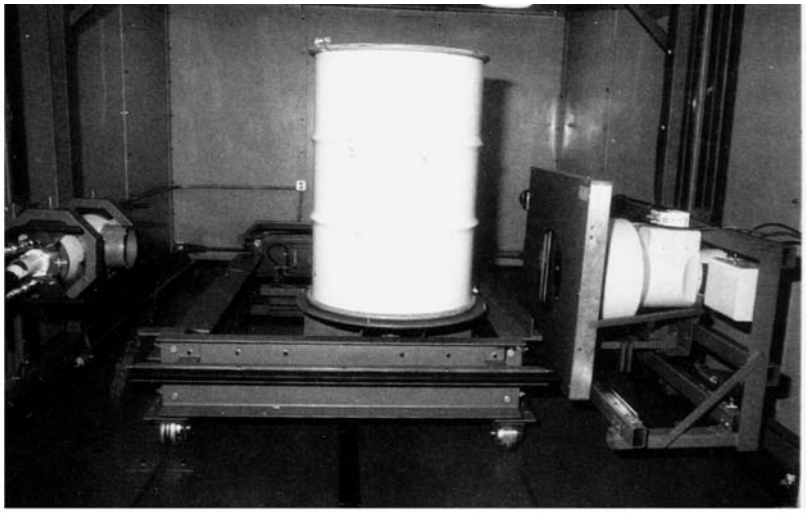


FIGURE 7.56 Photograph of a real time DR imaging system. From left to right are the X-ray source, staging with a drum, detector collimator, and an image intensifier lens coupled to a CCD detector.

Copyright © Marcel Dekker, Inc. All rights reserved.

forbidden. This real-time imaging system enables the operator to inspect drums for evidence of liquids sloshing as the drum is tilted.

Particle DR/CT Systems. Particles such as neutrons (56, 57), electrons, and protons (58) have been used to acquire digital radiographic and computed tomographic images. Here we provide a neutron DR/CT system application. As mentioned in the introduction of this chapter, neutrons can provide information that is inaccessible with X- and γ-rays. The most common neutron energy used for DR and CT is thermal (or near thermal). Thermal neutrons have the property of being heavily attenuated by hydrogen and certain strong absorbers while readily penetrating many common materials. Shields (59) used this attribute to image an elastomeric O-ring seal within a titanium assembly.

A 3-D rendering of the O-ring approximately 2.54-cm outer diameter is shown in Figure 7.57. The O-ring is mounted in a groove around a titanium piston with the O-ring and piston encased in a titanium cylinder. The noticeable kink in

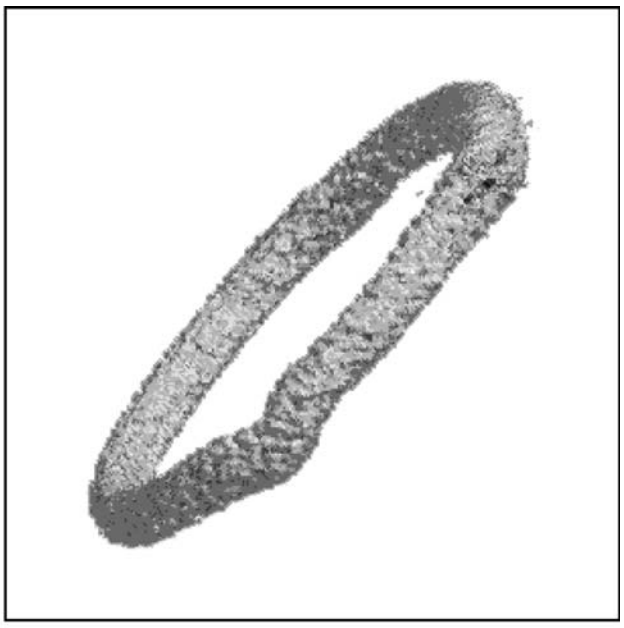


FIGURE 7.57 A solid rendering of an elastomeric O-ring contained within a titanium assembly. The image reveals a kink in the O-ring resulting from improper installation. This data was acquired using a neutron DR/CT imaging system could not be obtained by the use of X-rays. (Courtesy of K. Shields and McClellan Nuclear Radiation Center, McClellan, CA.)

Copyright © Marcel Dekker, Inc. All rights reserved.

the O-ring indicates a faulty O-ring installation. This information would have been very difficult to derive from X- and γ-ray imaging.

The neutron imaging data of the O-ring presented here was created at the McClellan Nuclear Radiation Center. The neutron CT system consists of a 2-MW nuclear reactor neutron source delivering a thermal neutron flux of 9.2×10^6 n/cm^2 sec. The beam was collimated so that the effective ratio of source size to source-detector-distance is 1/140. The neutron scintillator is a LiF-ZnS-Cu NE426 screen. Image data is collected using a CCD camera with a 1024×1024 pixel array and a pixel pitch of 0.1 mm at the object position. Each image was integrated for 90 s.

Energy Discriminating Detector-Based Systems

For data quality, the best imaging choice is the spectroscopy DR/CT based systems (60, 61). Sources of scatter are removed by collimation or from the energy resolution (discrimination) in the detector. However, the impact on system speed is dramatic (factors of a thousand slower than linear and area array-based systems). These systems are used mainly in research and not in day-to-day operations. Most energy discriminating detector-based DR and CT systems use only a single detector. However, there are some cases where multiple detectors are being used (62, 63, 64).

A schematic single-detector DR/CT system is shown in Figure 7.39. Since only *one* ray sum or gauge data point is acquired at a time, the object or source/detector is translated to obtain a single row (1-D) of a radiographic projection. For 2-D DR the object is translated to obtain the next row of the DR image. This is repeated until the entire object or area within the object is obtained—a *raster scan*. For CT the object or source/detector is translated to obtain all the ray sums for a 1-D radiographic projection. Usually each projection has at least one ray sum just outside the object to get a good measure of the incident intensity of the source. Next the object or source/detector is rotated then followed again by translation to acquire the next projection. This takes a long time to acquire all the data.

The source, detector, and object or source/detector manipulators are positioned and aligned by fixturing hardware. This alignment is important to get the best DR and CT data. However the alignment of the single detector DR/CT system is not as difficult as the linear- or area-array systems since it is only one detector and the spatial resolution of these systems is typically not as high. Typically single-detector systems spatial resolution is on the order of a fraction of a millimeter. Still any misalignments will result in skewed DRs and can result in artifacts in the final CT images. If everything is aligned properly only minimal processing (Table 7.7) is needed for single detector spectroscopy systems, since they are the closest physical realization of the ideal "ray-path."

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

The single detector systems are referred to as *first generation*. They are true parallel beam in geometry and the CT data are reconstructed by a filtered backprojection algorithm. Typically these systems use a radioisotopic source, e.g., ^{192}Ir, ^{60}Co, ^{241}Am and the energy discriminating detector is either a high purity germanium or NaI(Tl) (65). The object or source/detector manipulator must be able to translate along the x- and z-axes as well as rotate. PC or workstation computers are typically used for data acquisition, image reconstruction, and analysis.

The quality and usefulness of energy discriminating CT imaging is exemplified in the analysis of a DC motor. A motor contains several materials from plastics to lead. This technique is used to distinguish between the different materials that make up the motor. This may be useful for example if you needed to safely dispose of the motor. In this case the CT data could be used to best determine how to cut the motor into small pieces in an efficient manner to yield some pieces (e.g., aluminum and steel) that can be sorted for recycling and some (e.g., lead) that must be disposed of properly.

Many of these materials are shown in Figure 7.58. This data was acquired with an ^{192}Ir radioisotope and a energy-discriminating (high-purity germanium) detector. The results for the 317-keV gamma ray radiation peak reveal the material that make up the motor. The line-out in Figure 7.58 includes horizontal lines that represent the "theoretical-tabulated" linear attenuation coefficient (66) at 317 keV for several materials as labeled. As expected the housing, bearings and race are made of steel. There are also Al (just around the race) and plastic parts (circuit board) as shown. Another analysis method is to plot a histogram of the

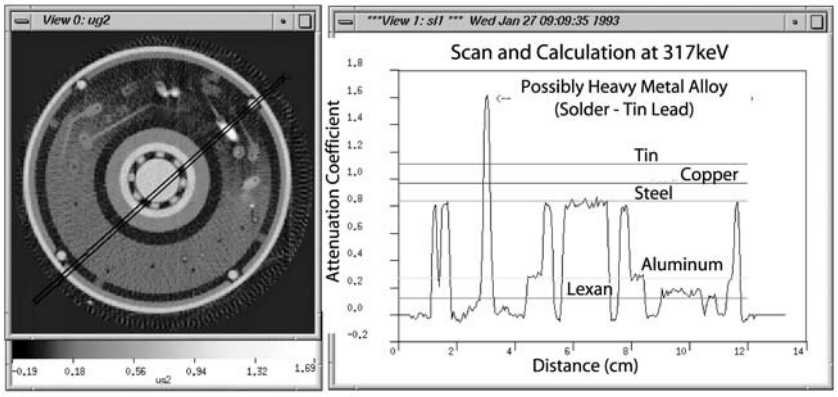


FIGURE 7.58 Representative cross sectional CT slice of the DC motor. Right: Line-out or 1-D profile for the line shown in the CT image. Note the materials contained within this CT image of the motor as labeled.

Copyright © Marcel Dekker, Inc. All rights reserved.

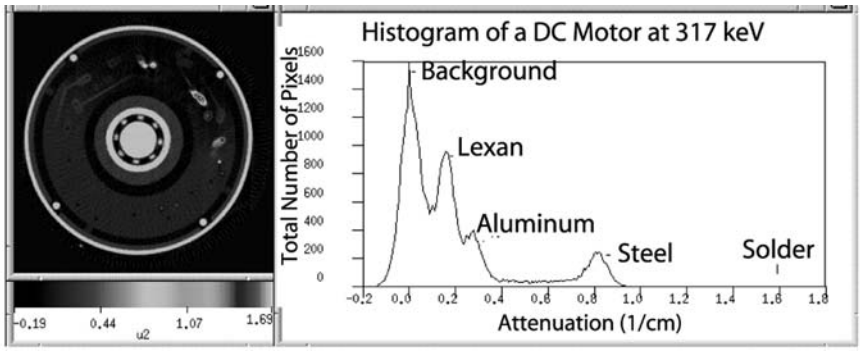


FIGURE 7.59 Cross-sectional CT slice of the DC motor shown in Figure 7.58 but with a different gray scale. Right: Histogram of the CT image. Note the area under the peak is related to the amount of materials contained within this CT image of the motor as labeled.

pixels within the slice plane. Figure 7.59 plots the histogram of the pixels within the slice plane shown in Figure 7.58. This reveals the amount of materials by a simple integration of the total number of voxels under each peak. For mono-energetic CT there is a simple correlation of the "theoretical/tabulated" and CT measured linear attenuation coefficients, which is not the case for polyenergetic sources. However, due to the time required to obtain this pristine data, these DR/CT systems are not often used.

7.4.5 Special Techniques

X-ray Diffraction

Nearly all metals in industrial use are polycrystalline solids. These are made up of individual crystals or *grains* that consist of atoms arranged with a highly ordered periodic structure. When parallel X-rays of a specific wavelength λ (energy) encounter a crystal, they will be diffracted at an angle θ (67) that depends on the atomic spacing d of the crystalline structure in accordance with *Bragg's Law*

$$\eta\lambda = 2d \sin\theta \tag{26}$$

where η is an integer order of refraction that expresses the number of wavelengths in the path differences between rays scattered by adjacent crystalline planes with spacing d (Figure 7.60).

The applications of X-ray diffraction mainly center on its use to identify crystalline spacing and therefore species or phases of a material. Applied or residual stress creates small lattice strain that can be measured using X-ray diffraction.

Copyright © Marcel Dekker, Inc. All rights reserved.

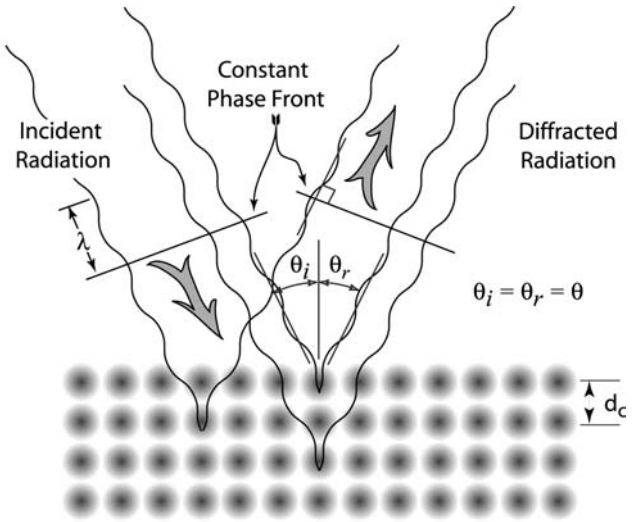


FIGURE 7.60 Diagram indicating the process of X-ray diffraction off rows of atoms in a crystalline material.

X-ray diffraction is not always classified as a nondestructive technique because penetration depth is slight and considerable sample preparation may be required.

X-ray Fluorescence (XRF)

In Electrons and Radioactivity in Section 7.2.2, we described how atoms emit characteristic X-rays as the atoms transition from an excited state to the ground state. Since the energies of these X-rays are characteristic of the element under transition, they can be used for elemental analysis. X-rays are used to excite the specimen, and the emitted characteristic X-rays are analyzed. Spectrometers for XRF may be either wavelength dispersive or energy dispersive. In a wavelength dispersive spectrometer, a single crystal of known spacing is used to disperse a polychromatic beam so that each energy appears at a different angle in accordance with Bragg's law. In an energy dispersive spectrometer, the detector and associated electronics determines energy by ionization produced in a semiconductor detector, usually lithium-drifted silicon.

Because air, detector windows, and the sample heavily attenuate characteristic X-rays from light elements, XRF is not generally useful for elements with atomic number less than ten. Sensitivity is typically $\sim$10 ppm for mid-Z elements. The technique requires extensive correction for absorption and specimen (matrix) effects, but these methods are well understood.

Copyright © Marcel Dekker, Inc. All rights reserved.

In the context of NDE, XRF is most often used as a surface probe to identify or sort unknown materials and to measure plating or coating thickness.

Depth Informing Techniques (other than CT)

There are many methods to gain depth information in an X-ray measurement. The simplest method is to take two or more X-ray images from different sides of the part. Two images taken at right angles from each other provide some limited depth information. There are many other techniques for obtaining depth information varying from stereo imaging to the use of a laminographic method that permits one to view the X-ray image of a specific internal plane while all other planes are blurred. More complex methods such as CT permit one to view many internal images and to construct a 3-D representation of the object in which all planes are in focus, i.e., no planes are blurred.

Stereo imaging of X-ray images requires the preparation of two radiographs taken after shifting the X-ray source or object between views (68). Figure 7.61 illustrates the shift of the X-ray tube approach and, in the lower view, the arrangement for viewing. The movement of the source in this simple arrangement should be the interocular distance, about 6–7 cm. The stereo views can be obtained with any area X-ray image device. Note that if a microfocus source and geometric magnification are used, the stereo image can be obtained with a much smaller shift of the X-ray source. This permits a real-time, electronic X-ray presentation, by shifting the X-ray focal spot electronically, showing each image as a separate field of the TV frame and viewing the monitor with special glasses that permit each eye to see the appropriate field (69). A novel X-ray image technique called *reverse geometry* also permits stereo viewing (70). In this case the X-ray source is scanned across the object and the detector is a point detector that can be placed in tight spaces. If two detectors are used, as shown in Figure 7.62, the observer again has two X-ray views from different perspectives and the opportunity for a stereo image.

Laminography makes use of a particular form of motion between the source and the detector to blur the radiographic image data for all but one plane. If one imagines a source and detector attached to opposite ends of a lever and the lever pinned somewhere along its length such that if the source moves North, the detector moves South, then if an exposure is made while the source and detector are both moved in this fashion, only the plane containing the lever pin will be in complete focus; this is illustrated in Figure 7.63. This technique is widely used in the electronics industry to image particular planes of circuit boards. If a new plane of focus is desired, a new exam must be made. These 3-D image techniques are discussed in a review article (71).

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

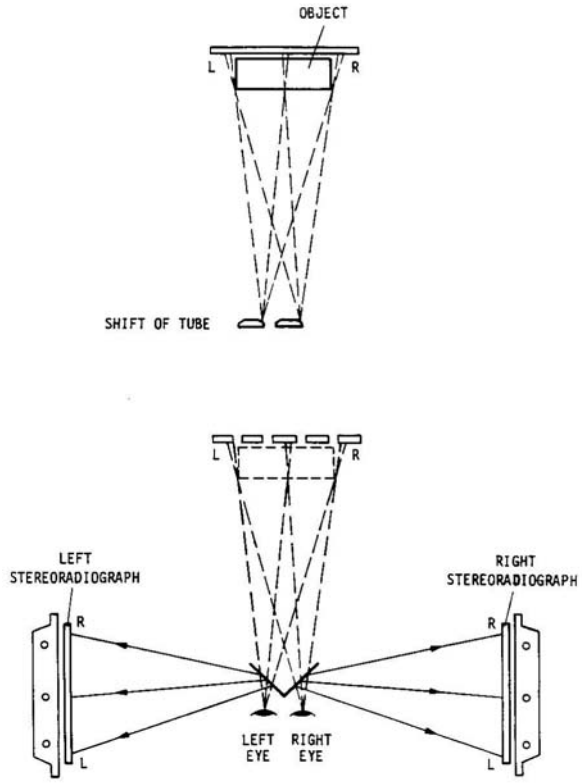


FIGURE 7.61 Exposure and viewing configuration for stereo radiography.

*Tomosynthesis** is the general term for the concept of accumulating radiographic information in specific pixels, corresponding to the projection of the same point in the object. The tomosynthesis technique is generally attributed to Ziedses des Plantes (72). The approach has been widely used in medical radiology (73, 74) but is now being used for industrial applications (75, 76, 77). Modern tomosynthesis is a technique that utilizes a computer process to simulate the laminographic process. Where laminography uses continuous motion to blur out-of-plane information, tomosynthesis uses a series of discrete source positions to produce a family of images to simulate or synthesize the laminographic process. Picture the source and detector being moved to discrete

*Computed tomography takes its name from the same source as tomosythesis, but the two methods use different approaches to the collection and analysis of the data.

Copyright © Marcel Dekker, Inc. All rights reserved.

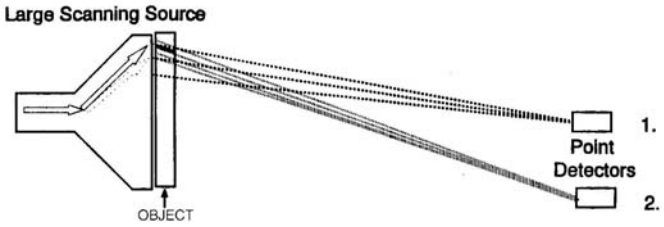


FIGURE 7.62 Reverse geometry X-ray imaging.

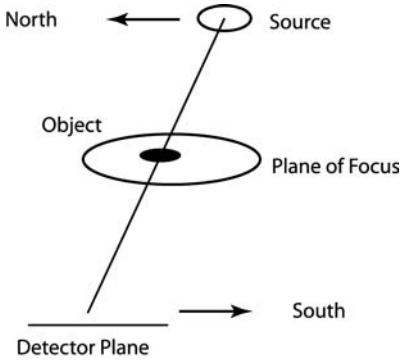


FIGURE 7.63 Movement arrangements for linear laminography.

locations along the path shown in Figure 7.63 and a radiographic image made at each location. The resulting radiographs could be viewed simultaneously in such a way that the plane of focus as above would be reinforced, or by using a different set of translations, a different plane of focus could be obtained. This technique can also be used with other patterns of source positioning, including circular and random, so long as the source position is known or can be determined for each image. Systems have been developed to acquire digital radiographic images and perform the plane of focus or "slice" calculations in a computer, permitting the user to dial through the thickness of an object.

7.5 SELECTED APPLICATIONS

The most common application for radiation NDE methods is detection of internal structure. Several X-ray NDE applications have been selected to help elucidate when and what X-ray imaging technique is best. Typically radiography is the best choice for planar types of objects although it is commonly used for a wide variety

Copyright © Marcel Dekker, Inc. All rights reserved.

of geometries. As examples of this NDE radiographic application we discuss the traditional inspection of welds used to join two plates together. Another pertinent example is the inspection of beryllium sheets for metal inclusions. Film radiography methods were used for both of these applications. There are times when the collapsing or superposition of a 3-D object onto a 2-D radiographic image is not sufficient to measure internal dimensions or material characterization and inspection of internal features. Inspection of bridge pins is used to show the application of combined techniques of digital radiography and computed tomography. One very good example where three-dimensional imaging is necessary is the use of CT for reverse engineering in prosthetic implant design. Lastly we present an application of a single detector energy-discriminating CT system to characterize explosive materials.

7.5.1 Traditional Film (Projection) Radiography

The most common application for X-ray radiography is the inspection of internal structure. It is especially applicable to finding voids, inclusions and open cracks and is often the method of choice for verification of internal assembly details. Single view radiography is *best* when applied to objects that are planar in extent. Additionally for simple feature detection (not true sizing) the maximum detection probability occurs when the feature's largest dimension is *parallel* to the direction of radiation propagation. Examples of planar objects inspected by radiographic methods include plates, sheet products, welds, adhesive joints, and electronic circuit boards. When the object being inspected becomes more three-dimensional in extent it becomes more difficult to use just a single radiographic image. For example in Figure 7.64 in one view it is impossible to determine the number of fingers in the hand while the 90° view of the hand clearly reveals the number of fingers on this person's hand. Therefore there are times when a few radiographic views can be used to adequately perform the NDE inspection.

Weld Inspection Using Film Radiography

Film radiography has been used for years to inspect welds in plates, pipes, etc. Here it is useful to show one example of weld inspection by film radiography. A 250 kV and 10 mA X-ray source was used to acquire several film radiographs of flat butt welds. Digitized images from film radiographs of some ASTM standards of two $\frac{3}{4}$-inch steel plates welded together using metal-arc welding are shown in Figure 7.65. The radiographic images reveal a suite of porosity, cracks, incomplete fusion, and undercutting defects in the flat butt weld. These standards are used to train radiographic weld inspectors. Inspection of pipe welds is a little more difficult due to their more complex geometry. Pipe welds are inspected by acquiring many radiographs at different angles around the pipe or sometimes an annular ring emitting X-ray machine source is inserted inside the pipe and film is

Copyright © Marcel Dekker, Inc. All rights reserved.



FIGURE 7.64 Digital radiographs of a women's hand at two angles, 0° (left) and 90° (right). In these radiographs both the hard (bone) and soft (flesh) tissues are observable.

wrapped around the outside. This technique helps reduce superposition and thus interpretation of the radiographic images. Another method used with pipe inspection is to position the radiographic direction at sufficient angle to the pipe axis to separate the near and far side welds. This eliminates *weld* superposition, but requires penetration of two pipe wall thicknesses.

Often, structures are designed with little consideration of inspectibility. In some cases, other design concerns are paramount, even when inspection is considered. The authors have experience with one instance where steel containers for a dangerous material were designed with a stepped weld joint for one flange. This design was fixed by code and regulatory approvals, so that no option existed except to devise a specialized method of radiographic inspection. The joint design of the 5-inch OD vessel is shown schematically in Figure 7.66.

The only method that could achieve the required sensitivity to weld defects was to destructively section the container and radiograph the joint at an angle through a single wall as shown in Figure 7.67. To assure complete circumferential coverage, five exposures are required. In this instance the specification developed required the use of lead shot to mark the limits of each segment.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

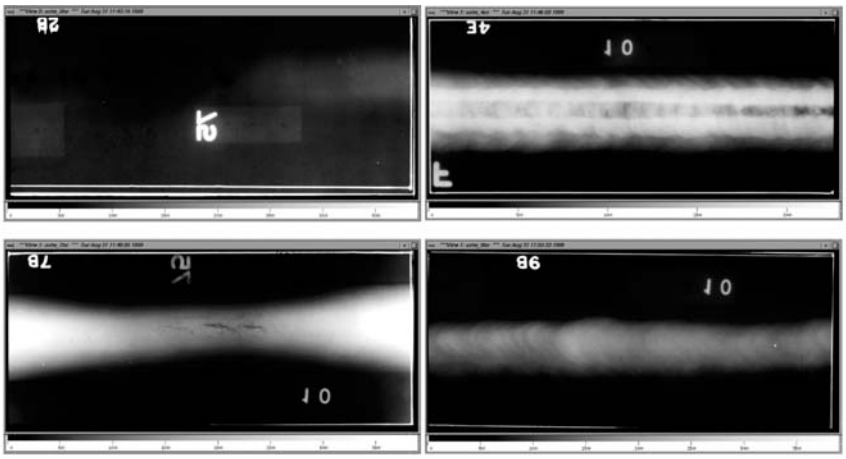


FIGURE 7.65 Digitized film radiographs showing common defects in four different flat butt welds. Upper left: the dark spots reveal coarse porosity (gas). Upper right: Dark central portion within the bright area reveals an incomplete fusion. Lower left: dark lines indicate long cracks in the weld. Lower right: Dark edges along the weld indicate under cutting. (Courtesy of the American Society of Testing and Materials.)

The alert student might question whether this is a nondestructive method in view of the fact that the part is cut up in order to do the inspection. That student would be correct. In this case, because the object could not be inspected and then returned to service, NDE methods had to be used to certify the *process*, not the actual component. Because of the potential consequences of weld failure, two vessels were made for every desired unit, with every detail of the weld held constant, especially the welder. One vessel is destroyed and radiographed, while the other one is placed in service. (Note this example represents an extreme inspection case.)

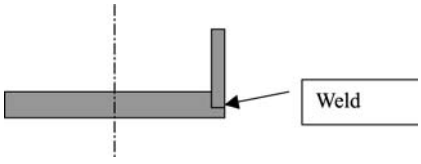


FIGURE 7.66 Schematic of a stepped weld joint design for a 5-inch outer diameter steel container.

Copyright © Marcel Dekker, Inc. All rights reserved.

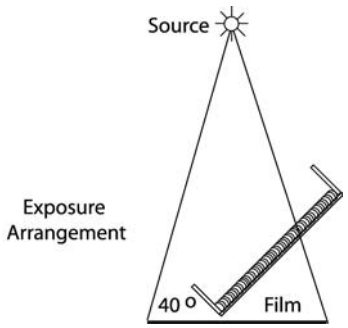


FIGURE 7.67 Schematic showing that the only method that could achieve the required sensitivity to weld defects was to destructively section the container and radiograph the joint at an angle through a single wall.

Another radiographic application is the detection of inclusions in beryllium sheets. Figure 7.68, from a film radiograph, shows a portion of a Be sheet that will be cut into disks for X-ray tube windows. The Be is on the left side of the image, while the right is bare film. Beryllium is the usual material for tube windows due to its low X-ray attenuation cross section and high strength. The thin X-ray window is intended to allow X-rays to exit the tube with minimal attenuation while maintaining the vacuum in the X-ray tube. The inclusion shown near the lower left of the image is highly attenuating and would create an objectionable shadow if it were used to manufacture a tube window.

In order to capture the image of Figure 7.68, a low tube potential of 7.5-kV was used. This provides reasonable contrast in the Be material. If air had been used for the space between tube and the Be sheet, the air would have dominated X-ray attenuation, so a He-filled enclosure was employed. Normal commercial film packaging contains paper that is imaged very effectively at this energy. This paper pattern would obliterate the desired information in the radiograph, so handmade packages were made of thin black polyethylene. The black blob in the upper right of the image is a spot where the polyethylene had a pinhole allowing light to leak in and strike the film. Another type of film artifact is visible on the original image, several subtle hortizontal streaks that extend through both the area of the Be sheet and the bare film area. These are caused by mechanical handling in the film processor.

7.5.2 Digital (Projection) Radiography and Computed Tomography

There are times when a few radiographic views are not sufficient to adequately perform the NDE inspection. This is typical for objects that are very complex in

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

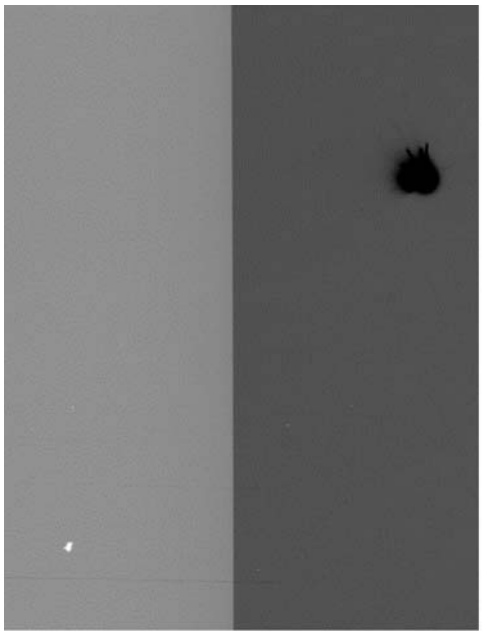


FIGURE 7.68 Digitized image taken from a film radiograph of a Be sheet intended for use in manufacturing X-ray tube windows. The left half (light gray) of the image is the Be sheet. The right half (dark gray) of this figure lies off the Be sheet. This area was exposed directly to incident X-rays. The bright spot at lower left is an inclusion that is not acceptable in a tube window. The faint horizontal streaks are artifacts on the film caused by film processing. The black blob is an artifact caused by a light leak in the film package (fortunately this is not within the area of the Be sheet).

all three dimensions. For these objects, CT or a combination of radiography and CT is often used.

Reverse Engineering Using CT

It is becoming increasing common that the inspection or characterization of an object in all three spatial dimensions must be determined accurately. Examples include accurate external and internal dimensional measurements of an object by CT that can be directly input into computer aided design (CAD) or computer aided machining/manufacturing (CAM) programs (Figure 7.69) (78, 79). This can be useful when for example a nuclear reactor valve is no longer available from the manufacturer. CT can be used to scan the valve to obtain the internal and external dimensions and this data is input into a CAM code to easily manufacture

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

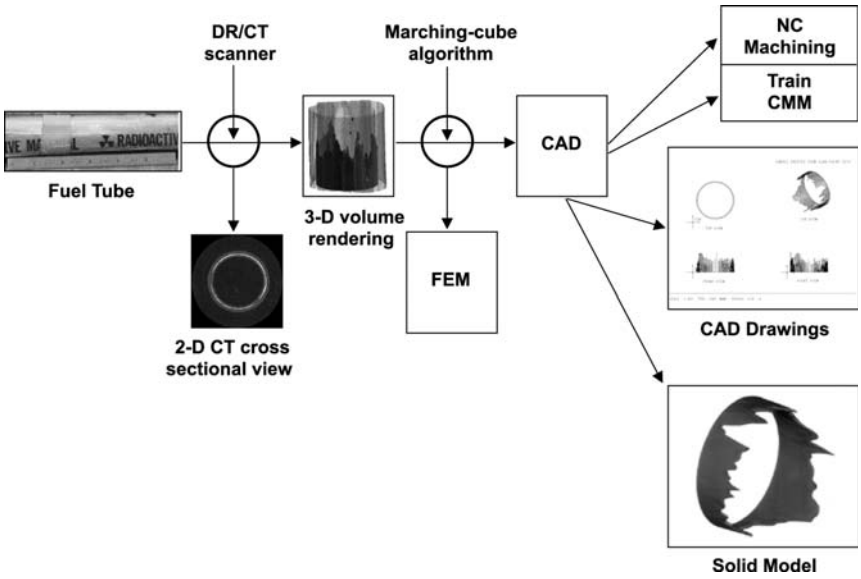


FIGURE 7.69 Flow diagram showing how CT data can be converted and used for numerical control (NC) machining, training a coordinate measurement machine (CMM), and for CAD drawings and models as well as finite element models (FEM). In this image is a reactor fuel tube. This tube contains uranium with an aluminum cladding.

the valve. Sometimes an object is generated ergonomically such as an airplane pilot's steering wheel that is molded by the seated polit's hands (80). The resultant wheel must then be dimensioned for production. CT provides a quick method to generate these measurements in all three dimensions. In most of these cases CT is the only method that can be used to obtain the *exterior* and *interior* dimensions of an object in all three dimensions nondestructively.

Here we describe an example where better implant designs are needed to mate the implants with the human joints. Human joints are commonly replaced in cases of damage from traumatic injury, rheumatoid diseases, or osteoarthritis. Frequently, prosthetic joint implants fail and must be surgically replaced by a procedure that is far more costly and carries a higher failure rate than the original surgery. Poor understanding of the loading applied to the implant leads to inadequate designs and ultimately to failure of the prosthetic (81). One approach to prosthetic joint design offers an opportunity to evaluate and improve joints before they are manufactured or surgically implanted. The modeling process begins with computed tomography data (Figure 7.70) of a female hand. This data

Copyright © Marcel Dekker, Inc. All rights reserved.

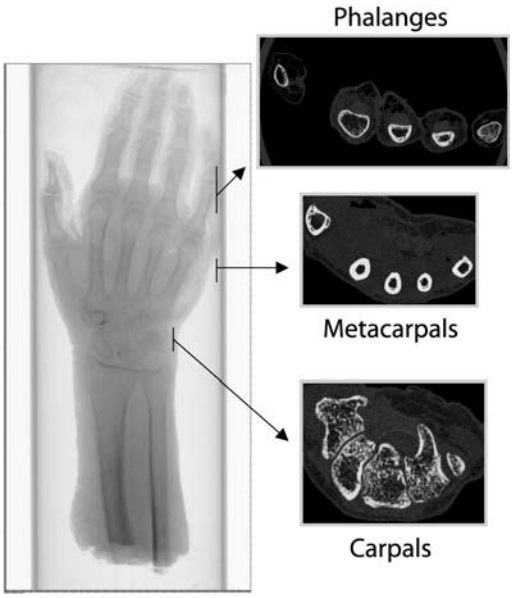


FIGURE 7.70 Digital radiograph of a human hand with some exemplar computed tomography cross-sectional images that reveal the true 3-D nature of computed tomography.

was acquired using a linear-array CT scanner called LCAT* (Figure 7.71) and a magnification of almost 3 × resulting in a voxel of ~160-μm × 160-μm × 160-μm. The CT data are used to segment the internal hard tissues (bone) as shown in Figure 7.72 (82). A three-dimensional surface of each joint structure of interest is created from the segmented data set (Figure 7.73).

An accurate surface description is critical to the validity of the model. The marching cubes algorithm is used to create polygonal surfaces that describe the 3-D geometry of the structures (most critical is bone) identified in the scans. Each surface is converted into a 3-D volumetric, hexahedral finite-element mesh that captures its geometry. The meshed bone structure for the women's hand is shown in Figure 7.73. Boundary conditions determine initial joint angles and ligament tensions as well as joint loads. Finite element meshes are combined with boundary conditions and material models. The analysis consists of a series of computer simulations of human joint and prosthetic joint behavior.

*LCAT or Linear Computed Axial Tomography.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

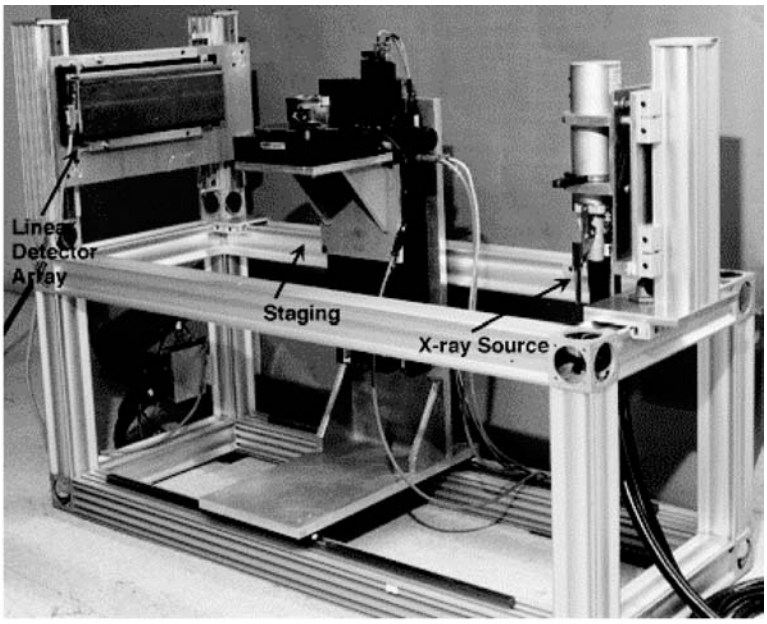


FIGURE 7.71 Photograph of a typical linear-array detector DR/CT system. From left to right are the detector, staging (elevation and rotation), and source. The fixturing structure is used to align the various components.

The simulations provide qualitative data in the form of scientific visualization and quantitative results such as kinematics and stress-level calculations (Figure 7.74). In human joints, these calculations help us to understand what types of stresses occur in daily use of hand joints. This provides a baseline for comparison of the stresses in the soft tissues after an implant has been inserted into the body. Similarly, the bone–implant interface stresses and stresses near the articular surfaces can be evaluated and used to predict possible failure modes.

Results from the finite element analysis are used to predict failure and to provide suggestions for improving the design. Multiple iterations of this process allow the implant designer to use analysis results to incrementally refine the model and improve the overall design. Once an implant design is agreed on, the device can be produced using computer-aided manufacturing techniques (83).

Bridge Pin Failure Mechanism Investigation

Here we use a bridge pin example where digital radiography and CT are both used in the inspection. Bridge pins were used in the hanger assemblies for some multi-span steel bridges built prior to the 1980s, and are sometimes considered fracture-

Copyright © Marcel Dekker, Inc. All rights reserved.

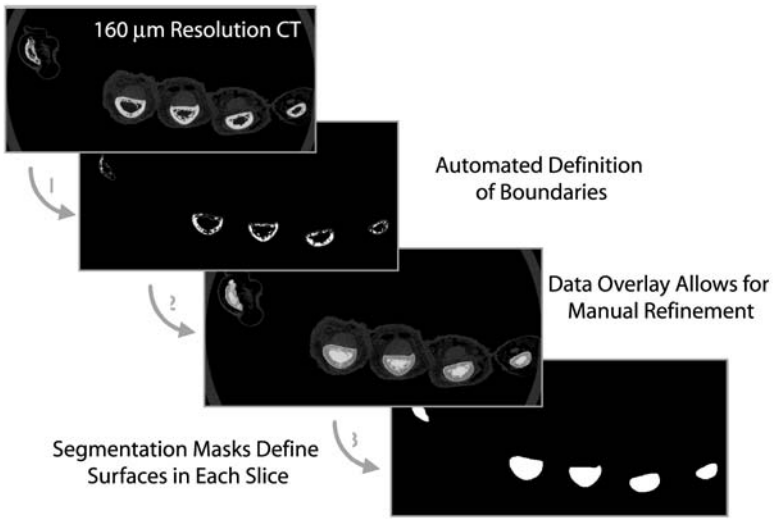


FIGURE 7.72 CT Segmentation: adult female hand. From left to right: One representative CT slice. Automatically defined boundaries for this CT slice. Semi-automated segmentation results overlaid onto its CT slice. Final surface segmentation results for this slice. After each of the 1084 slices is segmented they are investigated for interconnectivity and result in a 3-D segmentation of the bone surfaces within the women's hand.

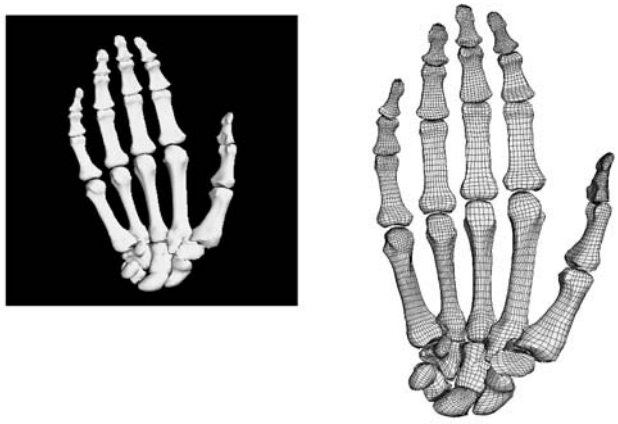


FIGURE 7.73 Left: A 3D surface-rendered image of the bone structure from the CT data of the women's hand. Right: 3-D hexahedral finite-element mesh of the segmented bone structure.

Copyright © Marcel Dekker, Inc. All rights reserved.

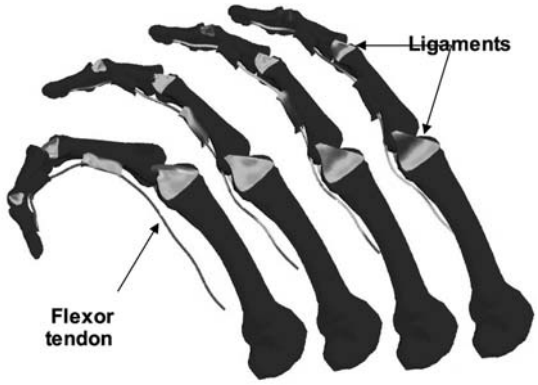


FIGURE 7.74 Three-dimensional model of the bone structure generated from CT data of the index finger. Different stages of finger flexion were studied by finite element analysis as shown. The different gray tones within the tendons and ligaments indicate stresses during flexion.

critical elements of a bridge. For example, a bridge pin failure was the cause of the 1983 collapse of a river bridge that resulted in the deaths of 3 people. Bridge pins (Figure 7.75) are typically made of steel. They are ~20-cm long with a 7.6-cm wide barrel and threaded ends of ~6-cm outer diameter. The pins typically

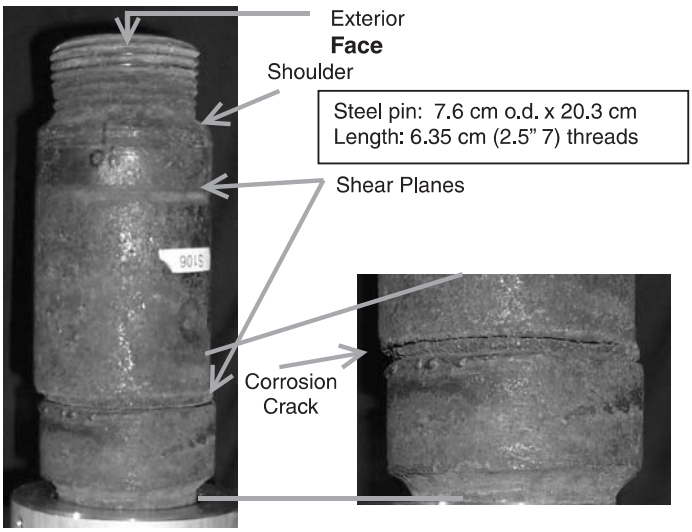


FIGURE 7.75 Photographs of one of the three bridge pins provided by the FHWA. This pin was the most damaged pin.

Copyright © Marcel Dekker, Inc. All rights reserved.

fail at the shear plane, the region where the shear forces from the hanger plate are transmitted to the web plate and where the majority of flaws so far are detected.

The only accessible part of the bridge pin available for inspection in the field is the exterior face of the threaded portion. This makes it difficult geometrically to inspect the pins. Ultrasonic methods are generally employed in the field to inspect the pin-hanger assemblies. These methods consist of manually scanning a hand-held probe over the exposed portion of the pin. Due to geometric constraints as well as limited access to the pins, the field inspection can be difficult to perform as well as interpret, due to the subjective nature of this method. For example, critical areas of the pin are often hidden by geometric constraints, corrosion groves at the surface of the shear planes typically occur and can be mistaken for cracks, and acoustic coupling between the pins and the hangers can occur creating erroneous results.

The Federal Highway Administration (FHWA) recently acquired several pins that were removed from a bridge in Indiana following an ultrasonic field inspection. Some of these pins were cracked and this provided the opportunity to better understand the failure mechanism(s) of bridge pins using laboratory NDE techniques. Visual inspection and laboratory inspection techniques such as ultrasonic C-scans and film radiography were performed by FHWA. Qualitatively these laboratory methods yielded similar results as found in the field inspections.

Digital radiography and computed tomography characterization methods were used to further study the bridge pins (84). A 9-MV LINAC X-ray source was combined with an amorphous-silicon, flat-panel detector (85) to acquire digital radiographs of three pins (Figure 7.76) and CT projections of one of the severely cracked pins. The CT data consisted of 180 projections over 180° and were reconstructed into several hundred tomograms (cross-sectional images) with a volume element (voxel) size of $\sim 127 \times 127 \times 127\,\mu m^3$. Two tomograms are shown in Figure 7.77 and three tomograms are shown in Figure 7.78.

It is useful to point out that far from the shear plane, Figure 7.78a, there are no apparent cracks or voids in the tomograms. Near the shear plane, Figure 7.78b, a crack appears within the pin at ~11 p.m. position within the pin. Within the shear plane, Figure 7.78c, we observe several cracks. The large surface-breaking cracks from 6–8 p.m. in Figure 7.78c are observed visually, radiographically and ultrasonically. Several smaller radial cracks are revealed in the tomograms that are not observed either visually or in the radiographic or ultrasonic data. These small radial cracks are only in the shear plane region and most likely lead to the larger cracks, which most likely cause pin failure.

Computed tomography reveals many more internal defects within the bridge pins than those observed in visual, radiographic, or ultrasonic inspection. For example, we observe large and small radial cracking, and are able to measure crack sizes in all three spatial dimensions. Initially, the field and laboratory ultrasonic inspection results were qualitatively corroborated by the large cracks observed in the X-ray tomograms. This provides some level of confidence in the

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

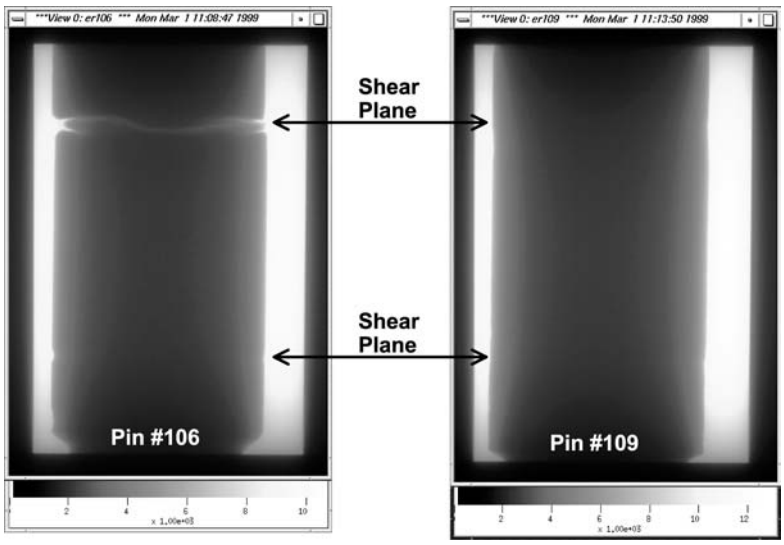


FIGURE 7.76 Digital radiographs of two of the three bridge pins provided by FHWA. Left: Severely cracked pin and Right: Uncracked pin.

ultrasonic field inspection method. Further work is required to quantitatively corroborate and correlate the X-ray tomographic data with the field and laboratory ultrasonic inspection results as well as to obtain a better understanding of bridge pin failure mechanisms and to improve field inspection reliability in predicting bridge pin failure.

Process Control of High Explosive Materials Using CT

Computed tomography techniques have been investigated to characterize high explosive materials and to improve current manufacturing processes for high explosives (HE). In order to revise the manufacturing process, it must be demonstrated that CT can provide information comparable to the techniques currently used: (a) shadow graphs, to determine gross outer dimensions; (b) radiography, to detect cracks, voids, and foreign inclusions; (c) dye penetrant, to identify surface cracks; (d) a wet/dry bulk-density measurement technique; and (e) destructive core sampling for exhaustive density measurements with crude spatial resolution. Application of X-ray CT could significantly improve safety and greatly reduce the process-flow time and costs, since CT could simultaneously determine rough outer dimensions, bulk density, density gradients, and the presence of cracks, voids, and foreign inclusions within HE (86, 87, 88).

Copyright © Marcel Dekker, Inc. All rights reserved.

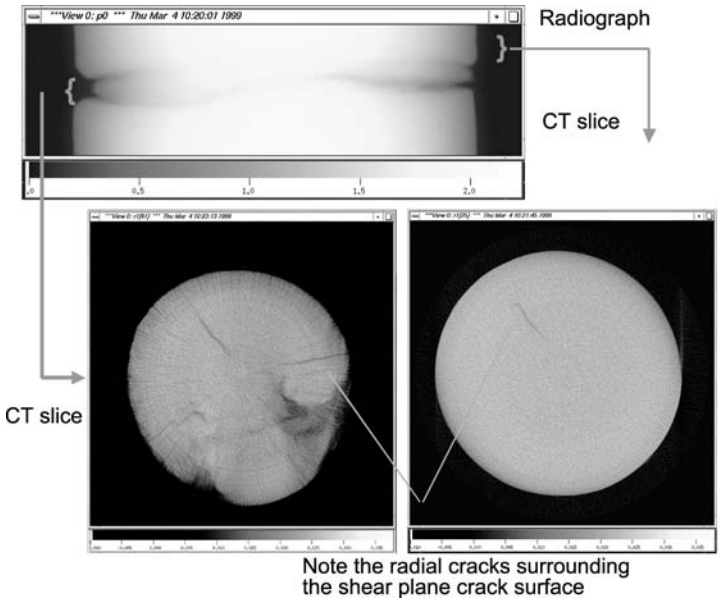


FIGURE 7.77 Digital radiograph of the severely cracked bridge pin and two CT slices showing the damage within the cracked portion of the pin.

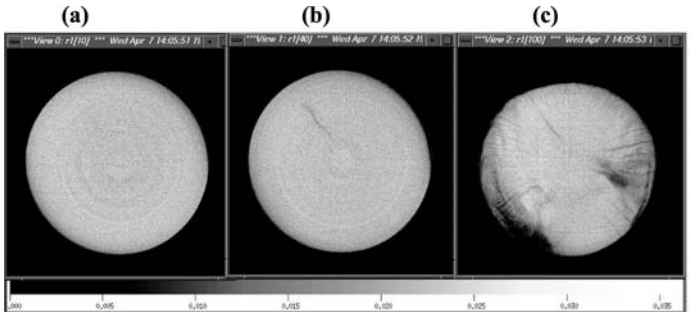


FIGURE 7.78 Tomograms along the longitudinal axis of a cracked bridge pin. The tomograms are (a) far from, (b) near, and (c) within the cracked shear-plane region. The gray scale relates color tones to relative attenuation in mm^{-1} units.

Copyright © Marcel Dekker, Inc. All rights reserved.

In the current HE manufacturing process, PBX9502* components are prepared by mixing a TATB† explosive powder with a Kel-F 800‡ plastic binder to ensure a solid pressing that has certain mechanical properties (tensile, compression, high strain rate tensile, etc.). PBX9502 material properties result from its two constituents and the pressing process. Using Kouris' et al. chemical formula mixture rule (89), the *effective-Z* (Z_{eff}) is calculated to be 7.1 for TATB and 12 for Kel-F. (The concept of effective atomic number (Z_{eff}) originated in medicine. Since the elements attenuate X-rays with distinctive energy dependence, a compound or a mixture can be assigned a Z_{eff} by comparison to the elements. But, since the energy-dependence of attenuation is complex, the value assigned depends on the details of the measurements and the standards. It is not an intrinsic property of the material.) The PBX9502 pellet volume proportions are 95% TATB and 5% Kel-F, resulting in an effective-*Z* of 7.6. Prior to the formulation, the density of TATB is $\sim 1.938\,g/cm^3$ and $\sim 2.02\,g/cm^3$ for Kel-F (90). The pressed PBX9502 density value is $\sim 1.9\,g/cm^3$. Different sets of explosive pellet pressings were fabricated. The explosives discussed here were cylindrical in shape $\sim$1.27-m diameter and $\sim$4-cm long, and are referred to as a pellet.

A single energy-discriminating detector, first-generation (Energy Discriminating Detection-Based Systems in Section 7.4.4) scan geometry system called PBCAT—Pencil-Beam Computed Axial Tomography—(91) was used to nondestructively evaluate the PBX9502 pellets. In this study, an $\sim$63 mCi ^{109}Cd (characteristic energy levels: 22, 25, and 88 keV) radioisotopic source, a high-purity intrinsic-germanium detector, γ-/X-ray spectroscopy electronics, and a multichannel analyzer were used to acquire the sinogram data. All PBCAT system reconstructed images are presented as values of absolute linear attenuation coefficients (cm^{-1}). Only the 22- and 25-keV characteristic lines were used in this analysis.

A PBX9502 pellet was scanned perpendicular to its longitudinal axis. This sinogram data set consisted of 90 projections at 2° intervals over 180°, and 40 ray sums with a 1-mm^2 source and detector aperture and 0.5-mm translational overlap, resulting in a spatial resolution of $\sim$1 mm. The total data acquisition time was 4 days. Photon counting statistics, on the average, were $\sim$1% and $\sim$3% for the 22 and 25 keV ray sum data respectively. The pellet sinogram data were reconstructed by a filtered backprojection algorithm into 40×40 pixel images (Figure 7.79).

*PBX is an acronym for plastic binder explosive; the number represents the explosive formulation.
†TATB represents 1,3,5-Triamino-2,4,6-Trinitro-Benzene ($C_6H_6N_6O_6$) (90).
‡Kel-F 800 consists of 75% chlorotrifluoroethylene polymer (C_2F_3Cl) and 25% vinylidenefluoride ($C_2H_2F_2$) with a calculated formula of $C_{2.1}H_{0.4}C_{0.6}F_{2.7}$ (90).

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

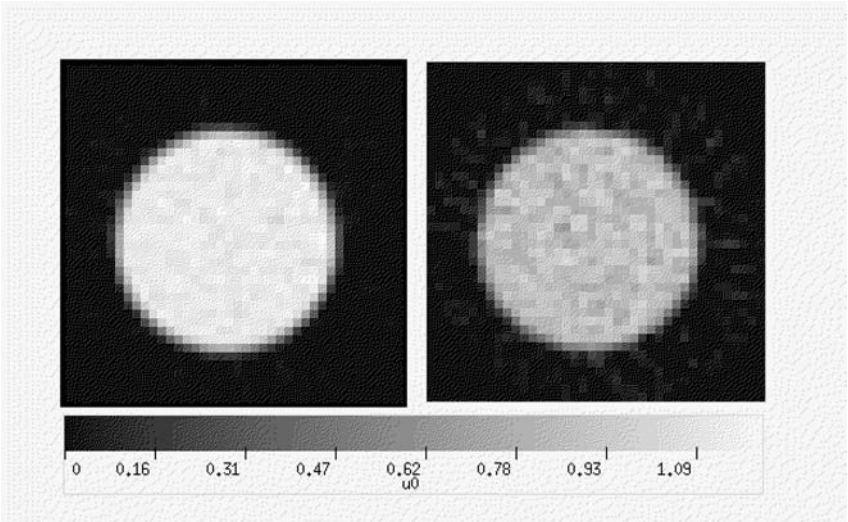


FIGURE 7.79 Reconstructed images from the PBCAT 22 (left) and 25 (right) keV sinogram data for the explosive-pellet. The grayscale is in cm^{-1}.

The pellet's PBCAT data resulted in mean and standard deviation, σ, values of 1.12 ± 0.03 and $0.87 \pm 0.7\,cm^{-1}$ at 22 and 25 keV, respectively. These data agree with the predicted values, 1.11 ± 0.03 and 0.86 ± 0.03 (92). Overall, the linear attenuation coefficient images are uniform (Figure 7.79), however this may not be true for the distribution of the elemental composition or Z and density, ρ. In PBX9502, changes in Z are a result of a variation in the distribution of TATB ($Z = 7.1$) and Kel-F ($Z = 12$) within the pressed pellet. Density variations can be a result of nonuniformities in the pressing process, or from higher or lower amounts of Kel-F, which has a slightly higher density than TATB. Quantitative multiple-energy data generated by PBCAT for the pellets were used to calculate *effective-Z*, *weight fraction*, and *density* images (Figure 7.80) (93, 94).

At the slice plane chosen, the experimentally determined mean effective-Z is 7.8 ± 0.5, which is within experimental error of the predicted value, 7.6. The variations in the effective-Z image are due solely to changes in the PBX9502 constituents (TATB and Kel-F) not upon density. This is just one component of the linear attenuation coefficient pixel-to-pixel variations (Figure 7.79). The weight fraction coefficient calculations involve both tabulated attenuation coefficients and nominal density values for TATB and Kel-F. The experimentally determined TATB weight fraction mean is 0.94 ± 0.03 (Figure 7.80). The expected TATB weight fraction is 0.95. The experimental and expected values are well within experimental error. The pixel-to-pixel variation in the weight

Copyright © Marcel Dekker, Inc. All rights reserved.

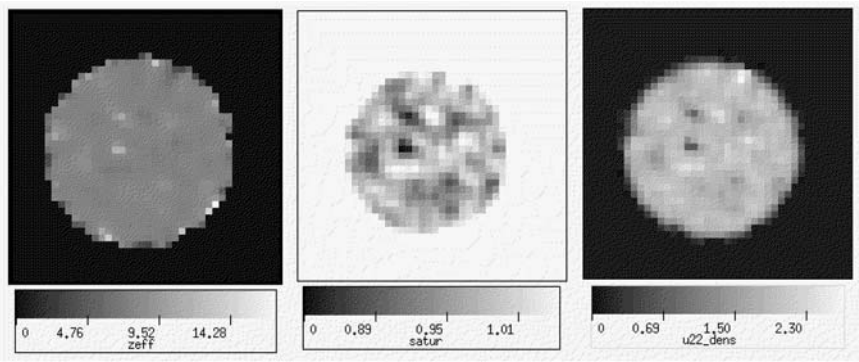


FIGURE 7.80 Resultant effective-Z (left), TATB weight fraction (middle), and density (right) images for the explosive-pellet. The gray scale is given in effective-*Z*, TATB weight fraction, and density (g/cm^3) units, respectively.

fraction image tracks with the variation in the effective-*Z* image; i.e., higher weight fraction values for Kel-F correspond with higher effective-*Z* pixels.

The density image for the medium pellet also is presented in Figure 7.80. The experimentally determined slice-plane mean density value is $1.9 \pm 0.2\,g/cm^3$ and is within the wet/dry bulk density measured value of $1.904\,g/cm^3$. This agreement provides some credibility to the density values on a pixel-to-pixel basis within the density image. There is considerably more variation in the pixel-to-pixel density values than in the linear attenuation coefficient images (Figure 7.79).

Comparison of all three images in Figure 7.80 reveals two small regions in the top left quarter (near the center) of each image that stand out. For both regions, the effective-*Z* values are high and the TATB weight fraction values are low implying a large concentration of Kel-F, but the density values are unexpectedly low. These low-density values can be explained by either TATB and/or Kel-F with a large void fraction. The effective-*Z* and weight fraction results suggest that the small regions are Kel-F with a larger than normal void fraction.

In summary, a preliminary comparison between CT and the current manufacturing inspection techniques reveals that CT may be able to replace most if not all of the current techniques used. If these pellet results are scalable to full size explosives and linear- or area-array non–energy discriminating DR/CT systems can provide similar results, then CT is a likely candidate for an automated high explosives inspection system. It should be pointed out that CT can be used to determine gradients in not only density but chemical constituents

Copyright © Marcel Dekker, Inc. All rights reserved.

as well. Therefore, it may be necessary to reinvestigate the current bulk density specification with respect to wet/dry measurement techniques.

7.6 RADIATION SAFETY

7.6.1 Why Is Radiation Safety Required?

Accidents Happen

Radioactive materials and radiations (alpha, beta, X-ray, and gamma) exist, and have existed, in our world since it was created. In the past century, we have learned to use these natural, as well as some manmade, radioactive materials in beneficial ways including medical diagnoses and therapy, industrial gauges, biological and agricultural research, analytical chemistry and applications such as nondestructive testing, smoke detectors, and luminescent signs and watches.

Along with the applications of radioisotopes and radiation have come accidents. Most are of little consequence but some overexposures result in injury and even death. Radiography involves the use of large radioactive sources ($100\,\text{Curies}/3.6 \times 10^{12}$ Becquerels) emitting penetrating radiation (usually gamma rays) or machines producing very large amounts of radiation (usually X-rays). Consequently, accidents with these sources can often be very serious.

Just as we are exposed to tiny amounts of energy such as heat, light, or natural radiation that appear to have no detrimental effect upon us, exposure to larger amounts of any kind of energy can cause negative health effects and even death. Radiation overexposure can produce sickness, loss of appendages, and death. Radiographers seem to have had an excessive number of accidents with overexposures and potential overexposures that has caused governmental agencies to regulate the use of radioisotopes and radiation. Refs. 95–97 have listings of radiographic accidents to illustrate how they happen and what the consequences may be. Generally, the accidents happen because a radiation survey meter is not used to determine radiation levels or an individual purposely defeats safety devices such as interlocks. In the later instance, an individual crawled under a door to prevent shutting down a radiation machine and attempted to change samples while it was running. In failing to beat the speed of light, the individual lost an arm and leg and required bone marrow transplants. Radiographers have frequently failed to determine with a survey meter that their radiographic source is properly shielded and have lost fingers when they were overexposed in taking their equipment apart or have overexposed other people when the source was lost from the equipment and left exposed for several days.

Governmental regulations now require specific procedures, training of the radiographers, and use of radiation detection equipment in an effort to reduce accidents and radiation overexposures in the performance of radiography. An

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

overview of these requirements and background for their formulation is presented in the following sections on radiation safety and responsibilities for safety.

Radiation Effects

As the consequence of our living in a naturally radioactive world along with the additional manmade sources of radiation exposure, all humans receive between 2.5 to 5.5 millisieverts [mSv- the SI unit of exposure] or 250 to 550 milliroentgens [mR] per year. Some areas of the world can produce exposures as high as 0.2 mSv/hr (20mr/hr) from the high uranium and/or thorium contents of the soil or the very high altitude of the area (thinner layers of air permit higher exposures to cosmic radiation).

The radiation exposure units have been developed over the years as the significance of radiation effects on materials and living things have become better understood. First, the Röntgen (R) was defined as the amount of X- or gamma radiation producing one electrostatic unit of electrical charge (2.09×10^9 ion pairs) in one cubic centimeter of dry air at standard temperature and pressure. The Röntgen was followed by the Röntgen equivalent in mammals (Rem) which was the radiation dose equivalent to the amount of damage done in biological material by one Röntgen of X- or gamma radiation. In the existing International System of Units (SI), the absorbed dose unit is the Gray (Gy) which is equivalent to absorption of one joule per kilogram of material and the dose equivalent unit is the Sievert (Sv) which is equal to 100 Rem.

Table 7.8 provides some insight into results of acute exposure (over short periods of time such as a few days) to the entire body of the individual. If the same exposure is spread out over large amounts of time, no effects are seen and deaths do not occur until much higher doses are received. This is indicative of the repair mechanisms existing in the body for repair of radiation damage if damage is small compared to the ability or speed of repair.

Localized doses require large doses such as 5 Sv (500 Rem) for permanent damage although reddening and discomfort in the exposed area will occur at

TABLE 7.8 Radiation Dose Versus Health Effect

Radiation Dose (Sv/Rem)[a]	Health Effect
0.25 Sv/25 Rem (or less)	None detectable
0.50 Sv/50 Rem	Small changes in the blood (temporary)
1.00 Sv/100 Rem	Nausea and fatigue
2.00–2.50 Sv/200–250 Rem	Single death occurs
5.00 Sv/500 Rem	Half of recipients die

[a]Dose received in a short time (few days at most).

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

lower doses. At and above 10 Sv (1000 Rem) permanent damage occurs and amputation may be required.

When radiation exposure is large but spread over long periods of time, delayed effects such as cancer may be observed. Radiographers may receive an occupational exposure of 0.02 Sv (20 Rem) over their working lives and show no detectable effects. However, the risk of cancer will increase to 20 cancer deaths from the radiation exposure in 10,000 radiographers receiving that amount. This does not count the 2,000 cancer deaths that will occur in that same population from natural causes. The cancer deaths from radiation exposure will not be identifiably different from the other cancer deaths.

Concern over genetic effects from exposure also exists. So far, no genetic effects have been observed in large exposures of large populations (Japanese exposed to atom bombs). Rapidly growing systems are more susceptible to genetic damage so pregnant women are limited to much lower levels of exposure than adults would be just to protect the fetus. Occupational exposures are also limited to persons 18 years or older with maximum annual exposures of 0.05 Sv (5 Rem).

Summary of General Regulations for Radiography

As the benefits of using radioisotopes and radiation began to grow, so did the number of accidents and the awareness of the consequences of overexposure to the radiation. At the conclusion of WW II and the use of the atom bomb, the U.S. federal government established the U.S. Atomic Energy Commission (USAEC) to handle regulations for the beneficial uses of radioisotopes and radiation. The Department of Health already had begun promulgating safety regulations for X-ray and other machines used in medical diagnoses and therapy. Some states, known as "Agreement States," began to take over the USAEC regulation of all radioisotope and radiation uses except for nuclear power plants. Agreement states at this time are: Alabama, Arizona, California, Colorado, Florida, Georgia, Illinois, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Mississippi, Nebraska, Nevada, New Hampshire, New Mexico, New York, North Carolina, North Dakota, Oregon, Rhode Island, South Carolina, Tennessee, Texas, Utah, and Washington. The states of Minnesota, Ohio, Oklahoma, Pennsylvania, and Wisconsin have sent letters of intent to the USNRC to become Agreement States. And finally, the USAEC was converted into the existing U.S. Nuclear Regulatory Commission (USNRC or NRC) with Agreement States programs compliant with that agency's programs and regulations. An Agreement State's regulations must comply with all of the USNRC's regulations but may be more stringent if that state wishes.

In some cases, the Agreement States have led in regulation reform. Texas began a formal certification program for industrial radiographers. The program required that radiographers receive specific training in radiography and safety

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

procedures followed by a state administered examination. Certification cards were given to those who passed the exam after proving they had the required training and experience. No radiographers could work in Texas on a regular basis without this certification. The American Society for Nondestructive Testing (ASNT) picked up the certification program and made it available in all states under the funding from the USNRC. The certification from ASNT is now available in all agreement states and USNRC regulated states as well as worldwide.

The USNRC required the certification of radiographers on June 27, 1999, that then under reciprocity became mandatory in all 50 of the United States. Other requirements include training (at least 40 hours are required by all certifying agencies and agreement states) on safety, use of radiation detection equipment for personnel dosimetry and surveys, proper radiographic procedures, control of radiographic devices (including storage, use, and disposal), and knowledge of the federal (or appropriate state) regulations concerning radiography and radioisotope use. Other general requirements include a minimum of two months of on-the-job training and a study of radiographic accidents and their results. The certification must be renewed by taking a new examination every five years. This is necessary because the regulations and what they require change frequently. Anyone getting into the field or working in the field of radiography needs to check the regulations often (many details and specific regulations on radioisotope handling and control of radiation exposures are not listed here) to stay abreast and in compliance with the complete and new requirements. Some, but not all and certainly none of the regulations developed since writing of this section, are presented in the following material. For complete information on current regulations, contact the USNRC or appropriate Agreement State. (See Ref. 98 for listings and addresses).

7.6.2 Radiation Safety Procedures

Reducing Radiation Exposure

One of the general regulations for conducting radiography, or using any radioisotope or radiation sources, is to keep exposures "*As Low As Reasonably Achievable*" (known as the ALARA concept). There are three procedures, *Time*, *Distance*, and *Shielding* that must be followed to do this. These three procedures require that the radiographer keep *time* during exposure as small as possible, put as much *distance* between the source and the individual as is possible, and use as much *shielding* as is physically possible. By doing these three things, radiation exposure can be minimized to reach ALARA radiation exposure levels. They are simple concepts but are often ignored in the swamp and mosquitoes of every day work schedules and deadlines.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Time is the simplest concept for reducing radiation exposure. The less time spent in a radiation field, the less exposure. A direct relationship:

$$Exposure = (time)(dose\ rate)$$

can be used to calculate exposure when the time of occupancy and dose rate mSv/hr (mR/hr or mrem/hr) are known. Radiographers frequently forget that very low dose rates become significant doses when times are large. An example is that the shield source may be stored close by for long periods of time such as when traveling to and from the job site. The dose rate may only be 0.010 mSv/hr (10mRem/hour) from the source in its shield at some location in the vehicle used to travel to and from the job site, but if the trip takes 4 to 5 hours each way, that can result in an exposure of as much as 1.0 mSv (100 mRem) and no work has been done! If the radiographer is limited to 0.05 Sv (5 Rem) per year, the 1.0 mSv (100 mRem) has become his week's dose. Likewise, the radiographer often exposes the source and stays at the crank position where the dose rate might be 0.1 mSv (100 mRem) per hour instead of staying for the same length of time where the dose rate might be a small fraction of that number. Remember that this is a direct relationship, twice as long gives twice the dose or half as long gives half the dose.

Distance is also a simple concept, further away from the source gives lower dose rates and closer gives higher dose rates. The relationship is known as the Inverse Square Law because the dose rate changes inversely with the square of the distance:

$$(\text{Dose rate}_1)(\text{distance}_1)^2 = (\text{dose rate}_2)(\text{distance}_2)^2$$

Obviously, increasing the distance reduces the dose rate very rapidly as compared to just reducing the time. For example, doubling the distance decreases the dose rate by four; tripling the distance reduces the dose rate by nine. Decreasing the distance causes rapid increases in the dose rate, especially at short distances such as less than a foot. The inverse square relationship results from the radiation emitted by the source traveling in all directions so as to cover the surface of a sphere around the source. As the distance from the source increases, the area of the sphere increases as the square of that distance. Since the same amount of radiation exists and it is covering an area increasing as the square of the distance, the decrease in radiation per unit area decreases as the square of the distance from the source.

Shielding involves a more complicated relationship but provides significant decreases in dose rate. The relationship of dose rate to the amount of shielding is exponential:

$$Dose\ rate_{\text{shielded}} = (dose\ rate_{\text{not shielded}} e^{-\mu l})$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

in which the exponents of the base e are μ, the linear absorption coefficient for the source and absorber materials, and l, the thickness of the absorber. As an example of the effectiveness of an absorber, change the shielding around a Co-60 source from nothing to two inches of lead (the thickness of lead that reduces the gamma radiation intensity by half is ~10 mm.). Two inches of lead would reduce the intensity of radiation from the source by 1/28 or to 0.035 of the nonshielded intensity. This reduction is true for any given distance or time. The 10 mm of lead to reduce radiation intensity of Co-60 by half is known as the half-value layer (HVL). Tables of half-value layers are given in radiographic texts and manuals for the more common sources and shielding materials used by radiographers (e.g. see Table 7.3). Physical constant handbooks for Chemistry and Physics often give the linear absorption coefficients that can easily be converted to half-value layers by:

$$\mathrm{HVL} = \frac{0.693}{\mu}$$

Linear absorption coefficients have units of reciprocal length. Mass absorption coefficients often found in some handbooks have units of mass per unit area and must be multiplied by the absorber density in appropriate units to obtain the linear absorption coefficient.

Radiographers use combinations of time, distance, and shielding to obtain dose rates that meet ALARA requirements as well as staying within their own exposure limits. The three concepts are also used to meet limits for exposure of the public. An example of a limit for public exposure is at the barricade for making a radiographic exposure; the radiation level must be such that the exposure is less than 2 mRem in any one hour to anyone standing at the barricade. Time, distance and shielding are all used to reach these kinds of values. Shielding of the source at an exposure position often requires a collimator that reduces the radiation in all directions except toward the specimen being radiographed. Collimators can greatly reduce the time and distance requirements for both the radiographer and the general public.

Reducing Accidents

Procedures required to reduce the number of radiographer accidents start with the training described above in certification. This training includes both the formal classroom training and the on-the-job training required before certification. It also includes continued training by Level III radiographers in the companies and the Radiation Safety Officer for the company. The after certification training is done to remind the radiographer of proper procedures, reduce the chances of cutting corners on safety procedures, and bring new requirements into the work procedures.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Recent requirements brought two person teams of radiographers into being where single individuals were the norm. Other new requirements brought on the use of alarming ratemeters/survey meters to reduce the chances of a radiographer working in a high radiation area without being warned. Previously, the equipment was visual only and a radiographer could easily walk into a radiation area without a warning. Radiographic exposure devices are designed to be more rugged and have an interlock system that will not permit unlocking the source for transfer to an exposure area unless the source is properly connected to the drive cable. Disconnected sources combined with failure to use survey meters produced most overexposures to radiographers in the recent past. Also required now is a mechanism in the exposure device that automatically locks the source into the shielded position when it is retracted from the exposed position.

The hardest thing to do has been to produce the good habit of always surveying so that the radiographer knows where the source is and what the radiation exposure level is. Training, company procedures, and safety training all require that the source be surveyed before it is transported to the work area. A survey is made as the source is exposed to assure that the source did actually move to the exposure area and that radiation levels are as expected from calculations and experience. On long exposures, the barricaded area is surveyed, as is the area from which the radiographer watches the area during exposure of the source. The retrieval of the source is surveyed, as is the shielded source immediately after retrieval (sometime the source has failed to be completely shielded by stopping too soon when it reaches the shield). The source is surveyed again when placed in the transport to storage and at storage. After long periods of time without incident, radiographers may cut corners on some of these surveys. When such behavior becomes habit, accidents happen.

7.6.3 Responsibilities for Safety

Radiographer

The radiographer maintains the first line of defense against accidents. That individual must follow all procedures and checks for safety that were in the training and regulatory requirements set up for making radiographs. The radiographer must see that no short cuts or bad decisions related to the proper exercise of safety occur even when pressures from clients and company would seem to move away from safety. The radiographer is responsible for his/her own safety as well as the public that could be affected by this work

Employer

The employer is responsible for good working conditions, equipment, and attitude of the employees. Any reduction in the above brings on circumstances in which proper procedures are not followed, equipment is not used or worse—

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

defective equipment is depended upon for safety, and bad attitudes produce bad and unsafe habits that lead to accidents. In explaining why they did not use survey meters or proper procedures, many radiographers have blamed the company for pressures to complete more work in less time or lose their jobs.

Clients

Clients frequently fail to require that work be done to some code or specification and that radiographers be certified to do the work. By selecting nonqualified companies and employees to do work or do the work at lower than competitive prices, unsafe procedures become the normal procedures and accidents happen. Then it is too late to add the proper qualifications for the supplier and its employees. A check should be made by the client of the company, its safety reputation, and credential of the employees. Observation of work procedures should be made at the beginning and continue at appropriate intervals during the job to assure that quality, safe work is being done.

Agencies

Agencies such as the NRC, U.S. Department of Health, and Radiation Control Departments of Agreement States must continue to evaluate regulations, add those that would contribute to increased safety and remove those that do not, and inspect for compliance with the regulations. At this period of time, the personnel in the agencies are better qualified and do a better job of fulfilling the above requirements than ever before. The agencies are better staffed and equipped to be of help in emergencies and have had specific training in doing just that. They are our last line of defense in preventing serious accidents.

PROBLEMS

1. You have available a 100 W X-ray source with a tungsten anode and a 2-mm spot size. You need to operate at 75 kV to get about 15% transmission through an aluminum plate. What is the maximum current you can operate at?
2. (a) If you operate a 100 W X-ray source with a tungsten anode and a 2-mm spot size at 45 kV how many characteristic lines will be produced? (b) How many characteristic lines are generated when operating this X-ray machine at 125 kV?
3. For a 100 W X-ray source with a tungsten anode and a 2-mm effective spot size what is the expected penumbra if you operate at a 10-mm object-to-detector distance and a 100-mm source to detector distance?
4. A radiographer takes a film radiograph of a steel gear. The customer requested a maximum film density of 1.0 for convenience in viewing on overhead cubicle lights. Does this compromise contrast? defect sensitivity?

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

5. A trial film radiograph of a cast Mg cylinder turns out to have film density ranging from 3.0 in the background to 2.5 through the diameter of the cylinder. Is this a quality job? Why? What should be corrected? Could film processing be poor?
6. A critical weld in a weldment is arranged so that it is 440 mm from the detector plane. The source available for the work has a spot size of 1.5 mm. The maximum distance between source and detector is 4.0 m because of the size of shielded room available. What is the source unsharpness for the image of the critical weld? Would a larger room improve unsharpness? Would this improve image quality if the pixel size on the detector is 127 μm?
7. The manufacturer of a Be-Cu spring for computer hard drives is considering radiography for inspecting the 0.1-mm thick sheet material prior to spring winding. The material density is 9.0 g/cm^3. The material is 2.0 wt% Be and the remainder Cu. What energy X-rays should be used? Estimate the relative attenuation of X-rays caused by Be and Cu. What applied voltage and filtration should be used with a W-anode tube to get X-rays of this approximate energy range?
8. A film radiograph has faint streaks that extend from edge to edge. What should you suspect causes this?
9. An ^{192}Ir source produced a satisfactory radiograph on May 11, 2001 of a component using an exposure time of 5 minutes in a specific configuration. What exposure time is required using this source in the same set up on July 4, 2001?
10. You are performing a CT scan of a deep-drawn stainless steel filter housing when you notice the object has tipped over. You should:
    (a) Start over.
    (b) Replace the object and continue from where the accident occurred.
11. You are looking at a data set for a single CT slice that was taken with a linear array detector. When displayed as a sinogram (I/I_0 plotted in angle-dimension space) you observe two vertical streaks (variable angle). What could cause these artifacts? What will they look like in the reconstructed data?
12. In performing a field radiography job, a technician receives 500 rem of dose equivalent over two working shifts. Is this acceptable?
13. What radiation method might be used to inspect the uniformity of gunpowder loading in a large military munition with a heavy brass cylindrical case?
14. Can you think of a reason to place a cylindrical part somewhat off the center of rotation when performing a CT scan?

GLOSSARY (99–101)

Absorbed dose: The amount of energy imparted by ionizing radiation per unit mass of irradiated matter. Denoted by "rad;" 1 rad = 0.01 J/kg. SI unit is "gray;" 1 gray = 1 J/kg.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Absorbed dose rate: The absorbed dose per unit of time; rads/s. SI unit, grays/s.

Absorption: The process whereby the incident particles of photons of radiation are reduced in number or energy as they pass through matter.

Absorption edge: A discontinuity in the absorption curve where, as the energy of an X-ray is increased, the absorption in the solid increases abruptly at levels where an electron is ejected from the atom.

Accelerating potential: The difference in electric potential between the cathode and anode in an X-ray tube through which a charged particle is accelerated; usually expressed in units of kV or MV.

Alpha particle: A positively charged particle emitted by certain radionuclides. It consists of two protons and two neutrons, and is identical to the nucleus of a helium atom.

Analog image: An image produced by a continuously variable physical process (for example, exposure of film).

Analog to digital converter (a/d *or* A/D): A device that changes an analog signal to a digital representation of the signal.

Angstrom (Å): Measurement of length, $1\,\text{Å} = 10^{-10}\,\text{m}$.

Anode: The positive electrode of a discharge tube. In an X-ray tube, the anode carries the target.

Anode current: The electrons passing from the cathode to the anode in an X-ray tube, minus the small loss incurred by the back scattered fraction.

Artifact: Spurious indication on a radiograph arising from, but not limited to, faulty manufacture, storage, handling, exposure, or processing.

Attenuation coefficient: Related to the rate of change in the intensity of a beam of radiation as it passes through matter.

Attenuation cross section: The probability, expressed in barns, that radiation (e.g. X-ray or neutrons) will be totally absorbed by the atomic nucleus.

Back scattered radiation: Radiation which is scattered more than 90° with respect to the incident beam, that is, backward in the general direction of the radiation source.

Barn: A unit of area used for expressing the area of nuclear cross sections.

$$1\text{barn} = 10^{-28}\,\text{m}^2$$

Becquerel (Bq): The SI unit of activity of a radioactive isotope, one nuclear transformation per second. $3.7 \times 10^{10}\,\text{Bq} = 1\text{Ci}$.

Betatron: An electron accelerator in which acceleration is provided by a special magnetic field constraining the electrons to a circular orbit. This type of equipment usually operates at energies between 10 and 31 MeV.

Blooming: In radiologic real-time imaging, an undesirable condition exhibited by some image conversion devices and television pickup tubes brought about by exceeding the allowable input brightness for the device, causing

Copyright © Marcel Dekker, Inc. All rights reserved.

the image to go into saturation, producing a fuzzy image of degraded spatial resolution and grey scale rendition.

Bremsstrahlung (white radiation): The electromagnetic radiation resulting from the retardation of charged particles, usually electrons.

Broad beam attenuation (absorption): The X-ray attenuation when contributions from all sources, including secondary radiation, are included.

Cassette: A light-tight container for holding radiographic recording media during exposure, for example, film, with or without intensifying or conversion screens.

Characteristic curve: The plot of density versus log of exposure or of relative exposure. (Also called the *D*-log *E* curve or the *H-D* curve.)

Cine-radiography: The production of a series of radiographs that can be viewed rapidly in sequence, thus creating an illusion of continuity.

Collimator: A device of radiation absorbent material intended for defining the direction and angular divergence of the radiation beam.

Compton scattering: When a photon collides with an electron it may not lose all its energy, and a lower energy photon will then be emitted from the atom at an angle to the incident photon path.

Compton scatter radiation: The scattered X-ray or gamma ray that results from the inelastic scattering of an incident X-ray or gamma ray on an electron. Since the ejected electron has short range in most materials, it is not considered part of the scattered radiation.

Computed radiology (photo stimulated luminescence method): A two-step radiological imaging process; first, a storage phosphor imaging plate is exposed to penetrating radiation; second, the luminescence from the plate's photostimulable luminescent phosphor is detected, digitized, and presented via hard copy or a CRT.

Contrast sensitivity: A measure of the minimum percentage change in an object which produces a perceptible density/brightness change in the radiological image.

Contrast stretch: A function that operates on the greyscale values in an image to increase or decrease image contrast.

Cumulative dose: The total amount of radiation received in a specified time.

Curie (Ci): That quantity of a radioactive isotope which decays at the rate of 3.7×10^{10} disintegration per second.

Densitometer: A device for measuring the optical density of radiograph film.

Density (film): A quantitative measure of film blackening when light is transmitted or reflected.

Density gradient (G): The slope of the curve of density against log exposure for a film.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Diffraction: Interference effects giving rise to illumination beyond the geometrical shadow. This effect becomes important when the dimensions of the apertures or obstructions are comparable to the wavelength of the radiation.

Digital image: An image composed of discrete pixels each of which is characterized by a digitally represented luminance level.

Digital image enhancement: Any operation used for the purpose of enhancing some aspect of the original image.

Dynamic range: Ratio of the largest to the smallest dynamic inputs along the linear (usable) portion of the film response curve.

Electron volt: The kinetic energy gained by an electron after passing through a potential difference of 1 V.

Equivalent I.Q.I. sensitivity: That thickness of I.Q.I. expressed as a percentage of the section thickness radiologically examined in which a 2T hole or 2% wire size equivalent would be visible under the same radiological conditions.

Equivalent Penetrameter Sensitivity: That thickness of penetrameter, expressed as a percentage of the section thickness radiographed, in which a 2T hole would be visible under the same radiographic conditions.

Exposure, radiographic exposure: The subjection of a recording medium to radiation for the purpose of producing a latent image. Radiographic exposure is commonly expressed in terms of milliampere-seconds or millicurie-hours for a known source-to-film distance.

Exposure range (latitude): The range of exposures over which a film can be employed usefully.

Film contrast: A qualitative expression of the slope or steepness of the characteristic curve of a film; that property of a photographic material which is related to the magnitude of the density difference resulting from a given exposure difference.

Film speed: A numerical value expressing the response of an image receptor to the energy of penetrating radiation under specified conditions.

Filter: Uniform layer of material, usually of higher atomic number than the specimen, placed between the radiation source and the film for the purpose of preferentially absorbing the softer radiations.

Fluorescence: The emission of light by a substance irradiated with photos of sufficient energy to excite the substance. The fluorescent radiation is monochromatic, having a wavelength that is characteristic of the fluorescing material.

Focal spot: For X-ray generators, that area of the anode (target) of an X-ray tube that emits X-ray when bombarded with electrons.

Fog: Generic term denoting an increase in optical density of a processed photographic emulsion caused by anything other than direct action of the image forming radiation and caused by one or more of the following:

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

(a) Aging—deterioration, before or after exposure, or both, resulting from a recording medium that has been stored for too long a period of time, or other improper conditions
(b) Base—the minimum uniform density inherent in a processed emulsion without prior exposure
(c) Chemical—resulting from unwanted reactions during chemical processing
(d) Dichroic—characterized by the production of colloidal silver within the developed sensitive layer
(e) Oxidation—caused by exposure to air during developing
(f) Exposure—arising from any unwanted exposure of an emulsion to ionizing radiation or light at any time between manufacture and final fixing
(g) Photographic—arising solely from the properties of an emulsion and the processing conditions, for example, the total effect of inherent fog and chemical fog
(h) Threshold—the minimum uniform density inherent in a processed emulsion without prior exposure

Fog density: A general term used to denote any increase in the optical density of a processed film caused by anything other than the direct action of the image-forming radiation.

Gamma-radiography: A technique of producing radiographs using gamma rays.

Gamma ray: Electromagnetic penetrating radiation having its origin in the decay of a radioactive nucleus.

Geometric unsharpness: The penumbral shadow in a radiological image which is dependent upon (a) the radiation source dimensions, (b) the source to object distance, and (c) object to detector distance.

Graininess: The visual impression of irregularity of silver deposit in a processed film.

Gray (Gy): The SI unit of absorbed dose. $1\,\mathrm{Gy} = 1\,\mathrm{Joule/kg} = 100\,\mathrm{rad}$.

Half-life: The time required for one half of a given number of radioactive atoms to undergo decay.

Half-value layer (HVL): The thickness of an absorbing material required to reduce the intensity of a beam of incident radiation to one half of its original intensity.

Image Quality Indicator (IQI): In industrial radiology, a device or combination of devices whose demonstrated image or images provide visual or quantitative data, or both, to determine radiologic quality and sensitivity. Also known as a penetrameter (disparaged). NOTE: It is not intended for use in judging size nor establishing acceptance limits of discontinuities.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Intensifying screen: A material that converts a part of the radiographic energy into light or electrons and that, when in contact with a recording medium during exposure, improves the quality of the radiograph, or reduces the exposure time required to produce a radiograph, or both. Three kinds of screens in common use are:

(a) Metal screen—a screen consisting of dense metal (usually lead) or of a dense metal compound (for example, lead oxide) that emits primary electrons when exposed to X- or gamma rays.
(b) Fluorescent screen—a screen consisting of a coating of phosphors which fluoresces when exposed to X- or gamma radiation.
(c) Fluorescent-metallic screen—a screen consisting of a coating of a metallic foil (usually lead) coated with a material that fluoresces when exposed to X- or gamma radiation. The coated surface is placed next to the film to provide fluorescence; the metal functions as a normal metal screen.

IQI sensitivity: In radiography, the minimum discernible image and the designated hole in the plaque type, or the designated wire image in the wire type image quality indicator.

keV (kilo electron volt): A unit of energy equal to one thousand electron volts, used to express the energy of X-rays, gamma rays, electrons, and neutrons.

Laminography: The specimen and the detector move in step so that only one image plane of the specimen remains relatively sharp on the detector.

Latent image: A condition produced and persisting in the image receptor by exposure to radiation and able to converted into a visible image by processing.

Line pair test pattern: A pattern of one or more pairs of objects with high contrast lines of equal width and equal spacing. The pattern is used with an imaging device to measure spatial resolution.

Linear Accelerator: An electron generator in which the acceleration of the particles is connected with the propagation of a high-frequency field inside a linear or corrugated waveguide.

Location marker: A number or letter made of lead (Pb) or other highly radiation attenuative material that is placed on an object to provide traceability between a specific area on the image and the part.

Low-energy gamma radiation: Gamma radiation having energy less than 200 keV.

Micro focus X-ray tube: An X-ray tube having an effective focal spot size not greater than 100 μm.

Net density: Total density less fog and support (film base) density.

Neutron radiography (NRT): A process of making an image of the internal details of an object by the selective attenuation of a neutron beam by the object.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Noise: The data present in a radiological measurement which is not directly correlated with the degree of radiation attenuation by the object being examined.

Nuclear activity: The number of disintegrations occurring in a given quantity of material per unit of time. "Curie" is the unit of measurement. One curie is equivalent to 3.7×10^{10} disintegrations per second.

Optical density (photographic density): The degree of opacity of a translucent medium (darkening of film) Expressed as follows: $\mathrm{OD} = \log (i_o/i)$.

Pair production: The process whereby a gamma photon with energy greater than 1.02 MeV is converted directly into matter in the form of an electron-positron pair. Subsequent annihilation of the positron results in the production of two 0.511 MeV gamma photons.

Pencil beam: A radiation beam which has little divergence, usually created by collimating an intense source of radiation. Penetrameter—alternative term for image quality indicator. Penetrameter sensitivity—alternative term for IQI sensitivity.

Penumbra: The penumbral shadow in a radiograph.

Photodisintegration: The capture of an X-ray photon by an atomic nucleus with the subsequent emission of a particle.

Photoelectric effect: A photon transfers its energy to an electron to eject it from the atom.

Photostimulable luminescence: The physical phenomenon of phosphors absorbing incident ionizing radiation, storing the energy in quasistable states and emitting luminescent radiation proportional to the absorbed energy when stimulated by radiation of a different wavelength.

Pixel: The smallest addressable element in an electronic image.

Pixel size: The length and width of a pixel.

Primary radiation: Radiation coming directly from the source.

Rad: Unit of absorbed dose of radiation equal to 0.01 J/kg in the medium.

Radiograph: A permanent, visible image on a recording medium produced by penetrating radiation passing through the material being tested.

Radiographic contrast: The difference in density between an image and its immediate surroundings on a radiograph.

Radiographic equivalence factor: That factor by which the thickness of a material must be multiplied in order to determine what thickness of a standard material (often steel) will have the same absorption.

Real-time radioscopy: Radioscopy that is capable of following the motion of the object without limitation of time (usualy 30 frames/sec).

Röntgen (Roentgen): Unit of exposure. One Roentgen (R) is the quantity of X- or γ-radiation that produces, in dry air at normal temperature and pressure, ions carrying one electrostatic unit of quantity of electricity of either sign.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Röntgen Equivalent Man (rem): The unit of dose used in radiation protection. It is given by the absorbed rad multiplied by a quality factor, Q.

Scintillators and scintillating crystals: A detector that converts ionizing radiation to light.

Screen: Alternative term for intensifying screen.

Secondary radiation: Radiation emitted by any substance as the result of irradiation by the primary source.

Source: A machine or radioactive material that emits penetrating radiation.

Storage phosphor imaging plate: A flexible or rigid reusable detector that stores a radiological image as a result of exposure to penetrating radiation.

Target: That part of the anode of an X-ray emitting tube hit by the electron beam.

Thermal Neutrons: Neutrons having energies ranging between 0.005 eV and 0.5 eV; neutrons of these energies are produced by slowing down fast neutrons until they are in equilibrium with the moderating medium at a temperature near 20° C.

Tomography: Any radiologic technique that provides an image of a selected plane in an object to the relative exclusion of structures that lie outside the plane of interest.

Total image unsharpness: The blurring of test object features, in a radiological image resulting from any cause(s).

Transmission densitometer: An instrument that measures the intensity of the transmitted light through a radiographic film and provides a readout of the transmitted film density.

Transmitted film density: The density of radiographic film determined by measuring the transmitted light.

Tube current: The transfer of electricity, created by the flow of electrons, from the filament to the anode target in an X-ray tube; usually expressed in unit of milliamperes.

VARIABLES

N_0	Number of nuclei at time, $t = 0$
N_t	Number of nuclei remaining at a later time, t
λ_r	Decay constant
τ	Half-life of a radioisotope and is defined as the time required for the number of radioactive nuclei to be reduced by one half
Λ	Activity of a radioisotope
ϕ	X-ray flux density from a source
S	Source strength (photons produced per unit time)
r	Distance from the source

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

μ/ρ	Mass attenuation coefficient (often tabulated in cm^2/gm) and is sometimes given by μ_m
μ	Linear attenuation coefficient (units of inverse length, for example cm^{-1})
σ_{cs}	Atomic cross section the most common unit is the barn (b), equal to 10^{-24} cm^2
b	Barn, equal to 10^{-24} cm^2
A	Atomic mass
Z	Number of protons
ρ	Material volume density common units g/cm^3
mfp	Mean free path units of length and is the inverse of the linear absorption coefficient
I	Number of photons per second reaching the detector in good geometry in presence of the test material
I_0	Number of photons per second that would be detected without the test material
l	Thickness of the test material and represents the path of attenuation of the radiation—ray path
R	Ray sum, which is dimensionless
R	Historical unit of exposure or Röntgen
x, y, z	Object axes
x_r, y_r, z_r	DR/CT system axes where y_r is along the X-ray beam direction
ϕ	Angle between the X-ray beam coordinate axes and the object coordinate axes
$\vec{r}$	Linear attenuation position vector
$R(\phi, x_r)$	Radon transform
D_r	Radiation dose, or more precisely called absorbed dose, and is defined in terms of energy absorbed from any type of radiation per unit mass of any absorber. As mentioned, the historical unit of dose is the rad, which is defined as 100 ergs/g of absorber material. The SI unit of dose is the gray (Gy) defined as 1 J/kg.
LET	Linear energy transfer
Q	Quality factor, the quality factors for X- and γ-ray photons are 1.0. Q for charged particles are higher (as high as 20) and Q for neutrons, which deposit energy by creating charged particles in-situ, are also higher and vary significantly with neutron energy
H	Dose equivalent and is used as a measure of potential biological effects from a radiation field
KE_e	Kinetic energy of the moving electron
m_e	Mass of the electron
v	Velocity of the accelerated electron
V	Bias voltage

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

e	Electronic charge on an electron (1.602×10^{19}C)
h	Planck's constant (6.626×10^{34} Js)
c	Speed of light in a vacuum (2.998×10^{8} m/s)
$\lambda_{\min}$	Wavelength corresponding to maximum energy radiated
P_g	Width of the penumbra
f	Effective focal spot size
d	Distance from the object to the image plane
D_o	Distance from the source target to the object
Λ_s	Specific activity of a radioisotope
τ	Radioactive half life, this is the time for half of the originally radioactive atoms to decay into daughter atoms and is commonly used to describe the average decay rate of radioisotopic sources
λ_r	Actual probabilistic decay rate
N_A	Number of atoms per mole of material (Avogadro's number)
A_W	Atomic weight of the specific atoms (grams per mole)
n	Number of disintegrations occurring at time t
n_0	Number of disintegrations occurring at an earlier time $t = 0$—when the source strength was first measured
i_{Film}	The light-beam intensity measured with film
i_0	Light-beam intensity measured without film
D	Photographic film density
C_s	Radiographic contrast
G	Film contrast or dynamic range
ε	Radiation exposure
t	Time
T	Transmission
SNR	Signal-to-noise ratio
ΔS	Absolute difference value between two signals, usually it is the mean of the two signals
m	Mean of a signal
σ	Standard deviation of a signal
N_P	Number of primary photons
N_S	Number of scattered photons
EPS	Equivalent penetrameter sensitivities
d_p	Penetrameter thickness
h	Diameter of the smallest detectable hole in a penetrameter
D_C	Dark current measured without radiation or light
λ	Specific wavelength
d_c	Atomic spacing of a crystalline object
η	Integer order of refraction that expresses the number of wavelengths in the path differences between rays scattered by adjacent crystalline planes with spacing d

Copyright © Marcel Dekker, Inc. All rights reserved.

Z_{eff}	Effective atomic number defined by Johns (102)
Rem	Röntgen equivalent in mammals
HVL	Half-value layer and is the thickness of a material required to reduce the radiation intensity by one half its incident intensity

REFERENCES

1. Electromagnetic Spectrum. http://csep10.phys.utk.edu/astr162/lect/light/spectrum.html.
2. DL Lewis, CM Logan and MR Gabel. Continuum X-Ray Gauging of Multiple-Element Materials, Spectroscopy, 7(3): 40–46, 1992.
3. VL Gravitis, JS Watt, LJ Muldoon, EM Cockrane. Long-term trial of a dual energy γ-ray transmission gauge determining the ash content of washed coking coal on a conveyor belt. Nucl. Geophys. 1: 111, 1987.
4. IL Morgan. Real-time Digital Gauging for Process Control, Appl. Radiat. Isot. 41(10/11), 935–942, 1990.
5. P Tábor, L Molnár, D Nagymihályi. Microcomputer Based Thickness Gauges, Appl. Radiat. Isot. 41(10/11): 1123–1129, 1990.
6. R Terrass. The Life of Ed Jarman: a historical perspective. Radiology Technology 66:5:291, 1995.
7. AW Goodspeed. Excerpt from Letter to O. Glasser dated February 15, 1929. Amer J Roent. 54: 590–594, 1929.
8. WC Roentgen and O Glasser. (translated). A new kind of ray (preliminary communication). Reprinted in Wilhelm Conrad Roentgen and the Early History of Roentgen Rays, San Francisco: Norman Publishing, 1993.
9. LR Sante. Manual of Roentgenologic Technique. Ann Arbor, MI: Edwards Brothers; 1952, p 1.
10. London Standard, January 7, 1896.
11. RL Eisenberg. Radiology: An Illustrated History St. Louis: Mosby Year Book, 1991.
12. PN Goodwin, EH Quimby, and RH Morgan. Physical Foundations of Radiology. 4th ed. New York: Harper and Row; 1970, p 2.
13. MI Pupin. Roentgen Rays. Science (NY and Lancaster, PA), N.S. 3:231–235, 1898.
14. R Halmshaw. Non-destructive Testing. London: Edward Arnold, 1987, p. 16.
15. Ruth Brecher and Edward Brecher, The Rays: A History of Radiology in the United States and Canada, Baltimore: Williams & Wilkins, 1969.
16. Electrical Review, April, 1896.
17. G Swenson, Jr., H Boynton, BD Crane, DH Harris, W Jackson, J Jinata, KR Laughery, HE Martz, K Mossman, and P Rothstein, "Airline Passenger Security Screening; New Technologies and Implementation Issues," Committee on Commercial Aviation Security, Panel on Passenger Screening, National Materials Advisory Board, Commission On Engineering and Technical Systems, National Research Council, Publication MAB-482-1, National Academy Press, Washington, DC. 1996.
18. ERN Grigg. The Trail of the Invisible Light, Springfield: Charles Thomas Press, 1964; reprint, 1991.
19. F Batelli. The reaction of Roentgen rays upon various human tissues. Atti Acad Med Fis Fiorent. March 14, 1896.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

20. P Brown, American Martyrs to Science through the Roentgen Rays, 1936.
21. PW Frame. Radioactive Curative Devices and Spas, Oak Ridger newspaper, 5 November 1989.
22. GN Hounsfield. Computerized Transverse Axial Scanning (Tomography): Part 1: Description of System. Brit. J. Radiology 46: 1016–1022, 1973.
23. WS Haddad, I McNulty, JE Trebes, EH Anderson, RA Levesque, L Yang. Ultrahigh-Resolution X-ray Tomography. Science 266, 1213, 1994.
24. PD Tonner and JH Stanley. Supervoltage Computed Tomography for Large Aerospace Structures. Mats. Eval. 12: 1434, 1992.
25. See the two web sites: http://www.alft.com and http://www.jmar.com.
26. RD Evans. The Atomic Nucleus, New York: McGraw Hill, 1955. RD Evans, The Atomic Nucleus, New York: Krieger, 1982.
27. DE Cullen, et al., Tables and Graphs of Photon-Interaction Cross Sections from 10 eV to 100 GeV Derived from the LLNL Evaluated-Photon-Data Library (EPDL), UCRL-50400, Vol. 6, Rev. 4, Lawrence Livermore National Laboratory, Livermore, CA, 1989.
28. HE Martz, DJ Schneberk, SG Azevedo, SK Lynch. Computerized Tomography of High Explosives. In: Nondestructive Characterization of Materials IV, CO Ruud, et al. Eds. New York: Plenum Press, 1991) pp 187–195.
29. http://www.llnl.gov/str May, 2001 issue.
30. HH Barrett, W Swindell. Radiological Imaging. New York: Academic Press, 1981.
31. http://www-xdiv.lanl.gov/XCI/PROJECTS/MCNP/index.html
32. http://www-phys.llnl.gov/N_Div/COG/
33. GF Knoll. Radiation Detection and Measurement, New York: Wiley, NY, 2000.
34. 1990 Recommendation of the International Commission on Radiological Protection, ICRP Publication 60, Pergamon Press, Oxford (1991).
35. JH Kinney, SR Stock, MC Nichols, U Bonse, et al. Nondestructive Investigation of Damage in Composites Using X-ray Tomographic Microscopy (XTM). Journal of Materials Research, 1990 May, V5 N5:1123–1129.
36. JH Kinney, TM Breunig, TL Starr, D Haupt, et al. X-ray Tomographic Study of Chemical Vapor Infiltration Processing of Ceramic Composites Science. 260(5109): 789–792, 1993.
37. Health Physics, 60(1); The entire issue is devoted to the Goiânia radiation accident, 1991.
38. M J Yaffe and JA Rowlands. Physics in Medicine and Biology 42(1), 1997.
39. RL Weisfield, MA Hartney, RA Street, and RB Apte. New Amorphous-Silicon Image Sensor for X-Ray Diagnostic Medical Imaging Applications, SPIE Medical Imaging, Physics of Medical Imaging, 3336, pp. 444–452, 1998.
40. K Dolan, C Logan, J Haskins, D Rikard, and D Schneberk. Evaluation of an Amorphous Silicon Array for Industrial X-ray Imaging in Engineering Research, Development and Technology FY 99, Lawrence Livermore National Laboratory, Livermore, CA, UCRL 53868–99, 2000.
41. C Logan, J Haskins, K Morales, E Updike, DSchneberk, K Springer, K Swartz, J Fugina, T Lavietes, G Schmid and P Soltani. Evaluation of an Amorphous Selenium

Copyright © Marcel Dekker, Inc. All rights reserved.

Array for Industrial X-ray Imaging. Lawrence Livermore National Laboratory, Engineering NDE Center Annual Report, UCRL-ID-132315, 1998.
42. GF Knoll, Radiation Detection and Measurement, New York: Wiley, 2000.
43. L Grodzins. Optimum Energies for X-ray Transmission Tomography of Small Samples. Nuc Instr Meth 206: 541–545, 1983.
44. P Hammersberg, and M Mångård. Optimal Computerized Tomography Performance. DGZfP-Proceedings BB 67-CD, Computerized Tomography for Industrial Applications and Image Processing in Radiology, March 15–17, 1999, Berlin Germany.
45. BH Hasegawa. The Physics of Medical X-ray Imaging, Madison, WI: Medical Physics Publishing, 1991.
46. "Standard Guide for Computed Tomography (CT) Imaging," American Society for Testing and Materials, Committee E-7 on Nondestructive Testing, Designation E 1441-91, December 1991 and "Standard Practice for Computed Tomography (CT) Examination," American Society for Testing and Materials, Committee E-7 on Nondestructive Testing, Designation E 1570–93, November 1993.
47. AC Kak, and M Slaney. Principles of Computerized Tomographic Imaging. New York: IEEE Press, 1987.
48. HE Martz, GP Roberson, DJ Schneberk, and SG Azevedo. Nuclear-Spectroscopy-Based, First-Generation, Computerized Tomography Scanners. IEEE Trans. Nucl. Sci. 38: 623, 1991.
49. A Macovski, Medical Imaging Systems, Englewood Clis, New Jersey: Prentice Hall, Inc., 1983.
50. Computed Tomography, HE Martz, DJ Schneberk, GP Roberson, and SG Azevedo, Lawrence Livermore National Laboratory, Livermore, Calif., UCRL-ID-112613, September, 1992.
51. DL Lewis, CM Logan and MR Gabel. Continuum X-Ray Gauging of Multiple-Element Materials. Spectroscopy, 7(3): 40–46, 1992.
52. "Assessment of Technologies Deployed to Improve Aviation Security, First Report," TS Hartwick (chair), R Berkebile, H Boynton, BD Crane, C Drury, L Limmer, HE Martz, JA Navarro, ER Schwartz, EH Slate, M Story, Panel on Assessment of Technologies Deployed to Improve Aviation Security, National Materials Advisory Board, Commission on Engineering and Technical Systems, National Research Council, Publication NMAB-482-5, National Academy Press, Washington, DC, 1999.
53. "The Practicality of Pulsed Fast Neutron Transmission Spectroscopy for Aviation Security," P Griffin (chair), R Berkebile, H Boyton, L Limmer, H Martz, and C Oster, Jr., Committee on Commercial Aviation Security, Panel on the Assessment of Practicality of Pulsed Fast Neutron Transmission Spectroscopy for Aviation Security, National Materials Advisory Board, National Research Council, Report No. DOT/FAA/AR-99/17, March 1999.
54. Committee on Commercial Aviation Security Second Interim Report. J Beauchamp, Jr., H Boynton, BD Crane, J Denker, P Griffin, DH Harris, W Jackson, HE Martz, JC Oxley, M Story, and G Swenson, Committee on Commercial Aviation Security,

Copyright © Marcel Dekker, Inc. All rights reserved.

National Materials Advisory Board, Commission On Engineering and Technical Systems, National Research Council, Report No. DOT/FAA/AR-97-57, July 1997.
55. BD Smith. Cone-beam Tomography: Recent Advances and A Tutorial Review. Opt. Eng. 29(5): 524–534, 1990.
56. K Shields, personal communication, McClellan Nuclear Radiation Center, Sacramento, CA, 1999.
57. J Hall, F Dietrich, C Logan, and G Schmid. Development of high-energy neutron imaging for use in NDE applications. SPIE 3769: 31–42, 1999.
58. AE Pontau, AJ Antolak, DH Morse, AA Ver Berkmoes, JM Brase, DW Heikkinen, HE Martz and ID Proctor. Ion Microbeam Microtomography. Nuclear Instruments and Methods, B40/41, 646, 1989.
59. K Shields, personal communication, McClellan Nuclear Radiation Center, Sacramento, CA, 1999.
60. HE Martz, GP Roberson, DJ Schneberk, and SG Azevedo. Nuclear-Spectroscopy-Based, First-Generation, Computerized Tomography Scanners. IEEE Trans. Nucl. Sci. 38: 623, 1991.
61. DC Camp, HE Martz, GP Roberson, DJ Decman, RT Bernardi. Non-Destructive Waste-Drum Assay for Transuranic Content by Gamma-Ray Active and Passive Computed Tomography, to be submitted to Nuc. Inst. Meth. B, 2001.
62. HE Martz, GP Roberson, K Hollerbach, CM Logan, E Ashby, and R Bernardi, Computed Tomography of Human Joints and Radioactive Waste Drums, In: Proceedings of Nondestructive Characterization of Materials^{TM} IX, Sydney, Australia, June 28–July 2, 1999, AIP Conference Proceedings 497, Melville, NY, pp 607–615.
63. GP Roberson, HE Martz, and D Nisius, Assay of Contained Waste Using Active and Passive Computed Tomography, 1999 Winter Meeting, American Nuclear Society Transactions, Long Beach, CA, November 14–18, 1999, 81, pp 229–231.
64. K Kalki, JK Brown, SC Blankespoor, BH Hasegawa, MW Dae, M Chin, C Stillson. Myocardial perfusion imaging in a porcine model with a combined X-ray CT and SPECT medical imaging system. J Nucl Med 38:1535–1540, 1997 and PA Glasow. Industrial applications of semiconductor detectors. IEEE Trans. Nucl. Sci. NS-29: 1159–1171, 1982.
65. HE Martz, GP Roberson, DJ Schneberk, and SG Azevedo. Nuclear-Spectroscopy-Based, First-Generation, Computerized Tomography Scanners. IEEE Trans. Nucl. Sci. 38: 623, 1991.
66. DE Cullen, et al. Tables and Graphs of Photon-Interaction Cross Sections from 10 eV to 100 GeV Derived from the LLNL Evaluated-Photon-Data Library (EPDL), UCRL-50400, Vol. 6, Rev. 4, Lawrence Livermore National Laboratory, Livermore, CA (1989).
67. BD Cullity. Elements of X-ray Diffraction. Reading, MA: Addison Wesley, 1978.
68. RC McMaster, ed. Nondestructive Testing Handbook, Vol. 1, Section 20, Ronald Press, New York, 1959.
69. D Polansky, T Jones, R Placious, H Berger and J Reed. Real-Time Stereo X-Ray Imaging, Program Book, ASNT Spring Conf, pp 193–195, March, 1990.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

70. RD Albert and TM Albert. Aerospace Applications of X-Ray Systems Using Reverse Geometry, Materials Evaluation, 51(12), 1350–1352, Dec. 1993.
71. WA Ellingson, H Berger. Three-Dimensional Radiographic Imaging, Research Techniques in Nondestructive Testing, RS Sharpe, ed. Vol. 4, pp 1–38, London: Academic Press, 1980.
72. Z des Plantes, Eine Neue Method zur Differenzierung in der Roentgenographie, Acta Radiol. 13: 182–192, 1932.
73. Z Kolitsi, N Yoldassis, T Siozos and N Pallikarakis, Volume Imaging in Fluoroscopy. A Clinical Prototype Systems Based on a General Digital Tomosynthesis, Acta Radiol 37(5): 741–748, 1996.
74. RJ Warp, DJ Godfrey and JT Dobbins, Applications of Matrix Inverse Tomosynthesis, Proc SPIE, 3977, pp 376–383, 2000.
75. U Ewert, B Redmer and J Muller, Mechanized Weld Inspection for Detection of Planar Defects and Depth Measurements by Tomosynthesis and Planartomography, In: DO Thompson and DE Chimenti, eds, Review of progress in Quantitative Nondestructive Evaluation, Vol. 19B, pp 2093–2098, AIP Conference Proceedings 509, Am. Inst of Physics, Melville, NY, 2000.
76. M Antonakios, P Rizo and P Lamarque, Real-Time and Tangential Tomosynthesis System Dedicated to On-Line X-ray Examination of Moving Objects. In: Vol. 19B, AIP Conference Proceedings 509, Am. Inst. of Physics, Melville, NY, 2000, pp 2099–2106, 2000.
77. TS Jones and H Berger, Performance Characteristics of An Electronic Three-Dimensional Radiographic Technique, ASNT Spring Conference Summaries, pp 29–31, March, 2000.
78. JH Stanley, RN Yancey, Q Cao and NJ Dusaussoy, Reverse Engineering and Rapid Prototyping for Solid Freeform Fabrication, SPIE 2455, 306–311, 1995.
79. GA Mohr and F. Little, Effective 3D Geometry Extraction and Reverse CAD Modeling, Review of Progress in QNDE, 14: 651–656, 1994.
80. AR Crews, RH Bossi and GE Georgeson, X-ray computed tomography for geometry acquisition, Air Force Wright Laboratory Report, WL-TR-93-4306, March 1993.
81. MCH van der Meulen, Mechanical Influence on Bone Development and Adaption. In Frontiers Of Engineering: Reports on Leading Edge Engineering from 1997 NAE Symposium on Frontiers of Engineering, Third Annual Symposium on Frontiers of Engineering, National Academy Press, Washington DC, 1998, pp 12–15.
82. "Application of 3d x-ray ct data sets to finite element analysis," P-L Bossart, HE Martz, HR Brand, K Hollerbach, In: DO Thompson, DE Chimenti, eds. Review of Progress in Quantitative Nondestructive Evaluation 15, New York: Plenum Press, 1996) pp 489–496.
83. K Hollerbach, and A Hollister, Computerized prosthetic modeling, Biomechanics, 31–38, September 1996.
84. AM Waters, HE Martz, CM Logan, E Updike, and RE Green, Jr., High Energy X-ray Radiography and Computed Tomography of Bridge Pins, Proceedings of the Second Japan-US Symposium on Advances in NDT, Kahuku, Hawaii, June 21–25, 1999, pp 433–438.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

85. RL Weisfield, MA Hartney, RA Street, and RB Apte, New Amorphous-Silicon Image Sensor for X-Ray Diagnostic Medical Imaging Applications, SPIE Medical Imaging, Physics of Medical Imaging, 3336: 444–452, 1998.
86. HE Martz, GP Roberson, MF Skeate, DJ Schneberk, SG Azevedo, and SK Lynch, High Explosives (PBX9502) Characterization Using Computerized Tomography, UCRL-ID-103318, Lawrence Livermore National Laboratory, Livermore, CA, 1990.
87. HE Martz, DJ Schneberk, SG Azevedo, and SK Lynch, Computerized Tomography of High Explosives, In: CO Ruud et al. Nondestructive Characterization of Materials IV, Eds. New York: Plenum Press, 1991, pp 187–195.
88. DE Perkins, HE Martz, LO Hester, G Sobczak, and CL Pratt, Computed Tomography Experiments of Pantex High Explosives, Lawrence Livermore National Laboratory, Livermore, Calif., UCRL-CR-110256, April 1992.
89. K Kouris, NM Spyrou, and DF Jackson, Materials Analysis Using Photon Attenuation Coefficients, In: RS Sharpe, et. Research Techniques in Nondestructive Testing Vol. VI, New York: Academic Press, 1982.
90. BM Dobratz and PC Crawford, LLNL Explosives Handbook Properties of Chemical Explosives and Explosive Simulants, UCRL-52997, Lawrence Livermore National Laboratory, Livermore, CA, 1985.
91. H. E. Martz, G. P. Roberson, D. J. Schneberk, and S. G. Azevedo, Nuclear-Spectroscopy-Based, First-Generation, Computerized Tomography Scanners, IEEE Trans. Nucl. Sci. 38: 623, 1991.
92. HE Martz, GP Roberson, MF Skeate, DJ Schneberk, SG Azevedo, SK Lynch, High Explosives (PBX9502) Characterization Using Computerized Tomography, UCRL-ID-103318, Lawrence Livermore National Laboratory, Livermore, CA, 1990.
93. HE Martz, GP Roberson, MF Skeate, DJ Schneberk, SG Azevedo, SK Lynch, High Explosives (PBX9502) Characterization Using Computerized Tomography, UCRL-ID-103318, Lawrence Livermore National Laboratory, Livermore, CA, 1990.
94. HE Martz, DJ Schneberk, SG Azevedo, and SK Lynch, Computerized Tomography of High Explosives, In: CO Ruud, et al. Eds. Nondestructive Characterization of Materials IV, New York: Plenum Press, 1991, pp 187–195.
95. F A Iddings. Back to Basics: Editorial, Materials Evaluation 37, Columbus, Ohio, American Society for Nondestructive Testing, Inc., 1979, p 20.
96. SA McGuire and CA Peabody. Working Safely in Gamma Radiography. Washington, DC, NUREG/BR-0024, June 1986.
97. USAEC. Public Meeting on Radiation Safety for Industrial Radiographers. Springfield, VA. National Technical Information Services. December 1978.
98. J Bush. Gamma Radiation Safety Study Guide. Columbus, Ohio: The American Society for Nondestructive Testing, Inc., 1999
99. Anon. E1316 Section D: Gamma- and X-Radiology. In: Annual Book of ASTM Standards, 03.03. Philadelphia: American Society for Testing and Materials, 2000.
100. Anon. Borders' Dictionary of Health Physics. The Borders' Consulting Group. www.hpinfo.org

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

101. BR Scott. Bobby's Radiation Glossary for Students. www.lrri.org/ radiation/ radgloss.htm
102. H Johns and J Cunningham. The Physics of Radiology, Fourth Ed. Charles C. Thomas, 1983.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

8

Active Thermography

Jane Maclachlan Spicer and
Robert Osiander
Johns Hopkins University, Laurel, Maryland

8.1 INTRODUCTION/BACKGROUND

8.1.1 Technique Overview

The thermal *nondestructive evaluation* (NDE) technique termed "active thermography" integrates infrared imaging with external heating to assess subsurface structure via the thermal response of the sample. It is important to distinguish this approach from that used in the predictive maintenance community where temperature distributions of operating machinery or electrical systems are imaged to locate hot spots indicative of operating problems. While this is a very important field on its own, it does not involve the use of an external heating source as is required in active thermography.

There are a number of different variants of thermal NDE techniques, each carrying its own name, acronym and particular characteristics. Since this field has developed principally under the direction of individual researchers, there is little standardization of the terminology as yet. We will use the term "active thermography" in this chapter as an encompassing term for all NDE work performed with an infrared camera, but Table 8.1 summarizes a number of the common terms used in describing thermal NDE techniques currently in use.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

TABLE 8.1 Descriptors of Thermal Inspection Methodologies

Technique	Acronym	Characteristics
Active thermography [1]		
Optical beam deflection [2]	OBD	
Photothermal radiometry [3]	PTR	Point measurement with single cell detector
Pulse echo thermal wave imaging [4]		
Pulsed infrared imaging		
Scanning photoacoustic microscopy [5]	SPAM	Requires enclosed gas cell
Stimulated infrared thermography [6]		
Thermal pulse video thermography [7]	PVT	
Time-resolved infrared radiometry [8]	TRIR	Employs step heating
Thermal nondestructive testing [9]	TNDT	
Transient thermography [10]		

Common to all active thermography techniques are four major components:

1. Controlled heating of the specimen, typically by optical absorption on the sample surface
2. Thermal transport of heat into and within the specimen
3. Imaging of the resulting specimen surface temperature distribution using an infrared camera
4. Interpretation of the spatial and temporal features of the temperature distribution to provide information about material and structural properties of the specimen

Each of these elements will be described in detail in the following sections.

As an initial example of the importance of understanding the temporal development of the temperature distribution in an active thermographic measurement consider the three infrared images shown in Figure 8.1 acquired at different times (3.9, 7.8 and 15.6 s) during heating of a series of flat-bottomed holes in a graphite epoxy panel. These holes were milled from the back surface of the panel to different depths, resulting in a series of subsurface defects at different distances from the top surface, but without any visible indication of their presence from the top surface. The defects are in groups of three at depths of 2.8, 1.4 and 0.8 mm

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

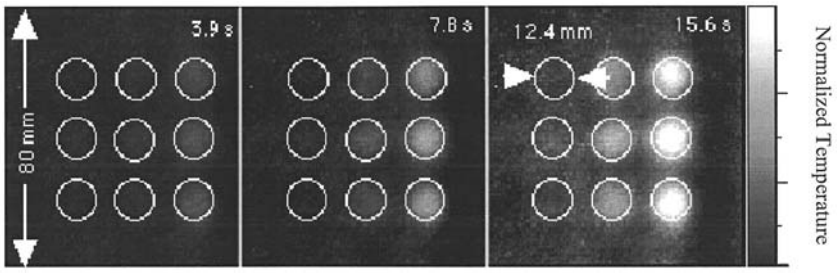


FIGURE 8.1 Infrared images of flat-bottomed holes, 12.4 mm diameter, at depths of 0.8, 1.4 and 2.8 mm, in graphite-epoxy composite panel at different heating times. The locations of the subsurface holes are delineated by the white circles.

starting with the left column. In the infrared images, the defects become visible at different times (the white circles in the images were added to show the hole positions) and the brightness of the response increases with time. Note that the shallower defects at the right hand side of the images appear at earlier times in the infrared images than the deeper defects on the left. This time-dependence of the surface temperature distribution as revealed in the infrared images is at the heart of how active thermography can provide quantitative NDE information.

8.1.2 Historical Perspective

Background

The use of heat flow for characterizing materials and defects has had a long and varied history. Many of the fundamental studies in NDE and materials characterization that lay the foundations for the active thermography methods currently practiced can be found in the fields of photothermal and photoacoustic phenomena. The different classes of photothermal characterization methods all exhibit common elements: optical heating of the sample, thermal transport in the sample, and subsequent detection of the surface temperature of the sample. Detection techniques used include the photoacoustic method and photothermal beam deflection (also known as the "mirage effect"). In another detection scheme, photothermal radiometry, the surface temperature is measured via the intensity of the emitted infrared radiation, a method that developed into active thermography as described in this chapter by replacing the single infrared detector with an infrared camera. All of these methods have been used in an imaging mode, where the sample or the exciting laser source is scanned mechanically and an image is generated serially. The rapid evolution of infrared imaging technology through the 1980s and 1990s allowed the development of techniques which permit real-

Copyright © Marcel Dekker, Inc. All rights reserved.

time full-field imaging using infrared cameras for detection. Simultaneously, the development of faster computers has allowed the manipulation of the very large data sets of time-dependent images to be managed much more readily than in the early 1980s.

The very earliest experiments that incorporated principles similar to those found in thermal NDE measurements were performed by Alexander Graham Bell in the 1880s. He constructed a device called a photophone (see Figure 8.2) which used a hearing-tube attached to a transparent chamber in which modulated sunlight was allowed to fall on the sample of interest. With this device he was able to investigate "photophonic phenomena" in common substances and also performed some spectroscopic measurements (11). The concept of a modulated heating source for use in the characterization of material properties was not new, even for Bell, and had been used for the measurement of thermal conductivity. Angstrom made the first measurements of this type in the 1860s (12). He periodically heated and cooled one end of a long thin bar until a steady state situation was achieved. Then, by measuring the temperature at two points along the bar he was able to calculate the thermal conductivity. Variations of this technique have been used by a number of researchers (e.g. Starr (13)) to measure thermal transport properties.

Photoacoustic Effect

A big revival of interest in thermal NDE measurements, in part due to the evolution of electronic instrumentation, came in the 1970s with the development

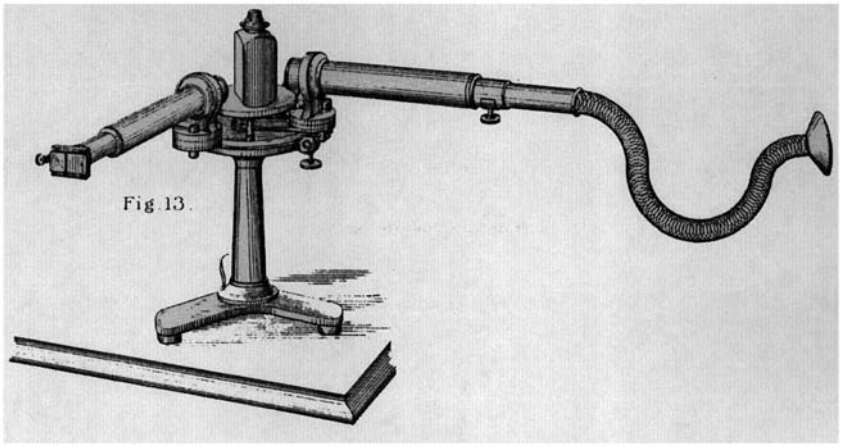


FIGURE 8.2 Alexander Graham Bell's photophone used to investigate photophonic phenomena.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

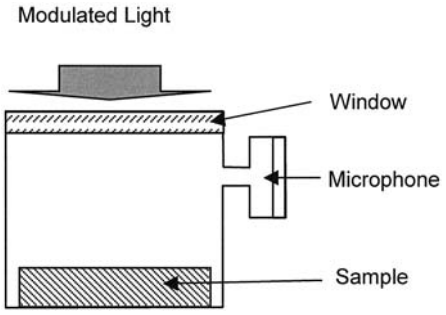


FIGURE 8.3 Photoacoustic gas cell.

of the field of *photoacoustic spectroscopy* (PAS) by Rosencwaig and others (14–17). In a photoacoustic measurement, the sample of interest is placed in an enclosed cell filled with gas (see Figure 8.3). The sample is then heated with a periodically-modulated light source, such as a chopped laser beam, and a microphone connected to the cell gas detects periodic pressure variations in the gas caused by periodic heat flow from the sample into the gas. The smallest amplitudes detectable with a microphone correspond to periodic surface temperature variations smaller than 10^{-3} Kelvin. Since the signal depends on the absorbed light intensity, the photoacoustic effect can be used for absorption spectroscopy measurements (PAS). The signal also depends on the thermal properties of the sample that control the flow of heat into the gas and the sample. A comprehensive analytical description for the photoacoustic effect has been given by Rosencwaig and Gersho (18), which is the foundation of almost all thermal NDE techniques. There were a number of early attempts at imaging using a photoacoustic cell and scanning either the cell or the focused laser beam mechanically. Such work has been termed *scanning photoacoustic microscopy* (SPAM) and from the early stages was applied to NDE problems. For example, Wong et al. imaged cracks in silicon-nitride ceramics used in turbine blades (19) and studied the ability to detect subsurface flaws in silicon-nitride (5). Subsurface flaws in metals were studied both theoretically and experimentally by Thomas et al. (20).

Optical Beam Deflection

Optical beam deflection (OBD) techniques for thermal NDE measurements were first developed around 1980 and have since been used in a wide range of spectroscopy and imaging applications. In this technique the deflection of a probe laser beam in the temperature gradient in the gas above the heated sample surface is used to monitor the specimen surface temperature (see Figure 8.4). Boccara

Copyright © Marcel Dekker, Inc. All rights reserved.

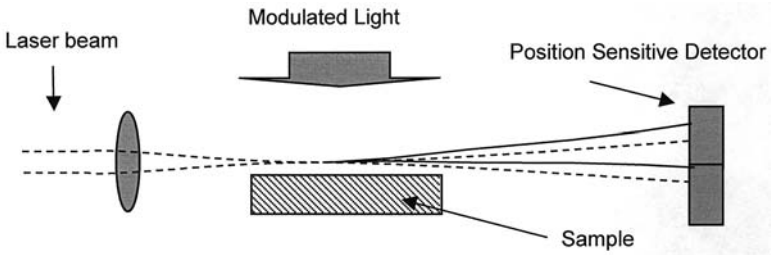


FIGURE 8.4 Schematic diagram of the ''mirage'' effect.

et al. (21) first reported the use of a deflection or "mirage"* technique for studying absorption spectra of materials and found that the surface temperature resolution of this technique is 10^{-4} K. Further developments of this technique were presented by Murphy and Aamodt (22–24), who also gave an analytical description based on Rosencwaig's models for the photoacoustic effect.

Optical beam deflection techniques are noncontacting and do not require an enclosed cell, which makes them suitable for imaging applications. Typically for such measurements the light source is a laser beam (pump beam) focused below the probe laser beam, and the sample is scanned mechanically as first described by Fournier (2). Early examples of imaging studies with OBD detection include detection of subsurface structure (25) and imaging of surface-breaking fatigue cracks (26).

Photothermal Radiometry

Another noncontacting method for measuring the surface temperature of a periodically-heated sample is to monitor the emitted infrared radiation. The use of infrared detection in thermal NDE measurements was demonstrated early on in the development of the field and has been described by the term *photothermal radiometry* (PTR) (3). In this technique the surface temperature variation of the sample is measured as a change in the infrared emission of the sample surface according to the Stefan-Boltzmann law. Early imaging results were reported by Busse (27) who imaged an integrated circuit in a transmission mode by placing the infrared detector on the opposite side of the sample from the heating laser source and scanning the sample. This approach provides the foundation of the active thermography methods which are the focus of this chapter.

8.1.3 Potential of Active Thermography Techniques

While many of the photothermal techniques have been used for spectroscopy, active thermography focuses on the detection of subsurface structures and defects

* Note that a desert mirage also is caused by a temperature gradient above the hot sand.

Copyright © Marcel Dekker, Inc. All rights reserved.

utilizing the variation in thermal properties between the defect and the host material. Active thermography has developed from photothermal radiometry with the evolution of infrared cameras, that allow images of the surface temperature distribution of a sample to be taken at video rates. Analytical models developed for the photothermal techniques allow interpretation of the surface temperature variations as a function of time and can determine the subsurface structure of the sample. A typical experimental setup is shown in Figure 8.5. A sample is heated with a modulated light source, in this case a laser beam, and the temperature distribution is recorded with the infrared camera as a function of time.

The variations between techniques used in active thermography are in the types of heating sources used and their different spatial and temporal heating patterns. We will find that each method has its advantages and disadvantages, but often can be chosen to match a specific problem. For all techniques, the temporal and spatial development of the surface temperature measured after a heat load is applied (typically by optical heating) is compared to an analytical model at each position resolved by the camera. Many defect types of interest such as voids, cracks, and delaminations in composites are very readily detected with active thermography. Among the strengths of active thermography is the ability to



FIGURE 8.5 Typical experimental setup for active thermography.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

conduct the measurements in a totally noncontacting manner with a large standoff distance if needed. This can be particularly important in environments with a moving product or when contact with the sample, or immersion in a fluid as used in ultrasonic scanning is not practical. Another strength of these techniques is that the material property being probed is its thermal behavior which is quite unique and different from the properties probed in other NDE methods such as ultrasonics (probes elastic properties), radiography (probes variations in Z), or eddy current (probes variations in electrical properties).

8.2 BASICS OF HEAT DIFFUSION

Active thermography relies on the fact that the conduction of heat in a structure allows the deduction of information about the structure itself and about its component materials. The structure and the thermal properties of the materials govern heat conduction. The thermal properties we will encounter in this section are summarized in Table 8.2. For the static case, the thermal properties of importance are the *specific heat*, c, and the *thermal conductivity*, κ. We will then examine how c and κ relate to the thermal properties in the case of time-dependent temperature changes and discuss a number of cases which are of importance to nondestructive evaluation including plates of finite thickness with different backing materials. Finally, we will compare the cases of pulsed and step heating, the two approaches most commonly used in active thermography.

8.2.1 Steady State Heat Flow

Specific Heat

The *heat capacity*, C, of a body describes how the temperature of a body changes as it is heated or cooled and is defined as

$$C = \frac{\Delta Q}{\Delta T} \tag{1}$$

TABLE 8.2 Summary of Thermal Properties

Symbol	Term	Units (cgs)	Units (SI)
Q	Energy	cal	J
ρ	Density	g/cm^3	kg/m^3
κ	Thermal conductivity	cal/s cm K	W/m K
c	Specific heat	cal/g K	J/kg K
α	Thermal diffusivity	cm^2/sec	m^2/sec
e	Thermal effusivity	$cal/cm^2\,sec^{1/2}\,K$	$W.sec^{1/2}/m^2\,K$
Γ	Thermal mismatch factor	(numeric)	(numeric)
t_{trans}	Thermal transit time	sec or $sec^{1/2}$	sec or $sec^{1/2}$

Copyright © Marcel Dekker, Inc. All rights reserved.

where ΔQ is the heat supplied to the body and ΔT is the resulting temperature change. One of the most important thermal properties of any material is the specific heat, c—the heat capacity per unit mass of the material of which the body is composed. The unit for the specific heat is J/kgK. The specific heat allows us to calculate the change in heat content, $\Delta Q(\mathrm{J/m^3})$, of a body of mass, m, from its temperature change, ΔT:

$$\Delta Q = cm\Delta T \tag{2}$$

For all materials, the specific heat is temperature-dependent. Values for the specific heat at room temperature for a number of common materials are given in Table 8.3. Note that while the values for specific heat for different materials vary by only about 1 to 2 orders of magnitude, other properties such as the thermal diffusivity and conductivity, which will be described in later sections, can vary by many more orders of magnitude.

Thermal Conductivity

When different parts of a body are at different temperatures, heat flows from the hotter parts to the cooler parts until the temperature distribution is homogeneous. This process is called heat conduction. Other transport mechanisms for heat, which we will not consider in the following discussion since they do not play an

TABLE 8.3 Thermal Properties at Room Temperature of Common Materials Listed in Order of Increasing Thermal Diffusivity

Material	Thermal diffusivity $\mathrm{cm^2/sec}$	Thermal effusivity $\mathrm{J/(cm^2\ Ks^{1/2})}$	Thermal conductivity J/(cm K s)	Specific heat J/(g K)	Density $\mathrm{g/cm^3}$
Polyisoprene	7.709×10^{-4}	5.57×10^{-4}	13.4×10^{-4}	1.90	0.913
Water	1.45×10^{-3}	6.02×10^{-3}	6.05×10^{-3}	4.18	0.997
Glass	3.43×10^{-3}	4.23×10^{-3}	7.79×10^{-3}	0.842	2.7
Zirconia	2.19×10^{-3}	4.60×10^{-3}	6.49×10^{-3}	0.582	5.1
Ni superalloy	0.0260	0.083	0.095	0.440	8.3
Air	0.221	7.43×10^{-8}	26.2×10^{-5}	1.01	1.18×10^{-3}
Aluminum alloy (2024-T4)	0.46	0.775	1.22	0.963	2.77
Pure aluminum	0.967	1.366	2.35	0.901	2.699
Silicon	1.08	0.641	1.70	0.679	2.33
Copper	1.17	3.30	4.01	0.385	8.936
Gold	1.26	1.92	3.18	0.131	19.32
Diamond	3.74	2.79	6.62	0.502	3.516

Copyright © Marcel Dekker, Inc. All rights reserved.

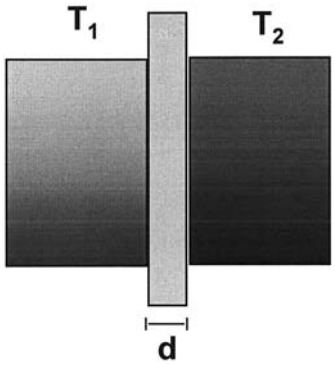


FIGURE 8.6 Schematic for heat conduction across a plate of thickness, *d*.

important role in active thermography at room temperature, are radiative and convective heat transfer. In radiative heat transfer heat is transferred via electromagnetic radiation such as infrared light*, while in convective heat transfer heat is transported by the motion of the heated substance such as a gas or liquid.

Imagine a situation as shown in Figure 8.6 where two heat reservoirs at different temperatures T_1 and T_2, which are held constant, are separated by a plate of thickness, d. Energy in the form of heat flows through the plate, and the heat flow, J in W/m^2, is proportional to the temperature gradient:

$$J = -\kappa \frac{\partial T}{\partial x} = -\kappa \frac{T_2 - T_1}{d} \tag{3}$$

The minus sign indicates that heat flows from the hotter to the colder region. In order to hold the temperatures of the two reservoirs constant, note that one reservoir has to be heated and the other needs to be cooled. The proportionality constant k is called the *thermal conductivity*. It is measured in W/m K and is temperature-dependent for most materials. Eq. (3) is also called *Fourier's law* and it only holds for the static or time-independent case.

The static case can be used to make nondestructive thickness measurements as shown in Figure 8.7. Imagine that one side of the plate under investigation is kept at a constant temperature. The surface temperature on the other side of the plate then depends on the heat flowing through the plate. If the temperatures on both sides of the plate are measured, and the power, P, per unit area, A, of the plate required to maintain the temperature at the heated side is recorded, the

*This will be discussed in Section 8.2.1 since this is the basis of infrared radiometry.

Copyright © Marcel Dekker, Inc. All rights reserved.

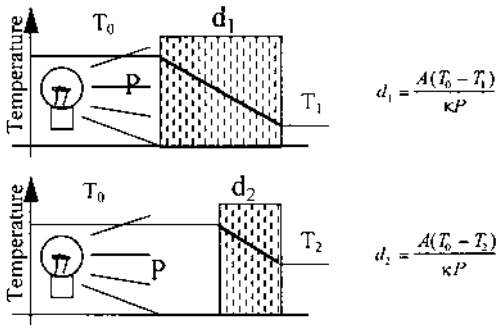


FIGURE 8.7 Thickness measurement using Fourier's Law.

thickness, d_x, of the material (if its thermal conductivity is known) can be determined:

$$J_x = \frac{P}{A} = \kappa \frac{(T_0 - T_x)}{d_x} \Rightarrow d_x = \frac{A(T_0 - T_x)}{\kappa P} \tag{4}$$

8.2.2 Conduction of Heat in One Dimension

Fourier's Law is only applicable for the case when the temperature distribution in the system has reached equilibrium. We can derive the time-dependent equation when we account for all heat sources and apply the concept of conservation of energy. We consider a plate of thickness Δx; density, ρ; specific heat, c; thermal conductivity, κ, and area, A, as shown in Figure 8.8. The heat entering the plate by diffusive heat flow at position $x = 0$ in the short time interval, Δt is given by Eq. (5)

$$Q_1 = J_1 A \Delta t = \left(-\kappa \frac{\partial T}{\partial x}\right)_{x=0} A \Delta t \tag{5}$$

and by

$$Q_2 = \left(\kappa \frac{\partial T}{\partial x}\right)_{x=\Delta x} A \Delta t \tag{6}$$

at position $x = \Delta x$ (note the sign change). The heat input from a heating source, H (in W/m^3), is:

$$Q_H = HA\Delta t \Delta x \tag{7}$$

and the change of heat in the plate for a temperature rise, ΔT, is given by

$$\Delta Q = c\rho \cdot A \Delta x \cdot \Delta T \tag{8}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

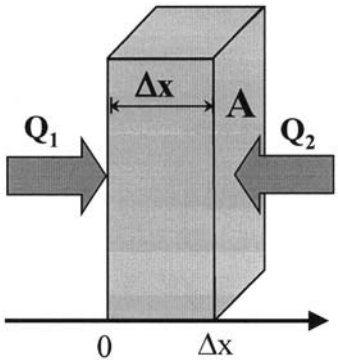


FIGURE 8.8 Schematic used to derive the differential equation for the conduction of heat.

The diffusive heat flow into the plate and the input from the external heating source has to equal the total heating in the plate, $Q_1 + Q_2 + Q_H = \Delta Q$. If we combine Eq. (5) through Eq. (8) we get

$$\left[\left(\kappa \frac{\partial T}{\partial x}\right)_{x=\Delta x} - \left(\kappa \frac{\partial T}{\partial x}\right)_{x=0}\right] A\Delta t + HA\Delta t\Delta x = c\rho \cdot A\Delta x \cdot \Delta T \tag{9}$$

Dividing by $A\Delta t\Delta x$ and taking the derivatives as the limit for small $\Delta t\Delta x$, gives:

$$\kappa \frac{\partial^2 T}{\partial x^2} - c\rho \cdot \frac{\partial T}{\partial t} = -H \tag{10}$$

Then introducing the *thermal diffusivity*, α, as

$$\alpha = \frac{\kappa}{c\rho} = \frac{\text{thermal conductivity}}{\text{specific heat density}} \tag{11}$$

we get the differential equation of conduction of heat in one dimension:

$$\frac{\partial^2 T(x,t)}{\partial x^2} - \frac{1}{\alpha} \cdot \frac{\partial T(x,t)}{\partial t} = -\frac{H(x,t)}{\kappa} \tag{12}$$

Eq. (12) is a diffusion equation for the temperature, and therefore the thermal diffusivity has the units of a diffusion coefficient, cm^2/s. In time-dependent heat flow problems, thermal diffusivity is the parameter of interest, not simply the thermal conductivity. Values for the thermal diffusivity of a number of common materials are shown in Table 8.3. Note that while metals (which have high values of thermal conductivity), also have high thermal diffusivities, and ceramics and glasses (which have lower thermal conductivities) also have low thermal

Copyright © Marcel Dekker, Inc. All rights reserved.

diffusivities, air, which has a *low* thermal conductivity, has a *high* thermal diffusivity due to its low density.

8.2.3 Periodic Solutions to the Heat Conduction Equation

A solution for a differential equation such as Eq. (12) can be composed from a general solution of the homogeneous equation (where the heat input, H is set to zero),

$$\frac{\partial^2 T(x,t)}{\partial x^2} - \frac{1}{\alpha} \cdot \frac{\partial T(x,t)}{\partial t} = 0 \tag{13}$$

plus one special solution which fulfills the appropriate boundary conditions with a given heat input. Assume a solution in the form of a temperature variation which is periodic in time,

$$T(x,t) = T(x)\exp(i\omega t) \tag{14}$$

If we substitute Eq. (14) into Eq. (13) we get a differential equation describing the spatial dependence of the temperature distribution,

$$\frac{\partial^2 T(x)}{\partial x^2} - \frac{i\omega}{\alpha} = 0 \tag{15}$$

which is solved by

$$T(x) = A\exp(\sigma x) + B\exp(-\sigma x) \quad \text{with} \quad \sigma^2 = \frac{i\omega}{\alpha} \tag{16}$$

where the constants A and B are to be determined by the boundary conditions. Substituting Eq. (16) into Eq. (14) gives the general solution for the homogeneous equation, Eq. (13), as

$$T(x,t) = [A\exp(\sigma x) + B\exp(-\sigma x)]\exp(-i\omega t) \quad \text{with} \quad \sigma = (1+i)\sqrt{\frac{\omega}{2\alpha}} \tag{17}$$

The solution given by Eq. (17) looks a lot like the solution of a wave equation, which has given rise to the use of the term "thermal waves" in describing heat flow resulting from a periodic heating source. There have been numerous experiments which have shown typical wave characteristics such as interference with thermal waves (28). Note that Eq. (17) describes a periodic temperature variation with positive and negative excursions about zero. This expression only describes the time-dependent temperature differences. Actually, a temperature increasing linearly in time, $T = T_0 + Cx + D(t + x^2/2\alpha)$ is also a general solution of Eq. (13) and should also be included in Eq. (17), since the periodic variation is actually about a continually increasing median temperature level. For simplicity, we will constrain ourselves to the solution shown in Eq. (17) which

Copyright © Marcel Dekker, Inc. All rights reserved.

has some very interesting properties. The temperature is attenuated to e^{-1} ($=0.368$) in a length δ_{th},

$$\delta_{th} = \sqrt{\frac{2\alpha}{\omega}} \tag{18}$$

which is called the thermal diffusion length. This parameter is frequency-dependent and the periodic temperature variations do not penetrate far into the sample for high heating source modulation frequencies and low specimen thermal diffusivities. It should be noted that the thermal diffusion length is analogous to the skin depth for electromagnetic waves in conductors.

We will now consider two special cases for Eq. (17), the infinite half-space and the plate of finite thickness. The plate of finite thickness will be representative of many situations in nondestructive evaluation such as coating layers debonding from substrates, composites with debonding between plies, or horizontal cracks beneath the surface of a test object.

Infinite Half-Space

Let us consider an infinite half-space as shown in Figure 8.9 with the sample heated periodically by light absorbed at its surface. We can determine the temperature variations in the air and in the sample, if we solve the boundary conditions for the constants A and B in Eq. (17). Since only a finite solution is of physical significance, we therefore must set $A_{sample} = B_{air} = 0$ in order to avoid an ever increasing temperature with distance into either the air or the sample. The solutions T_{air} ($x < 0$) and $T_{sample}(x > 0)$ are

$$\begin{aligned} x < 0: \quad & T_{air} = A_{air} \exp(\sigma_{air} x + i\omega t) && \sigma_{air} = (1+i)\sqrt{\frac{\omega}{2\alpha_{air}}} \\ x > 0: \quad & T_{sample} = B_{sample} \exp(-\sigma_{sample} x + i\omega t) && \sigma_{sample} = (1+i)\sqrt{\frac{\omega}{2\alpha_{sample}}} \end{aligned} \tag{19}$$

At the interface ($x = 0$), the temperature *must* be continuous and the heat flux in the sample and in air,

$$J = -\kappa \frac{\partial t}{\partial x} \tag{20}$$

must take into account the heat, $H_0 \exp(i\omega t)$, delivered at the surface. We get a system of two equations:

$$\begin{aligned} A_{air} &= B_{sample} \\ -\kappa_{air}\sigma_{air} A_{air} + H_0 &= \kappa_{sample}\sigma_{sample} B_{sample} \end{aligned} \tag{21}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

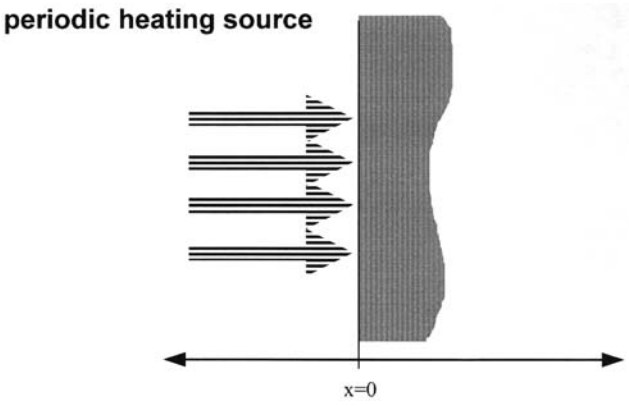


FIGURE 8.9 An infinite half-space subjected to periodic heating at the surface.

which can be solved and gives the temperature in the sample and the air as

$$T_{\text{sample}}(x, t) = \frac{H_0 \ \exp(-\sigma_{\text{sample}} x + i\omega t)}{(1+i)\sqrt{2\omega}\left(e_{\text{air}} + e_{\text{sample}}\right)} \tag{22a}$$

$$T_{\text{air}}(x, t) = \frac{H_0 \ \exp(\sigma_{\text{air}} x + i\omega t)}{(1+i)\sqrt{2\omega}\left(e_{\text{air}} + e_{\text{sample}}\right)} \tag{22b}$$

Here we have introduced a new thermal property, the *thermal effusivity*, e, which is defined as

$$e = \sqrt{c\rho\kappa} = \frac{\kappa}{\sqrt{\alpha}} \propto \sqrt{\text{heat capacity} \cdot \text{density} \cdot \text{thermal conductivity}} \tag{23}$$

Eq. (22) shows that the surface temperature is proportional to the inverse sum of the effusivity of the two materials. In other words, for a material with a high effusivity such as a metal the temperature response to a heat input is smaller than for a material with a lower thermal effusivity. The unit for thermal effusivity is $\text{J/m}^2\,\text{K}\,\text{s}^{1/2}$. Values for the thermal effusivities of different materials are given in Table 8.3.

Copyright © Marcel Dekker, Inc. All rights reserved.

Note that the temperatures in Eq. (22) are complex. Separating these expressions into magnitude and phase factors gives

$$T_{\text{sample}}(x,t) = \frac{H_0 \exp\left(-\sqrt{\frac{\omega}{2\alpha_{\text{sample}}}}x\right)}{\sqrt{\omega}\left(e_{\text{air}} + e_{\text{sample}}\right)} \exp\left[i\left(\omega t - \sqrt{\frac{\omega}{2\alpha_{\text{sample}}}}x - \frac{\pi}{4}\right)\right]; \tag{24a}$$

$$T_{\text{air}}(x,t) = \frac{H_0 \exp\left(-\sqrt{\frac{\omega}{2\alpha_{\text{air}}}}x\right)}{\sqrt{\omega}\left(e_{\text{air}} + e_{\text{sample}}\right)} \exp\left[i\left(\omega t + \sqrt{\frac{\omega}{2\alpha_{\text{air}}}}x - \frac{\pi}{4}\right)\right]. \tag{24b}$$

These equations show there is a phase shift of $\pi/4$ (45°) with respect to the heating source at all times. Figure 8.10 shows the temperature variation as a function of position in the infinite half-space and for the air above the half-space for a series of times. Note that the envelope of the curves demonstrates the strong damping of the thermal response with depth. The thermal diffusion length described in Eq. (18) is indicated on this graph.

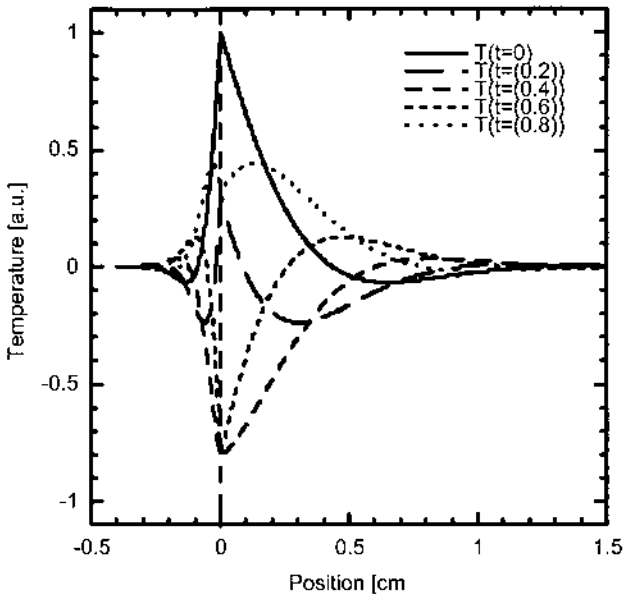


FIGURE 8.10 Temperature variation in the infinite half-space ($\alpha = 0.01\,\text{cm}^2/\text{sec}$) and for the air above the half-space ($\alpha = 0.24\,\text{cm}^2/\text{sec}$) at a periodic heating frequency of 1 Hz.

Copyright © Marcel Dekker, Inc. All rights reserved.

Plate of Finite Thickness

A model of practical interest for NDE applications is a plate of finite thickness, as shown in Figure 8.11, where a backing layer of a different material is attached to the plate of thickness, d. The temperature distributions in the three regions can be calculated using Eq. (17). Note that for the sample ($0 < x < d$) both coefficients A and B are nonzero because of the finite thickness, d of the plate.

$$
\begin{aligned}
&x < 0: \qquad T_{\text{air}} = A_{\text{air}} \exp(\sigma_{\text{air}} x + i\omega t)\\
&0 < x < d: \; T_{\text{sample}} = \left[A_{\text{sample}} \exp(\sigma_{\text{sample}} x) + B_{\text{sample}} \exp(-\sigma_{\text{sample}} x)\right] \exp(i\omega t)\\
&x > d: \qquad T_{\text{backing}} = B_{\text{backing}} \exp(-\sigma_{\text{backing}} x + i\omega t)\\
&\sigma_{\text{air}} = (1+i)\sqrt{\frac{\omega}{2\alpha_{\text{air}}}}; \;\; \sigma_{\text{sample}} = (1+i)\sqrt{\frac{\omega}{2\alpha_{\text{sample}}}};\\
&\sigma_{\text{backing}} = (1+i)\sqrt{\frac{\omega}{2\alpha_{\text{backing}}}}
\end{aligned}
\tag{25}
$$

With the same boundary conditions as before, but now at both $x = 0$ and $x = d$, continuity of temperature and continuity of heat flow with a source, H_0 at $x = 0$, we get a system of 4 equations that can be solved for the 4 variables A_{air},

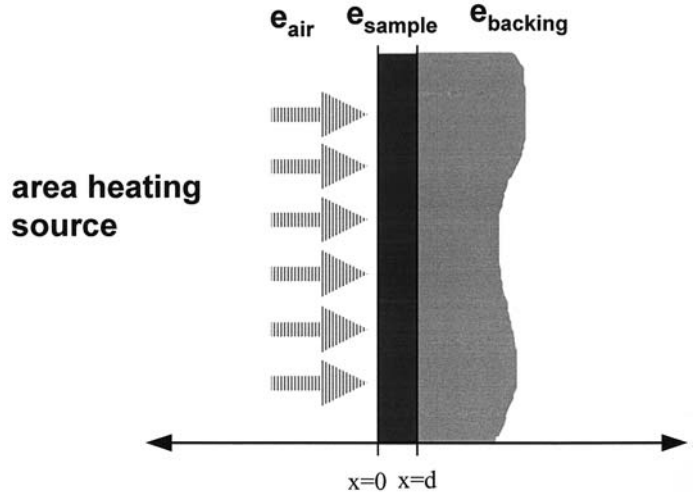


FIGURE 8.11 Schematic for a plate of finite thickness, d, with a backing layer.

Copyright © Marcel Dekker, Inc. All rights reserved.

A_{sample}, B_{sample}, and B_{backing}. With these variables the temperature at the surface, $T_{\text{sample}}(0, t)$ is

$$T(0,t) = \frac{H_0 \exp(i\omega t)}{\sqrt{2\omega}(1+i)(e_{\text{air}} + e_{\text{sample}})} \frac{1 - \Gamma_{\text{backing}} \exp(-2\sigma_{\text{sample}} d)}{1 - \Gamma_{\text{air}}\Gamma_{\text{backing}} \exp(-2\sigma_{\text{sample}} d)} \quad (26)$$

where we have introduced the *thermal mismatch factor*, Γ for the backing and for the air:

$$\Gamma_{\text{backing}} = \frac{e_{\text{backing}} - e_{\text{sample}}}{e_{\text{backing}} + e_{\text{sample}}}; \; \Gamma_{\text{air}} = \frac{e_{\text{air}} - e_{\text{sample}}}{e_{\text{air}} + e_{\text{sample}}} \quad (27)$$

The thermal mismatch factor describes how well heat flows from one material into another material. Because air has a very small effusivity we can set $e_{\text{air}} = 0$ and also set $\Gamma_{\text{air}} = -1$. If we then use the series expansion for $1/(1+x) = \sum_{n=0}^{\infty} (-x)^n$ we obtain

$$\begin{aligned} T(0,t) &= \frac{H_0 \exp(i\omega t)}{\sqrt{2\omega}(1+i)e_{\text{sample}}} \left[1 + 2\sum_{n=1}^{\infty} (-\Gamma_{\text{backing}})^n \exp(-2n\sigma_{\text{sample}} d)\right] \\ &= \frac{H_0 \exp(i\omega t)}{\sqrt{2\omega}(1+i)e_{\text{sample}}} \left[1 - 2\Gamma_{\text{backing}} \exp(-2\sigma_{\text{sample}} d) + \cdots\right] \end{aligned}$$

Eq. (28) can also be interpreted as a series of successive reflections of the thermal waves at the back surface with the magnitude of the reflection determined by thermal mismatch factor, Γ_{backing}. Each reflection is damped by a factor $\exp(-2\sigma d) \propto \exp(-2d/\delta_{\text{th}})$ per reflection where δ_{th} is the thermal diffusion length described in Eq. (18). For large sample thicknesses, $d > \delta_{\text{th}}$, only the first term in the expansion is significant and the expression is the same as that obtained for the semi-infinite case as shown in Eq. (22a) for $x = 0$. For smaller thicknesses, the reflections become significant and their contribution depends on the effusivity of the backing. When the backing has a lower effusivity than the sample, the thermal mismatch factor, Γ_{backing} becomes negative and the surface temperature increases substantially. Conversely, when the backing has a higher effusivity, the surface temperature increases by a smaller amount.

The magnitude and the phase angle for the surface temperature on a sample with conductive backings ($\Gamma = 0.9$ and 0.5) and reflective backings ($\Gamma = -0.9$ and -0.5) is shown in Figure 8.12 as a function of thickness in units of thermal diffusion length. Recall from Eq. (18) that the thermal diffusion length can be varied by changing the modulation frequency of the excitation source. Note how the response changes when the thermal diffusion length and the thickness of the plate are comparable (which corresponds to $d(\omega/2\alpha)^{-1/2} = 1$ in Fig. 8.12. This depth dependence is one of the reasons why thermal inspection methods can be used for nondestructive evaluation to provide quantitative measurements of thicknesses.

Copyright © Marcel Dekker, Inc. All rights reserved.

Early experiments with periodic heating have been widely used for photothermal experiments due to its high sensitivity. When a lock-in amplifier is used for detection, signals at the same frequency as a reference frequency, in this case the source modulation frequency, can be detected with high sensitivity. A disadvantage of periodic heating methods is the need to perform a frequency scan in order to obtain depth profiles. The periodic modulation has also been used in an imaging mode with infrared camera systems in recent years and has been termed "lock-in thermography" (29,30).

8.2.4 PULSED EXCITATION

Every function of time can be displayed as a sum of periodic functions of time, which is called its frequency or Fourier spectrum. The transformation between frequency and time domains is called Fourier transformation $F(\omega)$ and inverse Fourier transformation $f(t)$:

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) \exp(i\omega t) dt$$

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) \exp(-i\omega t) d\omega \tag{29}$$

This transformation allows us to use the solutions we obtained in the frequency domain for periodic modulation and calculate solutions in the time domain by integrating over all the frequency components contained in the heating source $H(t)$. An important temporal function is a very short heating pulse, mathematically expressed by a *Dirac's delta* function, $\delta(t)$. It is infinite at the origin and zero anywhere else. Integrated, it gives the value of the integrand at the origin,

$$\int_{-\infty}^{\infty} \delta(x) f(x) dx = f(0) \tag{30}$$

It is obvious from substituting Eq. (30) into Eq. (29), that $\delta(\omega) = 1/\sqrt{2\pi}$, or in words, the frequency spectrum of a delta function contains all frequencies with the same magnitude. The advantage of using a short heating pulse for excitation is that in one measurement the sample response for all frequencies can be determined. We can get the solution for Eq. (12) for a short heating pulse by integrating the solution in Eq. (24) or Eq. (28) over all frequencies. This is identical to performing an inverse Fourier transformation, shown in the second

Copyright © Marcel Dekker, Inc. All rights reserved.

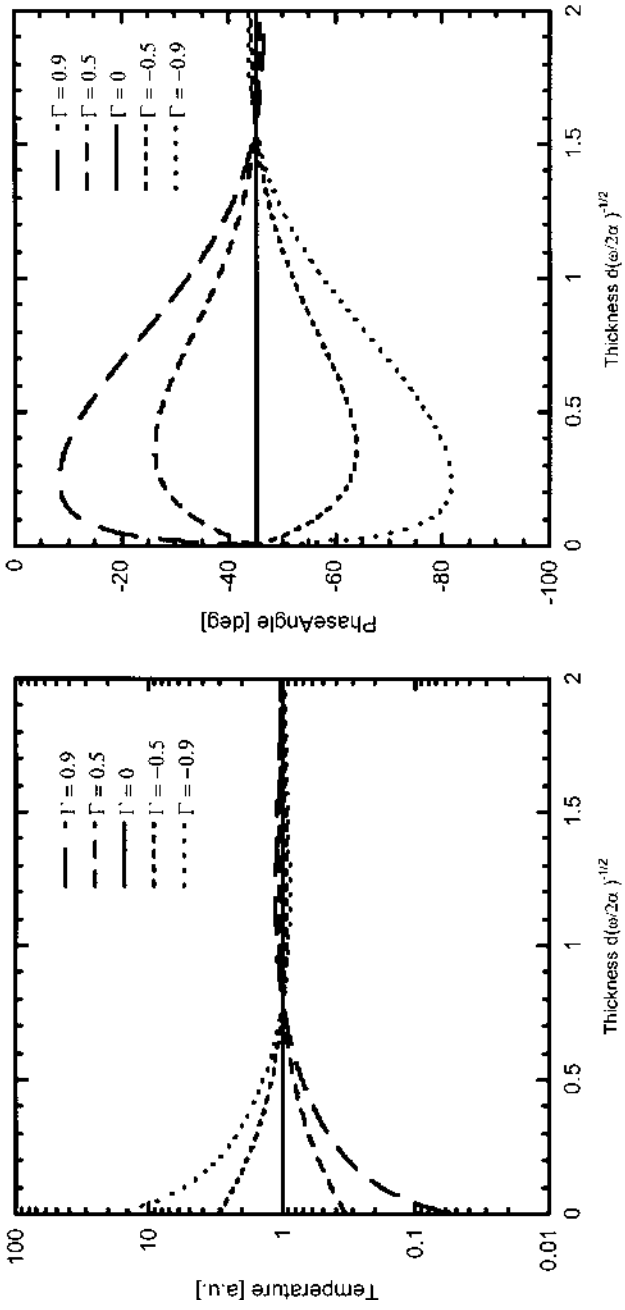


FIGURE 8.12 Magnitude and phase angle for the surface temperature on a sample with conductive ($\Gamma = 0.5, 0.9$) and reflective backings ($\Gamma = -0.5, -0.9$).

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

integral in Eq. (29). For most functions the Fourier transforms are tabulated. The pair of functions we are interested in is

$$F(\omega) = \frac{\exp\left[-(1+i)\sqrt{\frac{\omega}{2\alpha}}x\right]}{(1+i)\sqrt{\frac{\omega}{2\alpha}}x}$$

$$f(t) = \sqrt{\frac{\alpha}{\pi t}}\exp\left(-\frac{x^2}{4\alpha t}\right) \tag{31}$$

If we apply this transformation to Eq. (28), we get the surface temperature as a function of time for an ultrashort heating pulse of energy, I_0:

$$T(0,t) = \frac{I_0}{2e_{\text{sample}}\sqrt{\pi t}}\left[1 + 2\sum_{n=1}^{\infty}(-\Gamma_{\text{backing}})^n \exp\left(-\frac{n^2 d^2}{\alpha_{\text{sample}} t}\right)\right]$$

$$= \frac{I_0}{2e_{\text{sample}}\sqrt{\pi t}}\left[1 - 2\Gamma_{\text{backing}}\exp\left(-\frac{d^2}{\alpha_{\text{sample}} t}\right) + \cdots\right] \tag{32}$$

Since a very short heating pulse generates an entire spectrum of frequencies, photothermal measurements can be performed much more rapidly. The surface temperature as a function of time is shown in Figure 8.13 for conductive backings ($\Gamma = 0.9$ and 0.5) and reflective backings ($\Gamma = -0.9$ and -0.5). For short times the response is the same for both the conductive and reflective cases and shows the dependence observed for a semi-infinite specimen ($\Gamma = 0$), and which is described by the first term in the expansion in Eq. (32). The functional dependence of the response changes at a characteristic time when the heat pulse reflected from the interface returns to the surface. This is described mathematically when the second term in the expansion in Eq. (32) becomes significant. For a reflective backing the second term is additive and the temperature does not cool down as fast after the characteristic time, and for a conductive backing it cools down even faster.

This characteristic feature of the pulsed heating method has been widely used in thermographic inspection, aided by the availability of high power flash lamps that provide enough energy to investigate even challenging structures such as highly thermally-conductive aluminum aircraft skins (31,32).

8.2.5 Step Heating

Another important temporal heating function is continuous heating starting at time $t = 0$ with a constant power, P_0. Since this heating source is turned on at a known time, we refer to this as step heating. This can be considered equivalent to having a continuous array of short pulses starting a $t = 0$. Therefore a solution to

Copyright © Marcel Dekker, Inc. All rights reserved.

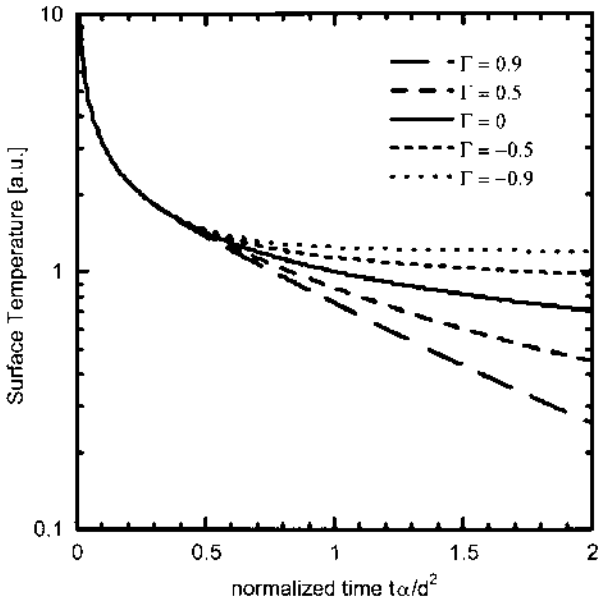


FIGURE 8.13 Surface temperature as a function of time for pulsed heating for a sample with conductive ($\Gamma = 0.5, 0.9$) and reflective backings ($\Gamma = -0.5, -0.9$).

the equation of conduction of heat for continuous heating can be obtained by integrating Eq. (32) over time, giving (33)

$$
\begin{aligned}
T(0,t) &= \frac{P_0\sqrt{t}}{e_{\text{sample}}\sqrt{\pi}}\left[1 + 2\sum_{n=1}^{\infty}(-\Gamma_{\text{backing}})^n\left[\exp\left(-\frac{n^2d^2}{\alpha_{\text{sample}}t}\right)\right.\right.\\
&\quad \left.\left. - \frac{nd\sqrt{\pi}}{\sqrt{\alpha_{\text{sample}}t}}\, erfc\left(\frac{nd}{\sqrt{\alpha_{\text{sample}}t}}\right)\right]\right]\\
&= \frac{P_0\sqrt{t}}{e_{\text{sample}}\sqrt{\pi}}\left[1 - 2\Gamma_{\text{backing}}\left[\exp\left(-\frac{d^2}{\alpha_{\text{sample}}t}\right)\right.\right.\\
&\quad \left.\left. - \frac{d\sqrt{\pi}}{\sqrt{\alpha_{\text{sample}}t}}\, erfc\left(\frac{d}{\sqrt{\alpha_{\text{sample}}t}}\right)\right] + \cdots\right]
\end{aligned}
\tag{33}
$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Here *erfc*(*x*) is the complementary error function which is defined as

$$erfc(x) = \frac{2}{\pi} \int_x^\infty \exp(-t^2)dt \tag{34}$$

Note that the temperature, for times where the sum is small, increases as a function of square-root time as described by the first term in the expansion in Eq. (33). This can be seen in Figure 8.14 where the surface temperature is plotted as a function of both time and square root time for conductive backings ($\Gamma = 0.9$ and 0.5) and reflective backings ($\Gamma = -0.9$ and -0.5). For short times where heat is only diffusing in the sample layer, the temperature increases linearly with square root time until a characteristic time called the thermal transit time, where the temperature rise extends to the backing and the functional dependence changes. The functional dependence also shows the thermal mismatch factor for the backing material. For a thermally conductive backing, the temperature increase becomes smaller, whereas for a thermally insulating backing the temperature increases much faster. Note that the thermal transit time is not affected by the value of the thermal mismatch factor.

8.2.6 Other Extensions

The analytical solutions shown in the previous sections cover only a small number of problems encountered in active thermography for NDE. Due to their simplicity in the approach and analysis it is advantageous to find arrangements which can be described by these solutions. There are also situations even in the one-dimensional approach, in which the analytical description has to be extended to include, for example, optical absorption of the excitation source into the sample, infrared transparency, subsurface heating, etc.

The solutions become more complex and only partially analytical when the lateral extent of the heat source or the sample is limited, and the three-dimensional diffusion of heat has to be considered. Solutions for a line or a point heating source can be found for all the temporal patterns described earlier, and allow the determination of lateral thermal conductivity and detection of vertical. Furthermore all of the one-dimensional systems have to be described by a three dimensional model as soon as the lateral dimensions of the sample, the defect structure, or the heat source are on the order of the diffusion lengths.

8.3 TECHNIQUES

8.3.1 Introduction

The different active thermography techniques in use today can be discriminated by the type of heating source, the temporal and spatial heating source distribution,

Copyright © Marcel Dekker, Inc. All rights reserved.

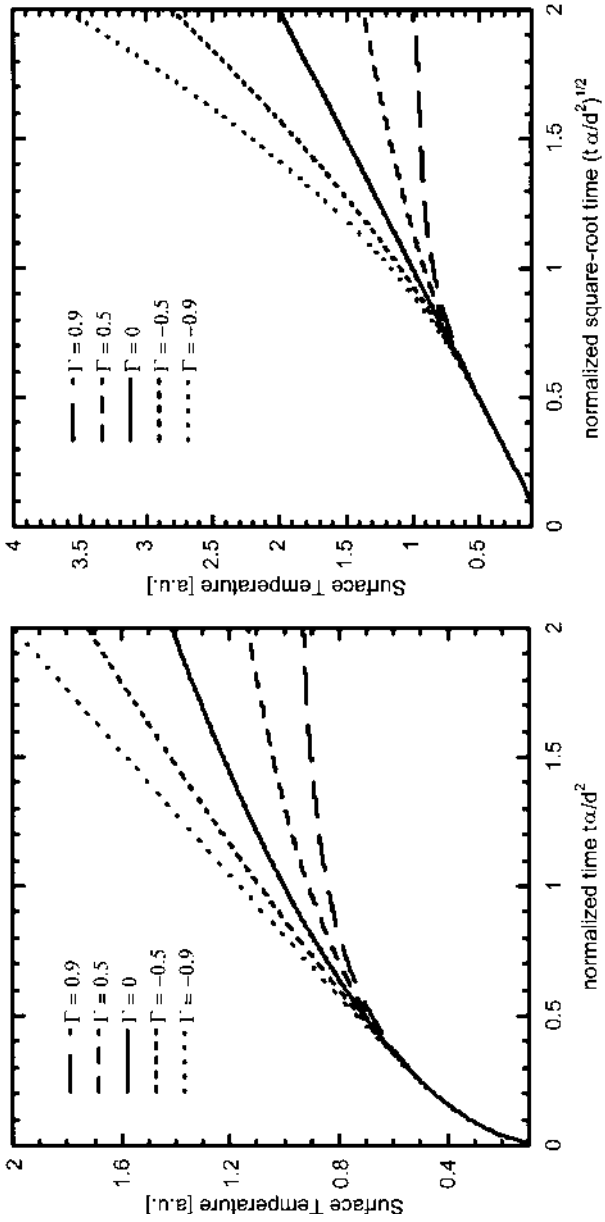


FIGURE 8.14 Surface temperature is plotted as a function of both time and square-root time on a sample with conductive ($\Gamma = 0.5, 0.9$) and reflective backing ($\Gamma = -0.5, -0.9$) for step heating excitation.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

and the temperature detection method. The common theme of all the techniques is that in order to study the transport of heat and obtain quantitative information about a structure, the structure has to be heated and the resulting temperature distribution on the structure has to be measured. Several devices and methods for temperature detection have been developed in the past (see Section 8.1.2) such as photoacoustic cells and optical beam deflection. While these earlier approaches perform extremely well in some special applications, they do not meet all the current requirements for a viable NDE method, which include the ability to survey large samples quickly and without contact to the specimen using operator-friendly instrumentation. The temperature detection method currently used for most active thermography applications is infrared radiometry in which the infrared emission of the sample surface is imaged as a function of position. With the development of fast infrared cameras with arrays of up to 512×512 infrared detectors, temperature distributions can be detected with high temporal resolution (speeds up to 1 kHz) and high sensitivity (better than 20 mK noise-equivalent temperature). A typical experimental setup for active thermography is shown in Figure 8.15. It is relatively simple, consisting of a switched heating source and an infrared detection method. In the following section, we will focus on infrared radiometry as a detection method, and discuss the important heating sources and, temporal and spatial patterns used in active thermography techniques.

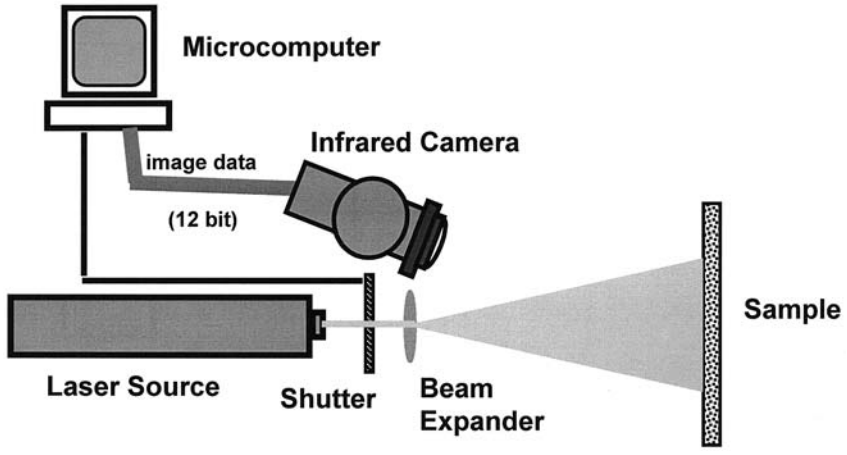


FIGURE 8.15 A typical experimental setup for active thermography.

Copyright © Marcel Dekker, Inc. All rights reserved.

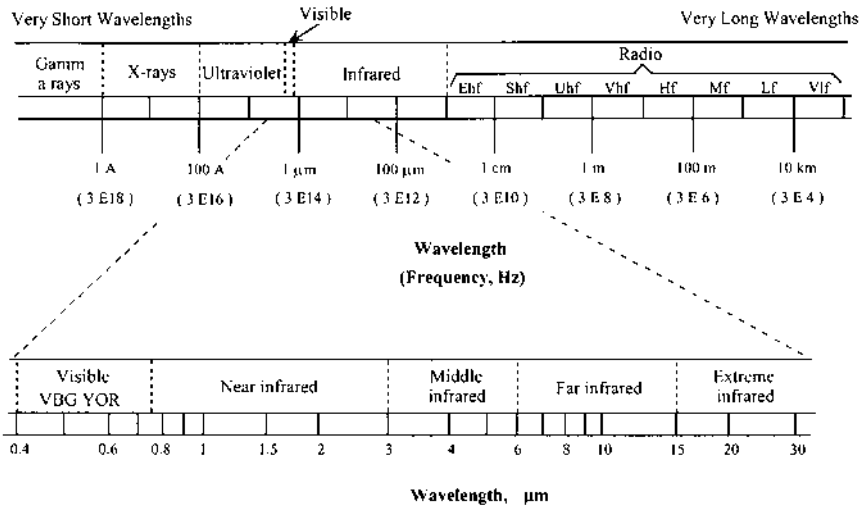


FIGURE 8.16 A schematic of the electromagnetic spectrum with radio waves at the long wavelength and X-rays at the short wavelength limits.

8.3.2 Infrared Radiometry

Radiation Laws

In infrared radiometry, the surface temperature distribution is determined by measuring the infrared emission of the sample. Figure 8.16 shows the electromagnetic spectrum with gamma rays at short wavelengths and radio waves at long wavelengths. The energy of the radiation is inversely proportional to its wavelength. Infrared radiation is electromagnetic radiation in the middle of the spectrum with wavelengths from 0.7–20 μm. According to Planck's law, the *spectral radiant emittance*, W_λ (the flux radiated into a hemisphere per wavelength interval $d\lambda$) of a body at an absolute temperature T is given by

$$W_\lambda = \frac{c_1}{\lambda^5} \frac{1}{\exp(c_2/\lambda T) - 1} \tag{35}$$

where

$$c_1 = 3.742 \times 10^4 \text{ W cm}^{-2} \mu\text{m}^4$$
$$c_2 = 1.439 \times 10^4 \ \mu\text{m K} \tag{35}$$

The spectral radiant emittance is plotted in Figure 8.17 as a function of wavelength in both linear–linear and log–log presentations. The log–log plot

Copyright © Marcel Dekker, Inc. All rights reserved.

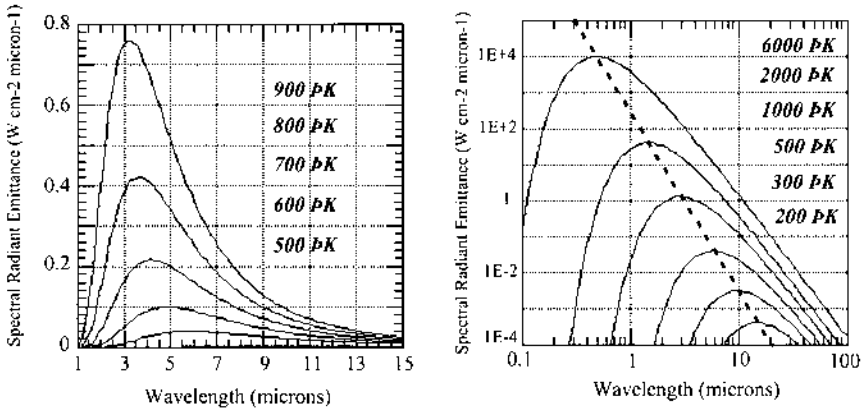


FIGURE 8.17 Spectral radiant emission as a function of the wavelength. (a) linear–linear plot, (b) log–log plot.

shows how the position of the peak maximum moves to shorter wavelengths with increasing temperature. This behavior is described by Wien's law,

$$\lambda_m T = a \tag{36}$$

where λ_m is the wavelength of maximum spectral radiant emittance and a is a constant given by 2897.8 μm K. At room temperature, 300 K, the maximum emittance occurs at a wavelength of 10 μm.

The spectral radiant emittance as described in Eq. (35) is only valid for a so-called "black body," which emits as much energy in the form of radiation as it absorbs from its environment at the same temperature. The ratio of the radiant emittance of any material, W', to the radiant emittance of a black body at the same temperature, W_{BB}, is called *emissivity* ε.

$$\varepsilon = \frac{W'(\lambda, T)}{W_{\mathrm{BB}}(\lambda, T)} \tag{37}$$

A material with $\varepsilon < 1$ is called a gray body. Note that for most materials, the emissivity is also wavelength-dependent. The emissivities of different materials, integrated over the all wavelengths, are shown in Table 8.4.

The total radiant emittance, i.e. the total flux radiated into a hemisphere above the body at temperature, T is given by Stefan-Boltzmann's law:

$$W = \int \varepsilon W_\lambda d\lambda = \varepsilon \sigma T^4 \tag{38}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

TABLE 8.4 Emissivities of Different Materials, Integrated Over All Wavelengths.

Material	Emissivity	Temperature°C
Graphite	0.98	20
Skin, human	0.98	32
Lacquer, matte black	0.97	100
Water, distilled	0.96	20
Paint, oil based	0.94	100
Glass, polished plate	0.94	20
Paper, white bond	0.93	20
Brick, red common	0.93	20
Concrete	0.92	20
Sand	0.90	20
Wood, planed oak	0.90	20
Stainless steel, 18-8:		
oxidized at 800°C	0.09	60
buffed	0.05	20
Cast iron:		
oxidized	0.64	100
polished	0.21	40
Aluminum:		
sheet as received	0.09	100
polished sheet	0.05	100
Silver, polished	0.03	100
Gold, polished	0.02	100

Source: Ref. 34.

where σ is the Stefan-Boltzmann constant, $\sigma = 5.67 \times 10^{-12}\,\mathrm{W\,cm^{-2}\,K^{-4}}$. The rapid increase of the total radiant emittance, the area under each curve, with increasing temperature can be seen in Figure 8.17.

Most infrared detectors have a limited wavelength range, D_λ over which they are sensitive, and the sensitivity is not necessarily uniform over this range. The important quantity for infrared radiometry is the change of infrared emissivity of the sample with temperature, $\varepsilon_\lambda\,(dW_\lambda/dT)$ integrated over the detector's spectral range, D_λ,

$$S = \int \varepsilon_\lambda \frac{dW_\lambda}{dT} D_\lambda d\lambda \tag{39}$$

The function dW_λ/dT as a function of wavelength is shown in Figure 8.18. The *detector spectral sensitivity*, D_λ, depends on the detector used and Eq. (39) can be

Copyright © Marcel Dekker, Inc. All rights reserved.

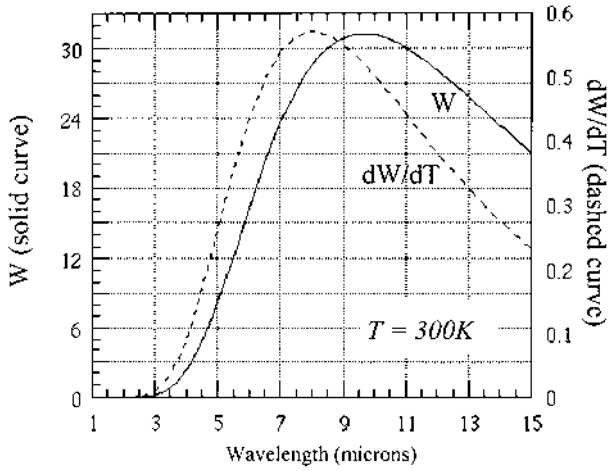


FIGURE 8.18 The function dW_λ/dT as a function of wavelength.

used to compare the sensitivities at different wavelengths for different cameras. For InSb with sensitivity between 3 μm and 5 μm, the emission of a blackbody at 290 K is 4 W/m^2, and the contrast of a 291 K object in a 290 K background is 0.039. For the range of 8 μm to 12 μm, typical for HgCdTe detectors, the emission of the black body at 290 K is 127 W/m^2, while the contrast is only 0.017. Therefore, the longer wavelength band has a better sensitivity for objects at ambient temperatures, whereas the shorter band performs better when looking for small temperature rises. The optimal choice of a detector depends further on the detector characteristics themselves, as discussed in the next section.

Infrared Detectors

There are many classes of detectors available for sensing infrared radiation. Most semiconductor-based detector systems, also called photon detectors, must be cooled to liquid nitrogen temperature (77 K) in order to be sensitive enough to detect radiation with wavelengths longer than 1 μm. Common infrared materials are indium antimonide (InSb) and platinum silicide (PtSi) for the range 3–5 μm, and mercury cadmium telluride (HgCdTe) for the range between 8–12 μm. These materials are available as single cell detectors and as scanners or focalplane arrays in infrared cameras. The use of thermal detectors is another method for detecting IR radiation where the temperature rise in a small absorber is measured. This approach has given rise in recent years to the construction of uncooled detectors and cameras.

The important characteristics of an IR detector are its *responsivity*, R, the ratio between the signal voltage and the incident flux level, and the *noise-*

Copyright © Marcel Dekker, Inc. All rights reserved.

equivalent power (NEP), the incident flux level required to produce a signal equal to the RMS noise. The NEP can be transformed into a *noise-equivalent temperature* (NET), which is the temperature rise at the sample required to produce a signal equal to the RMS noise. Typical values for the NET are in the order of up to a few mK. Another important characteristic is the speed of the detector, i.e. the fastest temperature rise the detector can follow. Photon detectors are typically faster than thermal detectors.

In most arrangements the IR radiation has to be focused optically onto the detector. This requires optical components such as lenses, which are transparent in the infrared range. Some typical materials and their transmission bands (35) are silicon (1.2–15 μm), germanium (1.2–23 μm), zinc selenide (0.5–20 μm), barium fluoride (0.25–15 μm), and sapphire (0.14–6.5 μm). Note that some of these materials are also transparent for visible light.

Infrared Cameras

In order to generate an image of the temperature distribution, the detector and the sample have to be moved with respect to each other. In earlier experiments, this was done by scanning the sample under a fixed detector, an approach only applicable to smaller samples which can be moved.

The first infrared cameras, called scanners, used galvanometer-based mirrors to scan the image onto a single detector or a linear detector array. A schematic of such a scanner is shown in Figure 8.19. The detection speed of this arrangement is determined by the time it takes for the mirror to scan through an

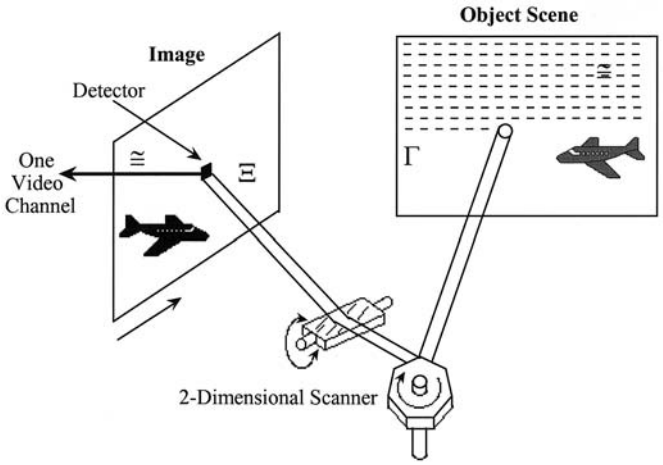


FIGURE 8.19 Infrared scanner schematic.

Copyright © Marcel Dekker, Inc. All rights reserved.

image which can be anywhere from a few seconds per image for a single detector up to video rates (60 images per second) when using a linear detector array.

In the last ten years, two-dimensional arrays of infrared detectors have been fabricated and combined with *charge-coupled device* (CCD) electronics for read out. These arrays are called focalplane arrays and they allow infrared cameras to be used like normal video cameras. High frame rates are achievable with these systems. Typical array sizes range from 128×128 up to 640×480 pixels and materials used include InSb, PtSi, and just recently, HgCdTe. The detector arrays have to be cooled, which can be done either with a dewar of liquid nitrogen, electrically with a thermoelectric cooler, or mechanically using a miniature Stirling cooler. A schematic of an infrared focalplane array camera is shown in Figure 8.20. A typical NET for an InSb camera is in the order of 10–20 mK.

More recent developments use micromachined arrays of mechanical devices termed micro-electro-mechanical systems (MEMS) which work as a thermal detector and allow infrared imaging with an uncooled detector. While these devices are still not very sensitive and are quite slow, they are expected to provide a low-cost solution for thermographic imaging in the very near future.

Data Analysis

An important part of a thermographic measurement, as compared to a simple survey inspection where the data can be stored on videotape, is the collection of the data into a computer for subsequent analysis. For single cell detectors, data are usually collected using an analog to digital converter and stored as a time record for each sample position of the scan. In most video cameras, the CCD output from the detector array is converted into a digital signal and normalization procedures to correct for different pixel gains are performed numerically in real time. Then the images are typically converted into a video format and are available on a video output. For infrared cameras with video outputs, common video rate frame grabbers can be used. They allow collection of 30 images per second (640 by 480 pixels) with a resolution of at least eight bits. For faster

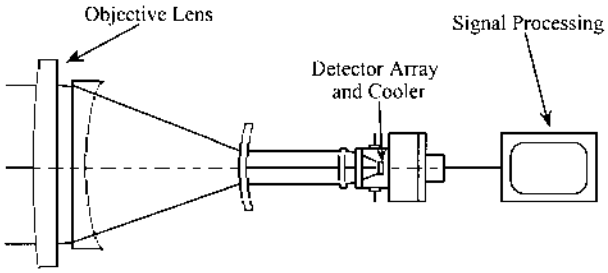


FIGURE 8.20 Focalplane array schematic.

Copyright © Marcel Dekker, Inc. All rights reserved.

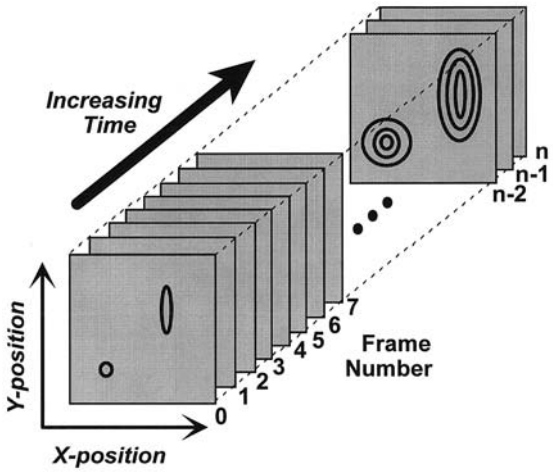


FIGURE 8.21 Schematic for data format used in active thermography.

cameras, for cameras with a different image size, or for better resolution, digital frame grabber boards are used to read the digital information out of the camera with a fast 16-bit parallel connection. In both cases the result is a block of data which is a stack of images collected as a function of time. A schematic of such a data block is shown in Figure 8.21. For most of the analysis, the important feature is the temporal dependence at certain pixel positions. Image processing algorithms can be used on the images to provide noise filtering or averaging.

8.3.3 Heating Sources

Heating Methods

The heating source is the most variable element in active thermography. The optimal choice of heating source depends on the sample material, the sample geometry, the defect structure, and the chosen technique with its specific temporal requirements. A number of possible heating sources are shown in Table 8.5 along with their wavelength range, potential modulation rates and whether periodic modulation is possible. Note that cooling, although not mentioned in the table, also is an option. All the heating sources described, except the hot air blower and the resistive heating, use electromagnetic radiation with the wavelength shown in the first column. Note that if the source emits radiation in the infrared range of the camera's sensitivity, the detector will be blinded while the source is on. A separation of the source and detector wavelength is therefore desirable in most cases. The second column shows the rate at which the source can be modulated. Depending on the technique and the properties of the material system under

Copyright © Marcel Dekker, Inc. All rights reserved.

TABLE 8.5 Methods for Specimen Heating

Method	Wavelength	Onset Time	Periodic Modulation Possible?
Laser	0.5–10 μm	psec–sec	✓
Flashlamps	IR	< msec	×
Quartz lams	IR	> sec	✓
Microwaves	1 mm–10 cm	> nsec	✓
Induction heating	10 cm–1 m	> msec	✓
Resistive heating	N/A	> msec	✓
Hot air blower	N/A	≫sec	×

study, it can be important to have good control over the onset time and the duration time of the heating source to allow quantitative interpretation of the thermographic results. The third column indicates whether the heating source can be periodically modulated, which is done in some techniques described in the next section.

The temporal requirements for the heating source are determined by the material properties. The temporal resolution required to measure the depth of a defect, L, in a material with thermal diffusivity, α, must be faster than the *thermal transit time*, $\tau = L^2/\alpha$. For 1 mm aluminum (high thermal diffusivity) the thermal transit time is about 10 ms but for 1 mm polypropylene (much lower thermal diffusivity) the thermal transit time is about 10 sec. For a measurement on a thin aluminum sample, only flashlamps or a laser can be switched fast enough. For the polypropylene sample, however, a simple quartz lamp could be used.

The optical absorption properties of the material under study are another important factor when choosing a heating source. Ideally, in order to use the simple models from Section 8.2, the radiation from the heating source is absorbed at the surface of the sample. This requires the wavelength of the source to be in an absorption band of the sample. Sometimes this can be achieved by putting a thin absorbing paint layer on the surface of the sample. Metals can also be heated effectively by induction heating with RF radiation, or just by resistive heating.

There are cases where subsurface heating is beneficial, although the mathematical description of this case is more complicated. Some kinds of embedded defects, e.g. entrapped water, can be heated very efficiently using microwave radiation in a water absorption band. An example of such a measurement will be shown later.

Spatial Heating Patterns

The spatial distribution of the heating source becomes important for detection of particular defect geometries. The equations developed in Section 8.2 are only

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

valid when the heat diffusion can be considered one dimensional, which is the case when the homogeneous extent of the heating source is much bigger than the depth of interest in the sample. In this case, diffusion need only be considered in the direction perpendicular to the sample surface. When either the sample or the heating source has finite lateral extent, lateral diffusion of heat must be considered and a three-dimensional model is required. While this limits the applicability of the simple models, it allows measurement of other interesting quantities such as the in-plane thermal diffusivity. This is shown in Figure 8.22, where the temperature images for a line heating source are shown at 0.1 sec and 2 sec for Delrin, a homogeneous sample, and for a ceramic matrix composite with higher in-plane conductivity.

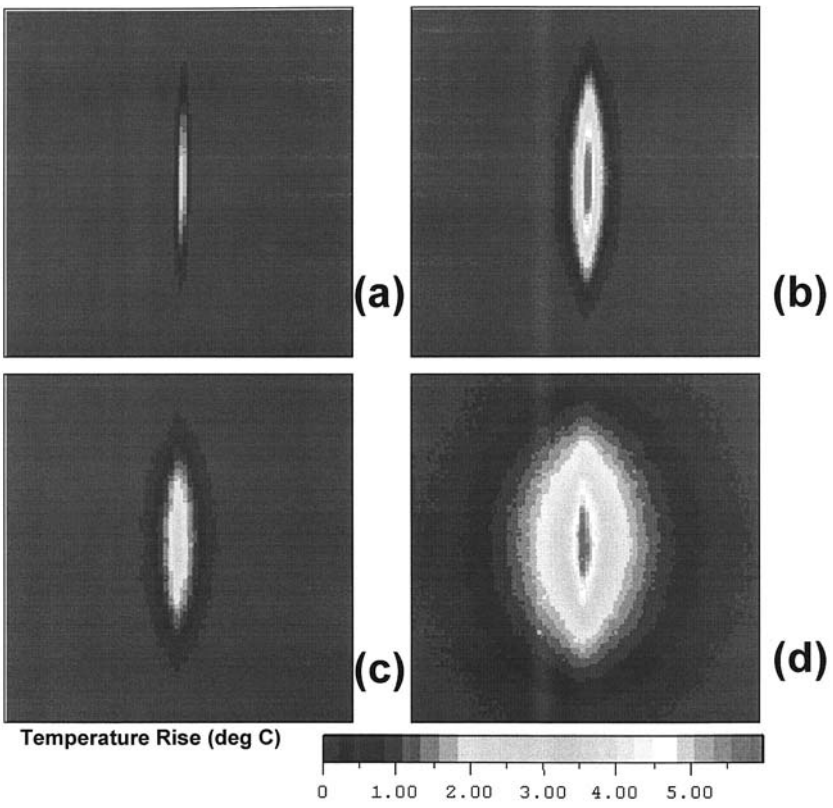


FIGURE 8.22 Thermal images for a line source heating are shown after 0.1 and 2 seconds of heating for Delrin (a), (b), and a ceramic matrix composite (c), (d) with higher in-plane conductivity.

Copyright © Marcel Dekker, Inc. All rights reserved.

8.3.4 Making Active Thermography Measurements

When presented with an NDE problem to which one would like to apply a thermal method, there are a number of factors to consider in setting up an approach. Since the contrast to be expected in a thermal image depends so strongly on the relative thermal properties of the defect and its surroundings, an initial understanding of these approximate values will provide insight into whether or not a thermal approach will be of value. It is important to recognize that there is a whole suite of NDE methods available, and not every method is equally successful on a specific problem.

It is also important to recognize that the ratio of a defect's depth to its lateral extent determines whether a thermal NDE method will be of use. As a general rule of thumb, it is not easy to detect any defect which is deeper in the specimen than it is wide. This is due to the fact that the diffusive nature of heat flow means that in a homogeneous material, the heat flows laterally as much as it flows down. It is interesting to note that step heating actually has less problem with smearing of the thermal images of defects than does pulsed heating.

The step heating approach requires a heating source that differs in the spectral content from the spectral range of the camera, which usually excludes all thermal sources. It also requires exact timing between the start of the heating source with the frame signal of the camera. For a typical experiment, the heat source can be a laser which is switched with a fast shutter, an acousto-optic modulator, or by direct control of the laser as is possible in the case of a laser diode. Two advantages of the step heating approach are the lower surface temperature rise required and the reduced lateral diffusion obtained. Disadvantages of step heating are the smaller contrast for interfaces and the requirement of wavelength separation between the heating source and the detector. This wavelength separation requirement is addressed by the use of a laser source which can provide enough power to homogeneously heat an area of typically about a 1 ft^2. Laser-diode arrays are now viable for this since their power levels have increased in the last few years. The requirement for wavelength separation is not a problem for nonoptical heating sources such as microwaves and inductive heating which will be described later.

Determining the speed of the measurement is another key requirement. For example, detection of an interface through 1 cm of aluminum will occur much more rapidly than through 1 cm of a ceramic and the measurement times must be adjusted accordingly. This is done by selecting the number of frames to be captured and the time delay between frames of the infrared camera. There is a limit to how thin a material one can inspect which is determined by the maximum frame rate of the camera to be used. With a conventional camera with video frame rates (60 Hz), the time between frames is about 17 ms.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Selection of using pulsed or step heating is partially determined by the required speed of the measurement. Step heating approaches are best suited to specimens with slower thermal responses while the pulsed method works well with specimens with faster thermal responses. Note that a step heating method also requires a heat source that is spectrally separate from the detection range of the infrared camera. A laser heating source is a good choice or some nonoptical heating approaches such as microwaves.

Selection of the heating method is another key element to be considered when contemplating using active thermography for a particular NDE problem. A particularly vivid example of this is when the defect can be selectively heated to help in providing contrast in the thermal image. For example, the use of microwaves to heat entrapped water under coatings or within composite materials is very effective.

Finally, the optical properties of the specimen surface need to be understood. This is important for both the optical absorption of a heating source and the infrared emissivity which determines how much infrared emission will be detected by the camera. Furthermore, if there is any spatial variation in these properties, the image will show these nonuniformities and compensation must be made.

Once the basic elements of step vs. pulse heating, the time duration for the measurement, and the choice of heating method have been determined, the test setup needs to be assembled and the experimental parameters need to be selected. These include:

Energy density of heating method on sample surface
Specimen surface emissivity and whether modifications are needed such as applying a washable high emissivity paint
Area to be inspected in one field of view and required spatial resolution for defect detection

The graphite-epoxy specimen with subsurface holes described at the beginning of this chapter for which a sample thermal image was shown in Fig. 8.1 will now be examined with both step and pulse heating to demonstrate the different types of thermal images that are obtained. Figure 8.23a shows images captured during step heating of this specimen after 1, 2, 4, 8 and 16 seconds. Plots of the temperature rise versus position across the holes are shown in Fig. 8.23b for each of these times. Note that for step heating, the contrast over the defects increases with heating time, although the total temperature excursion does not exceed 5000 mK (5°C). Corresponding images and plots for pulsed heating of this sample are shown in Figures 24a and b, also at times of 1, 2, 4, 8 and 16 seconds. Note here that the contrast of the defects is greatest at about 4 seconds. At 1 second the holes are not yet detected thermally, but by 16 seconds, the remaining temperature rise is too small to show much contrast at the defects.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

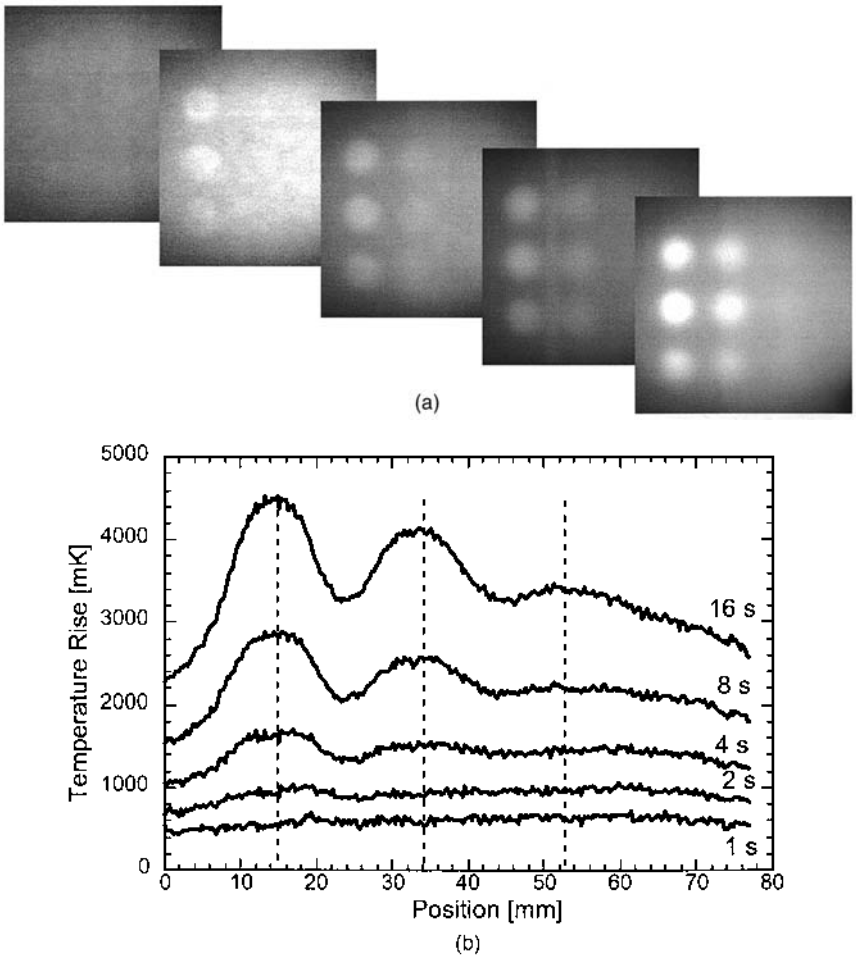


FIGURE 8.23 (a) A series of thermal images after 1,2,4,8 and 16 seconds of step heating for a graphite-epoxy specimen with flat bottomed holes. (b) Plots of temperature rise across the flat bottomed holes after 1,2,4,8 and 16 seconds of step heating.

These measurements were done with the same laser intensity, so in the pulsed case where the heating time is much shorter, the total temperature rise is smaller than in the step heating case.

For step heating cases, data analysis algorithms based on Eq. (33) can be used to generate images of subsurface defect depths in samples with known backings, or to characterize the backing materials for samples with known top layer thickness and thermal diffusivity. The time dependence of the temperature

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

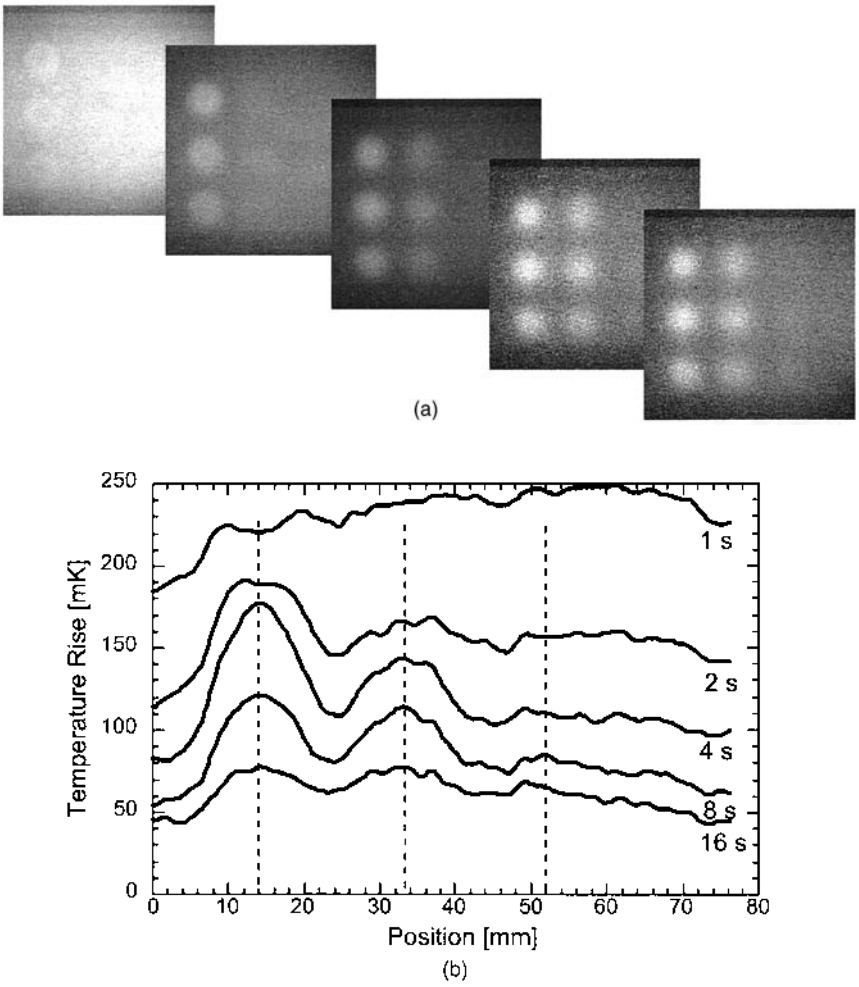


FIGURE 8.24 (a) A series of thermal images at 1,2,4,8 and 16 seconds after a heating pulse for a graphite-epoxy specimen with flat bottomed holes. (b) Plots of temperature rise across the flat bottomed holes at 1,2,4,8 and 16 seconds after a heating pulse.

with the semi-infinite response subtracted (this will be referred to as the normalized temperature) at a position at the center of the hole is plotted in Figure 8.25a for three different holes as a function of square root time. This subtraction leaves only the summation term in Eq. (33) which has a value of 0.18 (the horizontal line drawn on the graph) for $(t\alpha)^{1/2}/L = 1$. The depth of the defect can then be determined if the thermal diffusivity of the material is known. This

Copyright © Marcel Dekker, Inc. All rights reserved.

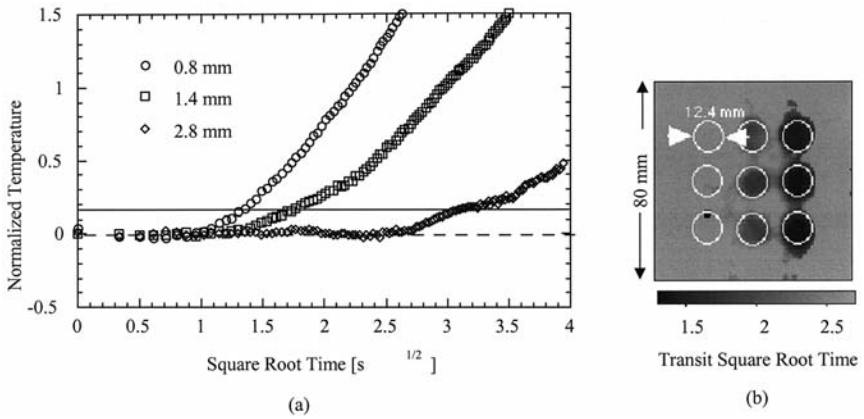


FIGURE 8.25 (a) Time dependence of the normalized temperature at a position at the center of the hole is plotted in Figure 8.18 (a) for three different holes. (b) A transit time image for the graphite-epoxy composite panel with flat-bottomed holes.

analysis can be done for each pixel thus creating an image of the thermal transit times. This was done in Figure 8.25b, which shows a transit time image for the graphite-epoxy composite panel with flat-bottomed holes. These images summarize the information contained in the stack of infrared images collected as a function of time during heating.

When two different materials are bonded together, the thermal transport will depend on the thermal transit time for the top layer and also on the thermal mismatch factor between the top layer and the backing material. For a homogeneous upper layer of constant thickness, the transit time is the same everywhere, and the normalized temperature images can be used to determine the thermal mismatch factor, which is characteristic of the backing material. Thermography measurements were performed on a test sample consisting of small plates of different materials bonded onto a 1 mm thick layer of fiberglass-epoxy composite. Figure 8.26a shows the normalized temperature image after a 2 sec heating pulse. The different plates are easily detected and the different materials show different and distinct behaviors.

The range of behaviors is shown more clearly in Figure 8.26b where the normalized temperature is plotted as a function of square root time for pixel locations on each of the different materials. As expected, the sign of the thermal mismatch factors for the thermally-conductive metals (steel, copper, brass, aluminum) and the thermally-insulating materials (Plexiglas, fiberglass) are different. The thermal transit time to the interface for all locations on the

Copyright © Marcel Dekker, Inc. All rights reserved.

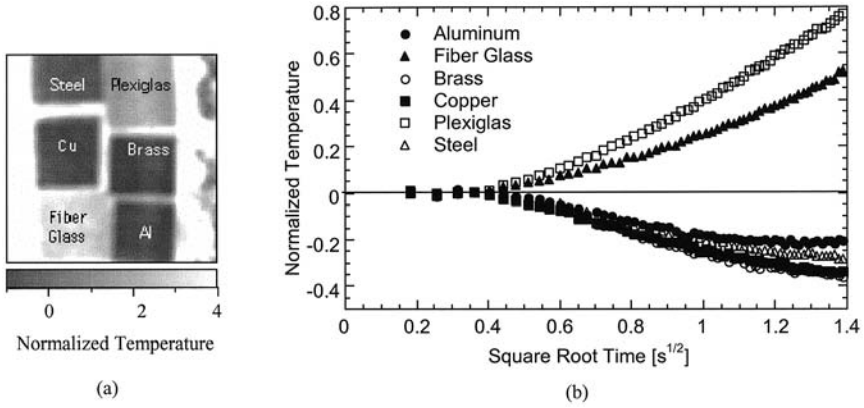


FIGURE 8.26 (a) Normalized temperature image of a test sample consisting of small plates of different materials bonded onto a 1 mm thick layer of fiberglass-epoxy composite. (b) Normalized temperature is plotted as a function of square root time for pixel locations on each of the different materials.

sample is the same since the fiberglass upper layer thickness was constant at 1 mm. A full thermal analysis at the location of the metal backing layers is more complicated, since the metal layer thickness on the sample is too thin for a two-layer model to be valid.

8.4 SPECIFIC APPLICATIONS

In this section a number of specific examples of the application of active thermography to actual NDE problems will be described. The examples encompass different heating methods including microwave and induction sources in order to demonstrate the important role of the heating source in active thermography.

8.4.1 Imaging Entrapped Water Under an Epoxy Coating

Microwave heating methods have provided some unique capabilities compared to heating with conventional optical sources (36,37). The use of microwave heating sources has distinct advantages for optically opaque but microwave-transparent materials containing localized absorbing regions, such as entrapped water in composites. For particular specimen geometries and material properties, the presence of the defect region can be imaged at higher contrast and better spatial resolution than obtainable with the surface heating technique. Since the heat has

Copyright © Marcel Dekker, Inc. All rights reserved.

only to diffuse to the surface, the characteristic thermal transit times for the measurement are shorter. Further, the spatial resolution in these measurements is determined by the IR wavelength and not by the microwave wavelength as occurs in conventional microwave imaging techniques. As a result, image resolutions of better than 30 μm can be obtained.

The measurements described here use microwaves at a frequency of 9 GHz and a maximum power of 2.3 W fed into a single-flare horn antenna through rectangular waveguide. The antenna has a beamwidth of about 50° and is placed about 15 cm from the sample. A 128 × 128 InSb focalplane array (Santa Barbara Focalplane ImagIR, 12 bit A/D converter) operating in the 3–5 μm band is used for detection of the IR radiation. The benefits of microwave heating in specific applications is demonstrated by the four images in Figure 8.27. The specimen is a section of steel pipe with an epoxy coating that has experienced some disbonding. This is a coating system widely used for corrosion protection of buried gas pipelines. Conventional laser source thermography is shown in Figure 8.27a and b for a disbond which was first dry and then filled with water. The disbonded region is much more clearly delineated when dry due to the high thermal contrast between the epoxy coating and the underlying air. When the disbond is filled with

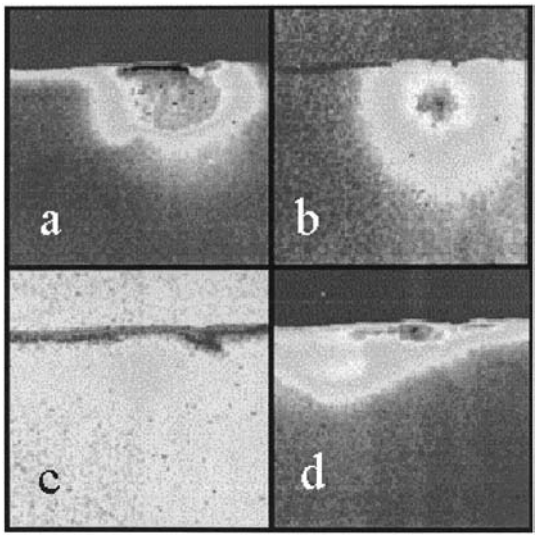


FIGURE 8.27 Thermographic images of a partially-disbonded epoxy-coated steel sample after 15 sec of laser heating for (a) the empty void and (b) the void filled with water, and after 30 sec of microwave heating for (c) the empty void and (d) the void filled with water.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

water, as is often encountered in a field situation, the thermographic image is predominately determined by the spatial distribution of the laser heating source, and does not clearly show the disbonded region. The thermographic image in Figure 8.27c shows microwave heating of the dry disbond and there is not any appreciable heat deposition in the specimen since the epoxy coating is microwave transparent. The image in Figure 8.27d was taken after the disbond region was filled with water. Here the water is readily heated by the microwaves and the infrared image of the surface of the coating provides an outline of the disbond region.

8.4.2 Detection of Carbon Fiber Contaminants

Another application for microwave heating is the detection of carbon fiber contaminants in fiberglass-epoxy composite materials (38). Such small linear conducting defects can occur in a composite fabrication environment. The quality of the composite structures depends on the purity of the different layers and is an important factor for applications in the aerospace industry. Thermographic detection with microwave excitation is a method which allows detection of the fiber and, using the analytical description of the surface temperature, determination of the depth of the fiber and the thermal diffusivity of the embedding material simultaneously and independently.

This approach can be implemented as an embedded sensor to monitor processes during which the thermal diffusivity of the embedding matrix material varies over time, as for example when looking at the curing of epoxy resin. Figure 8.28a shows infrared images of a 1 cm long carbon fiber embedded in uncured epoxy after 4 s of microwave heating. The time dependence of the temperature at different positions across the fiber is shown in Figure 8.28b. Fitting these curves using a model for a buried line heating source results in a thermal diffusivity of $0.84 \times 10^{-3}\,\mathrm{cm}^2/\mathrm{s}$ for the uncured and $1.48 \times 10^{-3}\,\mathrm{cm}^2/\mathrm{s}$ for the cured epoxy (39). Since the vertical position of the fibers changed due to the curing process, the independent determination of depth and thermal diffusivity was necessary.

8.4.3 USING INDUCTION HEATING FOR NDE OF REBAR IN CONCRETE

Another application with a nonoptical heating source is the use of RF induction heating for embedded metal structures. Figure 8.29a shows an infrared image (Amber Radiance, 256×256 pixels, 12 bit A/D converter) obtained of a concrete test specimen with rebar and an embedded defect imitating corrosion product after 1 min of heating with the induction heating source. (The coil wrapped around the cylindrical specimen is the induction heating coil.) The rebar appears bright in this infrared image since it is being heated about 10°C above ambient by the induction source. At 1 min of heating, there is no apparent variation in surface

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

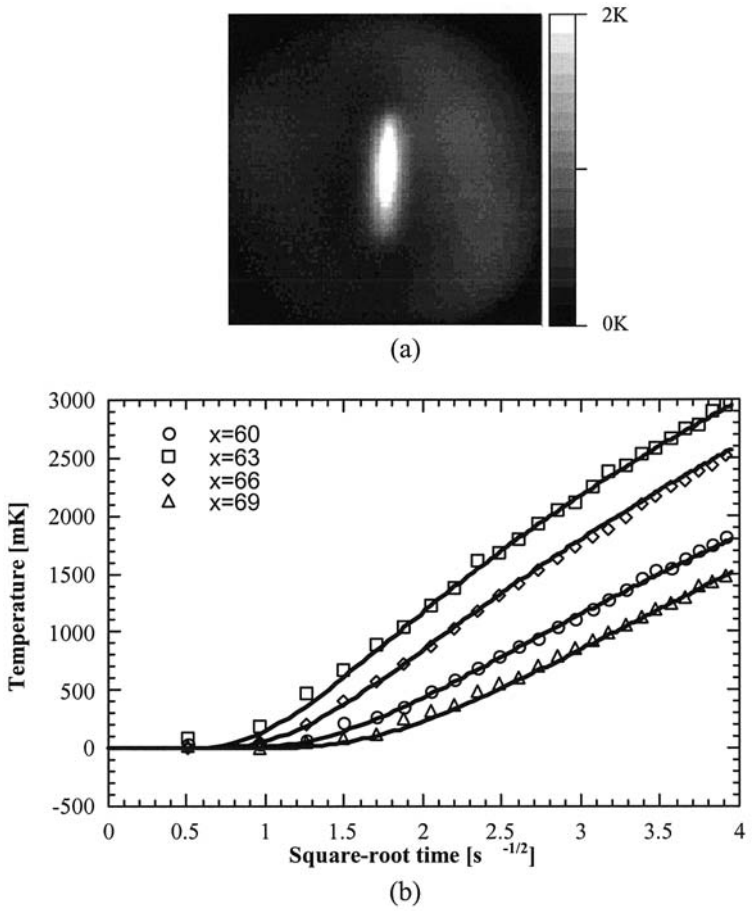


FIGURE 8.28 (a) Thermographic image of a 1 cm long carbon fiber embedded in uncured epoxy after 4 s of microwave heating. (b) Time dependence of the temperature rise at different *x*-pixel positions across the fiber in Fig. 8.10a for the uncured epoxy. Each pixel represents 0.2 mm in distance. The solid lines are calculated using a thermal diffusivity of $0.84 \times 10^{-3}\,\text{cm}^2/\text{sec}$.

temperature at the surface of the concrete. This is shown more clearly in the accompanying plot, Figure 8.29b, which displays the temperature as a function of position along the vertical rebar position in the infrared image. Also shown for comparison is the temperature distribution along a section of bare rebar without concrete after 1 min of heating. After 10 min of heating, however, the presence of the foam defect can be detected through the dip in the temperature plot from about 4 in. (101.6 mm) to 6 in. (152.4 mm).

Copyright © Marcel Dekker, Inc. All rights reserved.

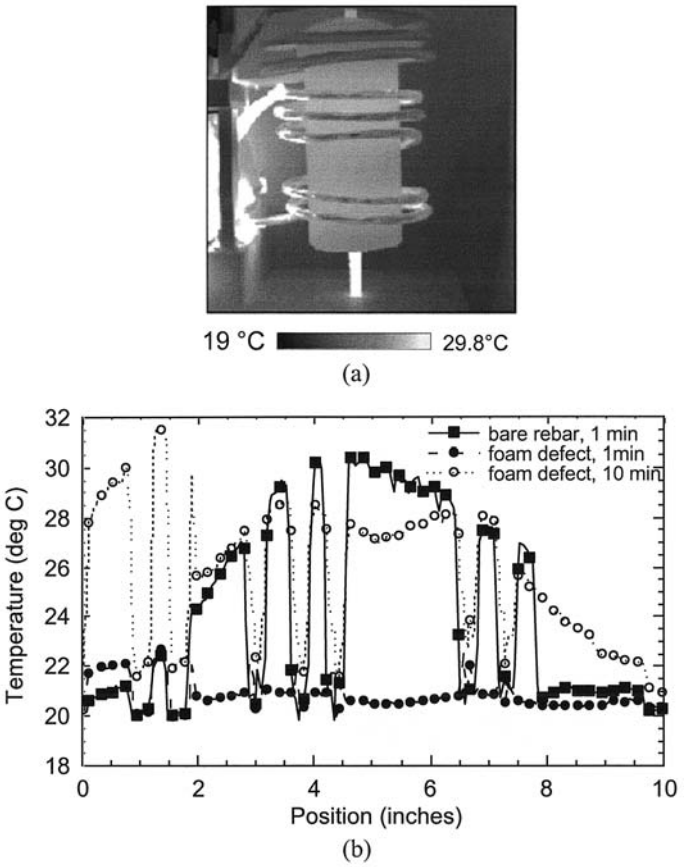


FIGURE 8.29 (a) Thermographic image of a concrete test specimen with rebar and an embedded defect imitating corrosion product after 1 minute of heating with an induction heating source. (b) Temperature as a function of position along the rebar in Fig. 12 (a) for a bare rebar sample after 1 min. of heating and for a concrete specimen containing rebar with a foam defect after 1 min and 10 min of heating.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

PROBLEMS

1. Calculate the time it takes to heat up a volume, V, of 1000 cm^3 of water, aluminum, and air from 10–70°C with a heating power, P, of 100 W (assuming no losses).
2. If you decrease the temperature of a liter of water by 1°C, what volume of air can you heat up by 1°C with the energy difference?
3. (a) How much power is lost through a 10 cm thick concrete wall of area 100 m^2, when inside the temperature is 20°C and outside it is 0°C? (Thermal conductivity of concrete 0.009 W/cm K).
   (b) How much energy is this per day?
4. How does this change when a 1 cm thick layer of Styrofoam ($\kappa = 0.0002$ W/cm K) is put on the outside? (Note that the temperature at the Styrofoam-concrete interface is continuous.)
5. (a) The concept of the thermal penetration length can also be applied to nature. The temperature changes during a day and a year can be considered as "thermal waves".
   (b) What are the depths at which the temperature changes penetrate into the soil (thermal conductivity 0.01 W/cm K, specific heat 0.8 J/g K, and density 2.5 g/cm^3) during the periods of a day and a year?
6. Equally absorbing black surfaces of aluminum and zirconia are illuminated with peridically modulated light at 100 Hz and a power density of 10 W/cm^2. What is the magnitude of the temperature variation at the surface, and depths of 100 μm and 1 mm?
7. Equally absorbing black surfaces of aluminum and zirconia are now illuminated with a very short light pulse of an energy density of 100 mJ/cm^2. What are the temperatures on the surface at times of 10 ms, 100 ms, and 1 s?
8. (a) How do the numbers for 100 ms calculated in Problem 8 change, if the sample thickness is 0.1 mm, 1 mm, and 10 mm? Consider only the first term in the sum.
   (b) Calculate the error by comparing to the size of the second term.
9. Equally absorbing black surfaces of aluminum and zirconia are now illuminated with a lightsource of a power density of 1 W/cm^2. What is the temperature at the surface at times of 10 ms, 100 ms, and 1 s after the light is turned on?
10. How do the numbers for 100 ms calculated in Problem 9 change if the sample thickness is 0.1 mm, 1 mm, and 10 mm? Consider only the first term in the sum. Use the closest number from the following table for the error function:

x	0	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	1	2	3
Erfc(x)	1	0.94	0.89	0.78	0.67	0.57	0.48	0.40	0.32	0.26	0.18	0.05	0

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

11. (a) What is the maximum time a bulk sample of wood ($\alpha = 0.0024\,cm^2/s$, $\kappa = 0.00013\,W/cmK$) can be heated with a power density of $1\,W/cm^2$ without the surface temperature exceeding 100°C
    (b) How much energy is in a 1 ms pulse that will achieve the same surface temperature after 1 ms?
    (c) What is the temperature after the time calculated for the step heating?
12. (a) How does the total emittance (integrated over all wavelengths) of a body change with temperature?
    (b) What is the wavelength with the maximum emittance at room temperature?
    (c) How does the emittance at this wavelength change for a temperature increase of 1 K?

VARIABLES

e Thermal effusivity($J/m^2\,K\,s^{1/2}$)
I Pulse energy density
J Heat flow (W/m^2)
L Layer thickness (m)
P Intensity (W/m^2)
Q Thermal energy (J)
R Optical reflectivity (unitless)
T Temperature (K)
U Internal energy (J)

Greek symbols

α Thermal diffusivity (m^2/s)
ε Radiant emissivity (unitless)
Γ Thermal mismatch factor (unitless)
κ Thermal conductivity (W/m K)
τ Pulse duration(s)
σ Complex thermal wave number
ω Modulation frequency (s^{-1})

Subscripts

i layer designation: 0 = top layer, 1 = second layer, etc.

REFERENCES

1. DP Almond, PM Patel. Photothermal Science and Techniques. Chapman and Hall, London, 1996.
2. D Fournier. The Mirage Effect in Photothermal Imaging. In: EA Ash, ed. Scanned Image Microscopy. London: Academic Press; 1980, pp 347–351.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

3. P-E Nordal, SO Kanstad. Photothermal Radiometry. Physica Scripta 20: 659–662, 1979.
4. DJ Crowther, LD Favro, PK Kuo, RL Thomas. Inverse scattering algorithm applied to infrared thermal wave images. Journal of Applied Physics 74(9): 5828–5834,1993.
5. YH Wong, RL Thomas, JJ Pouch. Subsurface Structures of Solids by Scanning Photoacoustic Microscopy. Applied Physics Letters 35(5): 368–369, 1979.
6. J-C Krapez, F Lepoutre, D Balageas. Early detection of thermal contrast in pulsed stimulated thermography. Journal de Physique 4(C7): 47–50, 1994.
7. WN Reynolds, GM Wells. Video-compatible thermography. British Journal of NDT 26(1): 40–44, 1984.
8. JWM Spicer, WD Kerns, LC Aamodt, JC Murphy. Measurement of Coating Physical Properties and Detection of Coating Disbonds by Time-Resolved Infrared Radiometry. Journal of Nondestructive Evaluation 8(2): 343–356, 1989.
9. V Vavilov. Infrared nondestructive testing of bonded structures: aspects of theory and practice. British Journal of Nondestructive Testing. 22(4): p 175–183, 1980.
10. CP Hobbs. Transient Thermography. Sensor Review 12(1): 8–13, 1992.
11. AG Bell. Upon the Production of Sound by Radiant Energy. Philosophical Magazine 11: 510–528, 1881.
12. HS Carslaw, JC Jaeger. Conduction of Heat in Solids. 2nd ed. Oxford: Clarendon Press, 1959.
13. C Starr. An Improved Method for the Determination of Thermal Diffusivities. Review of Scientific Instruments 8: p 61–64, 1937.
14. A Rosencwaig. Photoacoustic Spectroscopy: A New Tool for Investigation of Solids. Analytical Chemistry 47(6): p 592A–604A, 1975.
15. A Rosencwaig. Photoacoustics and Photoacoustic Spectrsocopy. New York: John Wiley and Sons, 1980.
16. CKM Patel, AC Tam. Pulsed Optoacoustic Spectroscopy of Condensed Matter. Reviews of Modern Physics 53(3): p 517–550, 1981.
17. GA West et al. Photoacoustic Spectroscopy. Review of Scientific Instrument. 54(7): p 797–817, 1983.
18. A Rosencwaig, A Gersho. Theory of the Photoacoustic Effect with Solids. Journal of Applied Physics 47(1): p 64–69, 1976.
19. YH Wong, RL Thomas, GF Hawkins. Surface and Subsurface Structure of Solids by Laser Phototacoustic Spectroscopy. Applied Physics Letters 32(9): 538–539, 1978.
20. RL Thomas et al. Subsurface Flaw Detection in Metals by Photoacoustic Microscopy. Journal of Applied Physics 51(2): p 1152–1156, 1980.
21. AC Boccara, D Fournier, J Badoz. Thermo-optical Spectroscopy: Detection by the 'Mirage Effect'. Applied Physics Letters 36(2): p 130–132, 1980.
22. JC Murphy, LC Aamodt. Photothermal Spectroscopy Using Optical Beam Probing: Mirage Effect. Journal of Applied Physics 51(9): p 4580–4588, 1980.
23. LC Aamodt, JC Murphy. Photothermal Measurements Using a Localized Excitation Source. Journal of Applied Physics 52(8): p 4903–4914, 1981.
24. JC Murphy, LC Aamodt. Optically Detected Photothermal Imaging. Applied Physics Letters 38(4): p 196–198, 1981.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

25. JGC Wetsel, FA McDonald. Subsurface Structure Determination Using Photothermal Laser-Beam Deflection. Applied Physics Letters 41(10): p 926–928, 1982.
26. RL Thomas et al. Thermal Wave Imaging for Nondestructive Evaluation. in IEEE Ultrasonics Symposium. IEEE, 1982.
27. G Busse. The optoacoustic and photothermal microscope: the instrument and its applications. In: E.A. Ash, Ed. Scanned Image Microscopy. London: Academic Press, 1980 P 341–345.
28. DP Almond, PM Patel, H Reiter. The testing of plasma-sprayed coatings by thermal-wave interferometry. Materials Evaluation 45(4): p. 471–475, 1987.
29. Z Ouyang, L Wang, X Wang, F Zhang, LD Favro, PK Kuo, RL Thomas. Lock-in Thermal Wave Imaging. In: Do Thompson, DE Chimenti, eds. Review of Progress in Quantitative NDE, NewYork: Plenum Press: p 383–387, 1997.
30. J Rantala, D Wu, G Busse, NDT of Polymer Materials Using Lock-in thermography with Water-coupled Ultrasonic Excitation. NDT&E International 31(1): p 43–49, 1998.
31. HI Syed, WP Winfree, KE Cramer. Processing Infrared Images of Aircraft Lapjoints. In Thermosense XIV: An International Conference on Thermal Sensing and Imaging Diagnostic Application. Orlando, FL: SPIE, 1992.
32. LD Favro, et al. Thermal-Wave Imaging of Corrosion and Disbonds in Aircraft Structures. In Nondestructive Inspection of Aging Aircraft San Diego, CA: SPIE, 1993.
33. JC Murphy, LC Aamodt, JWM Spicer, Principles of photothermal detection in solid. In: Principles and Perspectives of Photothermal and Photoacoustic Phenomena, A Mandelis ed. North-Holland: Elsevier. 1992, P 41–94.
34. RD Hudson. Infrared System Engineering. New York: John Wiley & Sons, 1969.
35. WL Wolfe, GJ Zissis, eds. The Infrared Handbook, Revised Edition ed. Department of the Navy: Washington, DC: Office of Naval Research, 1985.
36. JWN Spicer, R Osiander, JC Murphy. Time-resolved Infrared Radiometry Using Microwave Excitation. in Society of Experimental Mechanics Spring Conference, Baltimore, MD: SEM, 1994.
37. R Osiander, JWM. Spicer, J.C. Murphy, Thermal Nondestructive Evaluation Using Microwave Sources. Materials Evaluation 53(6): p 942–948, 1995.
38. MW Bowen, et al. Thermographic detection of conducting contaminants in composite materials using microwave excitation. In: Do Thompson, DE Chimenti, eds. Review of Progress in Quantitative NDE, New York: Plenum Press, 1995, p 453–460.
39. R Osiander, JWM Spicer, JC Murphy. Microwave-source time-resolved infrared radiometry for monitoring of curing and deposition processes. In: Thermosense XVII: An International Conference on Thermal Sensing and Imaging Diagnostic Applications. Orlando, FL: SPIE, 1995.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

9

Microwave

Alfred J. Bahr
SRI International, Menlo Park, California

Reza Zoughi
University of Missouri–Rolla, Rolla, Missouri

Nasser Qaddoumi
American University of Sharjah, United Arab Emirates

9.1 INTRODUCTION

As the need for NDE in our technology-based society grows, inspection requirements become ever more demanding and it is essential that we have as many NDE tools at our disposal as possible. Although not as widely known or understood as the NDE techniques discussed in earlier chapters, microwave NDE has proven to be very useful in certain applications. Microwave NDE may be defined as the inspection and characterization of materials and structures using high-frequency electromagnetic energy. For example, police routinely use radar "guns" to "nondestructively inspect" our moving automobile to determine its speed without having to be in physical contact (ride in the car with us). In addition, if the body of the car were made of a low-loss dielectric material, the microwaves would penetrate the car and could provide information on the

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

contents of the car as well. Noncontact inspection and the ability to penetrate dielectric materials are two of the most important attributes of microwave NDE.

9.1.1 Technical Overview

The microwave frequency region, although not rigidly defined, is generally taken to lie between a few hundred megahertz (MHz) and a few hundred gigahertz (GHz) (Table 9.1). The corresponding wavelengths in free space lie between 100 cm and 1 mm. The operating frequency can be selected to maximize the interaction of the electromagnetic energy with dielectric layers, voids, inclusions, surface flaws, material variations, and chemical species.

Microwave inspection generally consists of measuring various properties of the electromagnetic waves scattered by, or transmitted through, a test article. The signal strength inside a dielectric material is dictated by the incident power level, the loss factor of the material (ability to absorb microwave energy), and the frequency of operation. Specific characteristics of microwave measurements important in NDE are: (a) the incident and scattered waves are readily separated to enhance one's ability to measure small variations in the scattered signal, (b) the electrical phases of the waves vary rapidly with distance and/or frequency, thereby resulting in high sensitivity to variations in measured signal related to changes in distance or frequency, (c) the waves are polarized (electric-field vector lies along a specific line in the case of linear polarization), which allows selective

TABLE 9.1 The Electromagnetic Spectrum

Wavelength (m)	Frequency (Hz)	Photon Energy (eV)	Common Terminology
10^{-14}	3×10^{22}	10^{8}	Cosmic radiation
10^{-13}	3×10^{21}	10^{7}	X and Gamma radiation
10^{-12}	3×10^{20}	10^{6}	
10^{-11}	3×10^{19}	10^{5}	
10^{-10}	3×10^{18}	10^{4}	
10^{-9}	3×10^{17}	10^{3}	
10^{-8}	3×10^{16}	10^{2}	Ultraviolet
10^{-7}	3×10^{15}	10	
10^{-6}	3×10^{14}	1	Visible light
10^{-5}	3×10^{13}	10^{-1}	Infrared
10^{-4}	3×10^{12}	10^{-2}	
10^{-3}	3×10^{11}	10^{-3}	Microwaves
10^{-2}	3×10^{10}	10^{-4}	
10^{-1}	3×10^{9}	10^{-5}	
1	3×10^{8}	10^{-6}	

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

interaction with an elongated anomaly, and (d) the angular distribution of the scattering from an anomaly can be exploited for additional information. Spatial or lateral resolution perpendicular to the direction of wave propagation is either determined by the wavelength for far-field inspection, or by the dimensions of the transducer for near-field inspection. The spatial resolution along the direction of propagation is determined either by the measurement bandwidth (which can be large) or by the precision and stability of a single-frequency calibration made using a measurement standard (e.g., a layered plate with no defects).

Since microwave wavelengths are relatively small, the circuits used for signal transmission and analog processing tend to be distributed rather than lumped. In particular, hollow waveguides, coaxial lines, and printed microstrip lines are common. Also, like optical or ultrasonic waves, microwaves can be formed into beams and propagated between transducers. The transducers in this case are antennas, and their size is dictated by whether or not they are required to radiate efficiently. Efficient radiators have a size on the order of a wavelength, while antennas used for close-proximity inspection can be smaller. In the latter case, microwave and eddy-current inspection may be considered similar in some ways, except that eddy-current instruments operate at much lower frequencies.

9.1.2 Historical Perspective

One of the earliest examples of microwave NDE was discussed in a 1948 patent that describes a microwave technique for evaluating the moisture content in dielectric materials (1). This has been an important application of microwave NDE up through the present time because of the large changes in the microwave-frequency dielectric constant produced by the presence of water molecules in a material (2). The application of microwaves to NDE developed slowly, but a number of papers began to appear in the 1960s whose titles linked the words "microwave" and "nondestructive testing" (3–4). The applications discussed in these papers were primarily related to missile development and manufacturing. Applications during this period tended to be very specialized and limited, mainly because of inflexibility in the microwave instrumentation of the day and the need for specialized skills to design and operate a microwave NDE system. As we will discuss, modern microwave instrumentation has alleviated this situation.

9.1.3 Potential of the Technique

A partial list of microwave NDE applications that appeared during the 1960s and early 1970s would include the following: testing for delaminations in rocket casings, detecting defects in rocket propellants, measuring the burn rate of solid-propellant rocket motors, measuring the thicknesses of ablative shields for reentry vehicles, detecting voids in honeycombed ablative materials, detecting inclusions and porosity in ceramics and molded rubber, detecting surface cracks in artillery

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

shells, measuring epoxy-resin cure rates, measuring moisture content in dielectric materials, and measuring density variations in lumber. Most of these applications were in aerospace. Today, the inspection of new dielectric materials such as composites is an example of an important emerging application area for microwave NDE. Also, the evaluation of the properties and composition of mixtures, including the effect of curing in composites and resins, is receiving increased attention.

The modern era in microwave NDE was ushered in by the development of the microwave network analyzer in the early 1970s. This instrument provides the flexibility required for testing the use of microwave NDE in particular applications. In particular, it incorporates microprocessors that permit calibration and customization for each application, which make it easy to use. Another important instrumentation advance that occurred at this time was the development of stable, high-resolution, variable-frequency microwave synthesizers. These sources exhibit high spectral purity (which is important, for example, when it is necessary to maintain a background null) and are readily controlled by a separate computer.

The achievable resolution with microwave NDE can be considerably less than a free-space wavelength. For instance, in monitoring thickness variations in dielectric slabs and coatings, if resolution is taken to be the smallest thickness variation that can be detected, microwave NDE techniques have demonstrated measurement resolutions of a few micrometers at 10 GHz (a wavelength of 3 cm in free space). Alternatively, if resolution is taken to be the smallest spatial distance between two barely resolvable defects, then *near-field* microwave and millimeter-wave techniques have demonstrated resolutions of less than a tenth of a wavelength. Also, fatigue cracks on metal surfaces with widths in the range of a few micrometers have been detected at frequencies of 10 GHz and lower.

As in all NDE methods, it is important to be able to simulate the inspection process to optimize it or to extract quantitative information from the results (such as flaw dimensions or probability of detection). Electromagnetic theory is very well developed, so once the underlying theoretical foundation for the interaction of microwaves with a given medium is understood and modeled, electromagnetic software packages are readily available (or can be developed) that permit such simulations.

A list of demonstrated and potential applications of microwave NDE includes the following:

- Accurate thickness measurements of coatings, single dielectric slabs, and layered dielectric composites made of plastics, ceramics, or any other type of dielectric material whose loss factor (microwave energy absorption) is not too large.
- Detection of minute thickness variations in each layer of a stratified dielectric medium.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Detection of disbonds, delaminations, and voids in stratified media and (potentially) the determination of the depth of these flaws.
Inspection of thick plastics and glass-reinforced composites for interior flaws, fiber-bundle orientation and breakage, moisture content, etc.
Detection and estimation of porosity in ceramics, thermal barrier coatings, plastics, glass, etc.
Detection and evaluation of corrosion under paint and thick stratified composite-laminate coatings.
Detection and measurement of moisture content in wood, food grains, textiles, etc.
Impact damage detection and evaluation for reinforced composite structures, including graphite composites.
Fiber orientation determination in graphite and glass-reinforced composites.
Accurate characterization of the constituents (e.g., volume content, dielectric constant) in dielectric mixtures.
Detection and evaluation of cure states in chemically reactive materials such as carbon-loaded rubber, resin binder, cementious materials, etc.
Inspecting concrete for constituent determination, reinforcing-bar location, chloride detection, safety evaluation, etc. Detecting grout in masonry is also possible.
Detection and sizing of surface cracks (stress and fatigue) in metals. This capability extends to cracks filled or covered by dielectric materials such as paint, rust, dirt, etc.
Profiling and roughness evaluation of metal surfaces.
Imaging of localized and extended-area interior and surface defects.
Heating lossy dielectric materials for medical or thermographic-imaging purposes.
Radiometric detection of cancerous cells in the breast, skin, etc.

9.1.4 Advantages and Disadvantages

As mentioned earlier, an important feature of microwave inspection is that the transducer (antenna) need not be in contact with the object being inspected, thus permitting inspection of hard-to-reach areas or moving parts. However, only dielectric objects can be inspected internally; inspection of conducting objects is limited to their surfaces. One of the most powerful aspects of microwave NDE techniques is the availability of many different probes/sensors from which to select for best results. Large scan areas can be achieved by using arrays of sensors.

The network analyzer and other associated equipment mentioned above are well suited for carrying out laboratory experiments, but are relatively expensive.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

However, custom hardware systems designed for microwave NDE applications need not be expensive. One reason for this is that microwave signal-processing devices such as hybrid couplers, amplifiers, etc., as well as microwave sources, have become smaller and less expensive due to advances in microwave integrated-circuit technology. Thus, hardware for a specific application can be developed and built to be relatively inexpensive, simple in design, handheld, battery-operated, operator-friendly, and can be used for real-time on-line inspection. In a majority of microwave NDE applications where anomaly detection is the primary objective, there is very little need for complicated postsignal processing and the user interface is quite simple. The operator need not be a microwave expert to conduct microwave NDE measurements once a system has become operational. Since the required operating powers for most NDE applications are not more that a few mW (excluding ground-probing radars and microwave heating sources*) and since, in the majority of cases, the frequency bandwidth is narrow and the probes are small and located close to the object being inspected, microwave NDE systems do not produce detectable *electromagnetic interference* (EMI) nor are they affected by EMI from other sources. Finally, there are no known environment hazards or undesirable byproducts associated with these techniques.

Microwave NDE techniques have proven their worth in many NDE applications and should be considered when new inspection problems arise. In some cases one of these techniques will prove to be the technique of choice. In other cases it will prove to be a useful complement when used in conjunction with other methods. It is the objective of this chapter to provide the basic understanding needed to evaluate these possibilities.

9.2 BACKGROUND

At low frequencies, electromagnetic signals are conveniently represented by voltages and currents measured at points within a circuit. Radiation from a slowly varying current is negligible and all circuit elements are small compared to a wavelength (i.e., discrete). The ratio of voltage to current (called impedance) completely describes the interaction of a source with a load. On the other hand, microwave NDE is best understood in terms of the interaction between waves (or fields) and material structures rather than in terms of voltages and currents. At microwave frequencies (3×10^8 to 3×10^{11} Hz) radiation effects can be significant and circuit elements are typically on the order of a wavelength or larger in size (i.e., distributed). At such high frequencies it is usually convenient (and best) to represent electromagnetic signals by waves. Impedance is still a useful concept,

*The power level inside a typical microwave oven is several hundred watts.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

but now it is understood to be given by its more general definition as the ratio of certain electric- to magnetic-field components. Microwave NDE exploits the fact that an electromagnetic wave will have its amplitude and/or phase (delay) changed by passing through, or scattering from, a material body. The amount of change in these measurable wave characteristics is dictated by the body's dielectric, magnetic, and geometric properties.

In physical terms, the effects a material body can have on a microwave-frequency electromagnetic wave can be summarized as follows: (a) the material can absorb power from the wave, causing the wave to attenuate or the quality factor*, Q, of a resonator containing the material to decrease, (b) the material can change the velocity of propagation of the wave, and hence its wavelength, causing the wavefront to bend, the phase angle to change, and the signal to be delayed, (c) the material can have a wave impedance that differs from its surroundings, causing the wave to reflect from the surfaces of the material body, (d) the material can increase the stored energy in a resonator, causing the resonant frequency to shift, and (e) the body can scatter energy in many directions or from one mode (field configuration) into other different modes. All of these physical effects depend on the material's constitutive parameters, as well as on the body's shape and size. Thus, microwave NDE is capable of providing information on both material parameters and physical dimensions. This capability parallels that associated with ultrasonic testing, except that the material parameters in that case are mechanical rather than electromagnetic and ultrasonic wavelengths are five orders of magnitude smaller.

This section reviews some of the basic concepts of fields and waves that are needed to understand the principles of microwave NDE. We begin by reviewing the definitions of some material parameters and then discuss how fields and waves interact with material bodies.

9.2.1 Material Parameters

Skin Depth

All matter may be considered to be made up of atoms in which negatively charged electrons move in various orbits around a positively charged nucleus. The electrons are held in these orbits as a result of the attractive forces between the negatively charged electrons and the positively charged nucleus (5). In conductors (metals), electrons in the outer orbit (or the valance band) are only loosely bound and can easily move from one atom to the next. In addition, there are many such mobile electrons available in conductors. However, their movement is random

*Q is defined as the angular resonant frequency times the energy stored in the resonator, all divided by the average power lost in the resonator.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

until the presence of an impressed electric field causes them to move in concert, which gives rise to a conduction current.

For time-varying impressed fields, electromagnetic forces expel the conduction electrons from the interior and the current flows close to the surface of the conductor (i.e., the current is confined to a "skin" at the surface of the conductor). *Skin depth* is the name given to the parameter that characterizes the depth of penetration of this surface current. If the tangential components of the impressed fields do not vary rapidly over the surface in a distance that is small compared with the surface curvature of the conductor*, the skin depth, δ, is given by the following formula (6):

$$\delta = \frac{1}{\sqrt{\pi f \mu \sigma}} \tag{1}$$

where f is the temporal frequency of harmonic variation in Hz, μ is the magnetic permeability (equal to $4\pi \times 10^{-7}$ henry/m for vacuum, dielectrics, and nonferromagnetic conductors), and σ is the conductivity in siemen/m. Thus, the skin depth decreases as the frequency, or the permeability, or the conductivity increases. The important implication of this behavior for microwave NDE is that microwave electromagnetic techniques are not usually effective for inspecting the interiors of highly conducting materials. However, microwave NDE can still be useful for inspecting conductors in some circumstances, such as when the conductivity is not too large, the anomaly of interest is on or near the surface of the conductor (e.g., surface cracks, impact damage, etc.), or the material to be inspected is thin compared to a skin depth. Obviously, whether a signal is detectable or not after it has penetrated a conductor also depends on the strength of the illuminating field and the sensitivity of the measuring equipment.

Typical skin depths in good conductors at microwave frequencies are quite small. For example, let the material be copper ($\sigma = 5.7 \times 10^7$ siemen/m) and the frequency be 3×10^9 Hz. Eq. (1) becomes

$$\delta = \frac{1}{\sqrt{\pi \times 3 \times 10^9 \times 4 \times \pi \times 10^{-7} \times 5.7 \times 10^7}} = 1.2 \text{ micrometers!}$$

Because of the square-root relation, even an order-of-magnitude decrease in the conductivity or frequency only increases the skin depth by a factor of about 3.

Dielectric Constant

In dielectrics (insulators) the electrons are held tightly in place around the nucleus, and unlike in conductors they are not free to move from one atom to

*The skin-depth formula (Eq. (1)) is only strictly valid for plane waves and planar surfaces.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

the next (7). Consequently, no conduction current arises when a dielectric material is placed in a static electric field. However, since the electrons and the nucleus are oppositely charged and are separated from one another, a dielectric material can be viewed as a collection of randomly oriented dipole moments. Because of this random orientation, the resulting net macroscopic dipole moment is zero. Now, when a dielectric material is placed in a static electric field two things happen. First, the electrons (negative charges) and the nucleus (positive charge) of each atom in the material experience a force (Figure 9.1). This force results in a relative increase in the distance separating the centroids of the positive and negative charges, which, in turn, increases the dipole moment associated with the atom. Second, all the individual dipole moments align themselves with the impressed electric-field vector, $\vec{E}$, giving rise to an overall net dipole moment.

The slight stretching of the atoms in a dielectric increases the dipole moment of each atom and therefore increases their potential energy. Thus, a dielectric composed of many such stretched atoms can be thought of as containing stored electrical energy (8). This is the underlying principle behind the enhanced ability of a capacitor to store electrical energy when the air between the capacitor's electrodes is replaced by a dielectric.

The increase in net dipole moment caused by alignment with an impressed electric field is described by the term *polarization*. There are three basic mechanisms of polarization in homogeneous dielectric materials: (a) *electronic*, which occurs in most dielectrics as the centroids of the negative charges (electrons) and the nucleus (positive charge) in an atom experience a slight displacement in the presence of an electric field, (b) *orientational*, which occurs in polar materials that naturally possess finite electric-dipole moments (e.g., water molecules) that are randomly oriented in the absence of an electric field, but, in the presence of an electric field, align themselves with the field and produce a net polarization vector, and (c) *ionic* (molecular), which occurs in materials such as electrical ceramics that contain positive and negative ions (similar to electronic polarization).

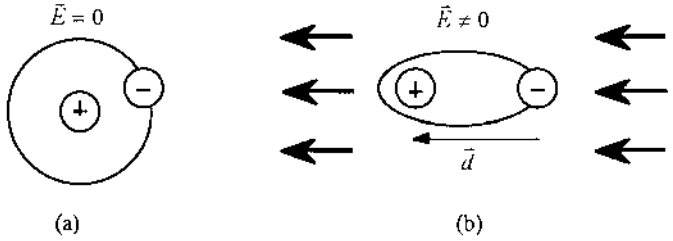


FIGURE 9.1 Macroscopic representation of a dielectric atom in the (a) absence and (b) presence of an impressed field.

Copyright © Marcel Dekker, Inc. All rights reserved.

It is helpful to consider the process of polarization from a macroscopic point of view. Assuming the positive and negative charges associated with the nucleus and the electrons, respectively, have magnitude q, the resulting *vector electric-dipole moment* in the presence of an electric field is given by

$$\vec{p} = q\vec{d} \qquad (C{\cdot}m) \tag{2}$$

where d is the distance between the charges and the unit "C" denotes coulombs. The macroscopic *electric polarization vector*, $\vec{P}$, associated with the dielectric material is consequently the limit of the sum of all dipole moments per unit volume, viz.,

$$\vec{P} = \lim_{\Delta v \to 0}\left[\frac{1}{\Delta v}\sum_{i=1}^{n}\vec{p}\right] \qquad (C/m^2) \tag{3}$$

where n dipoles are present in a differential volume, Δv. Therefore, the polarization vector represents the average effect of all the electric-dipole moments generated in the dielectric material.

The proportionality constant between the polarization vector and the impressed electric field is called the *electric susceptibility*, χ_e, of the dielectric material. Similarly, the proportionality constant between the *electric flux density*, $\vec{D}$, and the impressed electric field is called the *dielectric constant*, ε, (or, for the static case, the permittivity) of the dielectric material and is given by the sum of free-space and material contributions, viz.,

$$\varepsilon = \varepsilon_0(1 + \chi_e) \qquad (F/m) \tag{4}$$

where ε_0 is the *permittivity of vacuum* (8.854×10^{-12} F/m) and the unit "F" denotes farads. The electrical properties of a polarized material are completely described by its dielectric constant, and it is this material parameter that we most often wish to measure.

When a dielectric material is placed in a time-varying electric field its associated dipole-moment vectors change their directions in response to the changing direction of the electric-field vector. Consequently, the polarization vector associated with a dielectric material, and the resulting dielectric constant, are influenced by the alternating nature of the field. The model of electronic polarization in which the stretching of the atoms increases potential energy is analogous to a classical resonant mass–spring system (with friction) or to an electrical R–L–C circuit (9). The effect of frequency on the material properties may then either be modeled mechanically in terms of an equivalent spring coefficient, friction, and mass (usually the electron mass is used, since the nucleus is much heavier and moves very little compared to the electrons; however, for ionic polarization, resonance tends to occur at lower frequencies due to the heavier mass of the ions), or modeled electrically in terms of an equivalent resonant frequency and quality factor (or damping). As mentioned above,

Copyright © Marcel Dekker, Inc. All rights reserved.

harmonic variation causes a periodic change in the polarity associated with the polarization vector. Since the nucleus of an atom is much heavier than the electrons orbiting it, the electron cloud primarily moves back and forth around the nucleus (8). The time delay in this movement can be thought of as being associated with friction in the system, which translates to a certain amount of the impressed energy being absorbed by the material (converted to heat). Therefore, in the time-varying case both the energy storage and energy absorption must be characterized by the dielectric parameters of a material.

The absorbed energy is macroscopically related to the static conductivity (which is very small in dielectric materials) as well as to the dielectric-hysteresis effect associated with the rotation of the polarization vector in an alternating electric field. For good dielectrics, this latter effect is a greater cause of energy absorption than is the static conductivity. In general, the dielectric properties of a material, under harmonically time-varying conditions, are described by its effective complex dielectric constant (10).

$$\varepsilon^* = \varepsilon' - j\varepsilon'' = \varepsilon' - j\frac{\lceil \sigma_e \rceil}{\lfloor \omega \rfloor} \tag{5}$$

where ε' is the *absolute permittivity* (due to displacement current) characterizing the ability of the material to store energy, ε'' is the *absolute loss factor* (due to the static and the alternating conductivities) characterizing the ability of the material to absorb energy, σ_e is the *equivalent conductivity* due to both the static and the alternating conductivities, $j = \sqrt{-1}$, and ω is the *angular frequency.* It is convenient to define a relative complex dielectric constant by normalizing the complex dielectric constant to ε_0, the permittivity of vacuum, viz.,

$$\varepsilon_r = \frac{\varepsilon^*}{\varepsilon_0} = \varepsilon_r' - j\varepsilon_r'' \tag{6}$$

where ε_r' is referred to as the *relative permittivity* and ε_r'' is referred to as the *relative loss factor*, or simply, loss factor. Another commonly used material parameter is the *loss tangent*, $\tan\delta$, defined as the ratio of the loss factor to the permittivity, $\varepsilon_r''/\varepsilon_r'$. Lossless materials have a loss tangent of zero, while $\tan\delta \ll 1$ indicates a low-loss material and $\tan\delta \gg 1$ indicates a high-loss material.

The frequency dependencies of these dielectric parameters are determined by the different sorts of relaxation processes that are present and are generally more complicated than the simple inverse-frequency dependence indicated in Eq. (5) above. For more detailed information regarding the effect of time-varying electric fields on dielectric materials the reader is referred to (11).

The macroscopic dielectric constant of a material composed of several constituents (i.e., a dielectric mixture or a composite material) is dependent upon the dielectric constant of each of its constituents, their volume fractions, their distribution in the mixture, the orientation of the constituents with respect to the

Copyright © Marcel Dekker, Inc. All rights reserved.

impressed electric-field vector, any polymerization (i.e., molecular bonding and curing) that may have occurred during the production of the material, and the operating frequency. Thus, the following types of information are obtainable when using microwave NDE to evaluate the dielectric properties of a composite material or a mixture:

1. Dielectric constant of a constituent in the mixture
2. Volume fraction of a constituent (e.g., porosity in polymers and ceramics)
3. Cure state
4. Anisotropy associated with the mixture
5. Physical and mechanical properties which are related to the cure state of the material (e.g., compressive strength in concrete)

There are various theoretical and empirical dielectric mixing models that can be used to extract this kind of information (11).

There are numerous microwave-NDE methods available for measuring the dielectric properties of materials. Which technique is best depends on the following considerations:

State of material to be evaluated (i.e., liquids, solids, gases)
Real-time requirements
Required measurement accuracy
Measurement environment
Loss tangent of the material (i.e., low-loss vs. high-loss)
Nondestructive requirements
Noncontact requirements
Geometry of the material under test (i.e., cylindrical fibers, sheets, etc.)
Type of information being sought

Examples of some of the more prominent measurement approaches appropriate to different applications will be discussed in later sections.

9.2.2 Basic Electromagnetic Wave Concepts

Plane Waves and Characteristic Impedance

The propagation of electromagnetic waves in linear media is governed by Maxwell's equations (12). These equations are the differential equations for the vector components of the electric and magnetic fields, and the equations take different forms depending on which orthogonal coordinate system (Cartesian, spherical, etc.) is used in expressing the vector fields. It is sufficient for our purposes here to consider only the simplest case, i.e., uniform plane *transverse electromagnetic* (TEM) waves in rectangular coordinates. It is also convenient to restrict ourselves to fields that vary with time according to the complex

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

exponential function (harmonic) $e^{j\omega t}$ where $j = \sqrt{-1}$, ω is the angular frequency, and t is the time. There is little loss of generality in using such a time function since any physically realizable time variation can be decomposed into a spectrum of such functions by means of the Fourier integral (13). Also, in practice, microwave sources typically generate sinusoidal signals.

Plane waves are defined to have no spatial variation in the plane transverse to the direction of propagation. Taking the z-direction as the direction of propagation, the differential equation (wave equation) for the *transverse electric-field vector*, $\vec{E}_t$, in the x–y plane is

$$\frac{d^2\vec{E}_t}{dz^2} + k^2\vec{E}_t = 0 \tag{7}$$

where $k^2 = \omega^2\mu\varepsilon$. The solution of this equation has the form $A^{\pm}e^{\mp jkz}$ where $A^{\pm}$ are complex coefficients that are determined by boundary conditions. Since a time function $e^{j\omega t}$ was assumed, $A^{+}e^{-jkz}$ represents propagation along the positive z-axis and the other solution represents propagation in the opposite direction (e.g., a wave reflected from a material boundary). The *propagation constant* or *wave number*, k, is given by

$$k = \omega\sqrt{\mu\varepsilon} = \frac{\omega}{v_c} = \frac{2\pi f}{v_c} = \frac{2\pi}{\lambda} \tag{8}$$

where v_c is the *phase velocity* of electromagnetic waves in a medium with parameters μ and ε, and λ is the wavelength corresponding to frequency f in the same medium. Note that the phase velocity is solely determined by these material parameters. The magnetic fields associated with these forward- and backward-propagating waves are perpendicular to the corresponding electric fields and are given by

$$H_x = \mp\frac{E_y}{\eta}$$

$$H_y = \mp\frac{E_x}{\eta} \tag{9}$$

The factor

$$\eta = \sqrt{\frac{\mu}{\varepsilon}} \tag{10}$$

has the dimensions of ohms and is called the *characteristic* or *intrinsic impedance* of the medium. In free space (vacuum) it is equal to $120\,\pi = 377$ ohm. For uniform plane waves, the electric and magnetic fields are orthogonal to each other as well as to the direction of propagation (TEM *waves*). Consequently, the vector

Copyright © Marcel Dekker, Inc. All rights reserved.

cross product between the electric and magnetic field vectors ($\vec{E}_t \times \vec{H}_t$) points in the direction of propagation.

Standing Waves

Consider the case where two plane waves, having with the same electric-field direction but different complex amplitudes, are traveling in opposite directions. The total electric field at a point in space is then just the sum of the electric fields in the two waves, i.e.,

$$E_{\text{total}} = E_1 e^{-jkz} + E_2 e^{jkz} \tag{11}$$

where the harmonic temporal variation, $e^{j\omega t}$, has been suppressed. If we compute the magnitude of the total electric field (assuming $|E_2| \leq |E_1|$) we see that it varies as the $\cos 2kz$ and has a maximum and minimum value of $|E_1| + |E_2|$ and $|E_1| - |E_2|$, respectively. Shifting the phase of one of the waves with respect to the other causes the positions of these extrema to shift along z. This varying amplitude pattern caused by interference between the two waves is called the *standing-wave pattern* and the ratio

$$\text{SWR} = \frac{|E_1| + |E_2|}{|E_1| - |E_2|} \tag{12}$$

is called the *standing-wave ratio*. This standing-wave pattern repeats every one-half wavelength.

Thus, if the amplitude and/or phase of one of the waves has been influenced by interaction with a material or structure, direct measurements of the corresponding standing-wave pattern (magnitude and extrema positions) will contain information about the material or structure. Historically, such measurements were very common in the microwave art, but are less common today because of the enhanced measurement capabilities provided by the modern microwave network analyzer.

Reflection and Transmission at a Planar Interface (Normal Incidence)

Let the half-space $z > 0$ be filled with a lossless homogeneous dielectric material with a characteristic impedance η, and let the half-space $z < 0$ be free space with characteristic impedance η_0. From the region $z < 0$ let a plane wave be incident normally on the interface. When this wave strikes the interface, $z = 0$, a portion of the wave will be reflected and a portion will be transmitted into the region $z > 0$. Now, the tangential electric and magnetic fields must be continuous across the boundary (14). Application of this boundary condition allows us to calculate the fractions of the incident wave that are reflected and transmitted. We find that

Copyright © Marcel Dekker, Inc. All rights reserved.

the *complex reflection coefficient* of the boundary (the ratio of reflected to incident electric field) is

$$\Gamma = \frac{E_{\text{refl}}}{E_{\text{inc}}} = \frac{\eta - \eta_0}{\eta + \eta_0} \tag{13}$$

and the *complex transmission coefficient* (the ratio of transmitted to incident electric field) is

$$\tau = \frac{E_{\text{trans}}}{E_{\text{inc}}} = \frac{2\eta}{\eta + \eta_0} \tag{14}$$

Finally, we can write the standing wave ratio in terms of the reflection coefficient as

$$\text{SWR} = \frac{1 + |\Gamma|}{1 - |\Gamma|} \tag{15}$$

In practice, it is usually more convenient to measure the reflection and transmission coefficients rather than the electric (or magnetic) fields themselves. We see that both coefficients contain information about the material properties and thus are useful parameters to measure in carrying out the microwave NDE of materials.

If a composite structure is composed of layers of different materials, multiple internal reflections will occur and the interference phenomenon becomes more complex. However, using multiple-frequency measurements of the reflection and transmission from and through such structures, respectively, it is possible to extract information about the layer thicknesses as well as their material properties by combining these measurements with a suitable multilayer model (15).

Scattering Parameters

The reflection and transmission coefficients defined above are examples of a more general representation of microwave networks called *scattering parameters*. A detailed discussion of scattering parameters is beyond the scope of this chapter, but since they form the basis of modern microwave measurements using a network analyzer, at least a brief introduction to the concept is in order.

Voltages, currents, and impedances cannot be measured directly at microwave frequencies. The quantities that *can* be measured directly are the amplitudes and phase angles of the scattered, e.g., reflected, waves in the transmission lines (or waveguides) that are connected to a microwave structure (test sample). These scattered-wave quantities are defined at selected reference planes in the transmission lines and are measured relative to the incident-wave amplitudes and phase angles at these reference planes (note that the term "scattered" includes waves that are transmitted *through* the structure). In view of the linearity of the field

Copyright © Marcel Dekker, Inc. All rights reserved.

equations and most microwave structures, the complex scattered-wave amplitudes are linearly related to the complex incident-wave amplitudes. The matrix describing this linear relationship is called the *scattering matrix* (16).

We envision the ith wave incident on the microwave structure (or a port) as being associated with an equivalent voltage, V_i^+, normalized such that $(1/2)|V_i^+|^2$ is equal to the power transmitted by the wave. Similarly, the kth wave reflected from the structure is associated with a voltage, V_k^-. The relation between reflected and incident (voltage) waves can be written as

$$\begin{bmatrix} V_1^- \\ \vdots \\ V_N^- \end{bmatrix} = \begin{bmatrix} S_{11} & \cdots & S_{1N} \\ \vdots & \ddots & \vdots \\ S_{N1} & \cdots & S_{NN} \end{bmatrix} \begin{bmatrix} V_1^+ \\ \vdots \\ V_N^+ \end{bmatrix} \tag{16}$$

where the S_{ki} are the scattering parameters (reflection and transmission coefficients) and there are N transmission lines (ports). Thus, S_{11} is the reflection coefficient at port 1 when all the other ports are terminated in a matched load (no signal is inputted to them or reflected from them), S_{21} is the transmission coefficient between ports 1 and 2 when a wave is incident on port 1 and all other ports are terminated in a matched load, etc. A network analyzer measures the scattering parameters directly as ratios between reflected or transmitted and incident waves.

Scattering by a Finite-Size Material Body

In the example discussed above, we calculated the plane-wave scattering from a plane interface between two different materials. In microwave NDE we are also interested in measuring the scattering from a material body (or a flaw within a material) of finite size. Obviously, the magnitude and phase of this scattering depends not only on the material properties but also on the dimensions and shape of the body or flaw. In general, this is a complicated situation to model. However, we can illustrate the point for the simple case of scattering from a sphere of radius, a_s, where a_s is much smaller than a wavelength (Rayleigh scattering).

Consider a plane monochromatic wave to be incident on the spherical scatterer. The incident fields induce electric ($\vec{p}$) and/or magnetic ($\vec{m}$) dipole moments in the small scatterer and these dipoles radiate energy in all directions. Far away from the scatterer, the scattered electric and magnetic fields are (17)

$$\vec{E}_{sc} = k^2 \frac{e^{-jkr}}{r} \left[(\hat{n} \times \vec{p}) \times \hat{n} - \eta_0 \hat{n} \times \vec{m} \right] \tag{17a}$$

and

$$\eta_0 \vec{H}_{sc} = \hat{n} \times \vec{E}_{sc} \tag{17b}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

respectively, where $\hat{\boldsymbol{n}}$ is a unit vector in the direction of observation and r is the distance from the scatterer.

If the scatterer is a nonmagnetic dielectric sphere, the electric dipole moment induced by the incident electric field, $\vec{E}_{\text{inc}}$ is

$$\vec{p} = \left(\frac{\varepsilon_{\text{rel}} - 1}{\varepsilon_{\text{rel}} + 2}\right) a_s^3 \vec{E}_{\text{inc}} \tag{18}$$

where ε_{rel} is the dielectric constant of the sphere relative to the surrounding dielectric material. Thus, as expected, the scattering depends on both the relative dielectric constant and size of the sphere. On the other hand, if the scatterer is a perfectly conducting sphere, both an electric and a magnetic dipole moment are induced. The electric dipole moment is

$$\vec{p} = a_s^3 \vec{E}_{\text{inc}} \tag{19}$$

and the magnetic dipole moment is

$$\vec{m} = -\frac{a_s^3}{2} \vec{H}_{\text{inc}} \tag{20}$$

In this case, only the sphere size affects the scattering. Also, if a number of such conducting spheres were to be dispersed throughout a dielectric material, we see that both the effective permittivity and permeability of the composite material would be affected.

Resonance

Microwave-resonator sensors are easily constructed from partially closed metallic cavities or from a high-dielectric-constant material body (such as a sphere) immersed in another material with a lower dielectric constant (such as air). Resonance occurs at frequencies where the electromagnetic waves within these structures reflect back and forth many times from the structure boundaries in a way that satisfies the boundary conditions (in fact, the number of significant reflections is equal to the resonator quality factor). For purposes of microwave NDE a portion of the electromagnetic fields supported by the resonator structure must interact with the material or structure being characterized or inspected. Hence, if the resonator sensor is a closed cavity, the material under test must be within the cavity and this approach is often used for precise measurements of dielectric constant and permeability. More often, however, a resonator sensor is used that is partially open so that some electromagnetic fields extend outside of the physical cavity structure. This open structure is then brought into proximity with the material or structure under test. In this way the microwave sensor (resonator) and test sample need not be in contact and they can be easily moved relative to one another.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

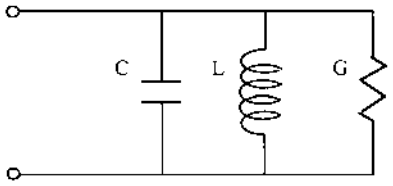


FIGURE 9.2 Parallel-resonant circuit.

At low frequencies resonators are formed from either a series or parallel connection of a capacitor, C, an inductor, L, and a resistor, R (or conductance, G). Consider the parallel connection case shown in Figure 9.2. The *input admittance* (inverse of impedance) in this case is

$$Y_{in} = G\left[1 + 2jQ_0(\Delta f/f_0)\right] \tag{21}$$

where

$$Q_0 = \frac{2\pi f_0 C}{G} \tag{22}$$

is the *resonator quality factor*, Δf is the *frequency deviation from resonance*, and f_0 is the resonant frequency given by

$$f_0 = \frac{1}{2\pi\sqrt{LC}} \tag{23}$$

Now suppose that the capacitor is filled with a dielectric material whose dielectric constant changes (e.g., due to curing). Then if we measure the change in resonant frequency caused by a change in the real part of the dielectric constant, $\Delta\varepsilon'$, it can be shown from the above equations that

$$\frac{\Delta\varepsilon'}{\varepsilon'} = -2\frac{\Delta f_0}{f_0} \tag{24}$$

Note that a *decrease* in the resonant frequency corresponds to an *increase* in the dielectric constant. Physically, this behavior comes about because an increase in dielectric constant corresponds to an increase in stored electric energy, and thus the resonator appears to be electrically larger and has a lower resonant frequency.

Similarly, a change in the loss component of the dielectric constant (imaginary part), $\Delta\varepsilon''$ produces a corresponding change in reciprocal Q_0, $\Delta(1/Q_0)$, and we have

$$\frac{\Delta\varepsilon''}{\varepsilon''} = \Delta\left(\frac{1}{Q_0}\right) \tag{25}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Thus, measurement of these two resonator parameters provides a direct measure of variations in dielectric constant.* As an alternative to $1/Q_0$, one could measure changes in the input resistance at resonance. It should be noted that the accuracy of resonator techniques degrades as the power lost in the material under test increases.

At microwave frequencies, short-circuited or open-circuited sections of transmission lines or waveguides are often employed as resonators whose resonances occur when the section length is an even or odd multiple of a quarter wavelength, respectively. When measured at special locations in the transmission line(s) connected to these resonators, the input impedance near resonance of such devices (as well as other types of resonator cavities) behaves exactly like that of a low-frequency lumped-circuit resonator (18). Thus, we can model a microwave resonator as a lumped circuit (near resonance). If we generalize to the case where the material under test only partially fills the resonator, we can write

$$C = C_0[1 + +K(\varepsilon' - 1)] \tag{26}$$

where K is an *empirical filling factor* and C_0 is the *effective resonator capacitance* when the material under test has been removed. Substitution of this expression for C into that for the resonant frequency (Eq. (23)) leads to the result

$$\frac{\Delta\varepsilon'}{1 + K(\varepsilon' - 1)} = -\frac{2}{K}\,\frac{\Delta f_0}{f_0} \tag{27}$$

This expression only differs from Eq. (24) in the appearance of a filling factor. The resonant frequency of the microwave resonator can be determined by observing either the transmission through the resonator (if it has two ports), or the reflection from the resonator, as a function of frequency.

Similarly, the change in the loss component of the dielectric constant can be conveniently monitored by observing the change in the input resistance of the resonator at resonance (this quantity is related to Q_0). When normalized to the reference impedance of the measurement system (typically 50 ohm), this input resistance, r_0, is given in terms of the input reflection coefficient at *resonance*, Γ_0 (a real number), by

$$r_0 = \frac{1 + \Gamma_0}{1 - \Gamma_0} \tag{28}$$

*Resonator quality factor can be determined from a measurement of the frequency bandwidth between points on the resonator response curve that are one-half the peak amplitude.

Copyright © Marcel Dekker, Inc. All rights reserved.

Assuming that material losses dominate other resonator losses (e.g., ohmic wall losses), the change in the imaginary part of the dielectric constant is given by

$$\frac{\Delta\varepsilon''}{\sqrt{1+K(\varepsilon'-1)}} = M\Delta\left(\frac{1}{r_0}\right) \tag{29}$$

where M is another empirical (calibration) factor. This expression for the change in the loss component of the dielectric constant is analogous to the one obtained for a resonated material-filled capacitor (Eq. (25)).

Thus, measurements of the resonant frequency and input resistance of a microwave resonator provide information about the dielectric properties of a material that interacts with the resonator fields.

9.3 MICROWAVE EQUIPMENT

Any microwave inspection system or equipment is comprised of several microwave components and devices that are assembled efficiently to provide proper and useful output signals. Laboratory microwave equipment is generally bulky, relatively expensive, and can provide for a wide range of functions throughout a relatively wide frequency band. Custom-designed microwave equipment, for nondestructive testing purposes, can be made to be small, modular, frequency specific, and relatively inexpensive. However, it is very important to note that laboratory equipment is essential for designing and building custom inspection systems. There is a host of microwave equipment, components, and devices offering different functionalities. The reader is referred to (19) for further information. In this section the functions and attributes of several of the more prominent pieces of microwave equipment, components and devices will be presented.

Microwave Oscillators. This is one of the most essential pieces of equipment in a laboratory. This equipment produces a microwave signal at a prescribed frequency and power level. As a laboratory piece of equipment, it is usually capable of producing a single-frequency signal (known as continuous wave or CW signal) or it can produce a swept frequency signal which is linearly ramped between a start and an end frequency in a prescribed duration. Consequently, this piece of equipment is usually referred to as a sweep oscillator. Modular and small oscillators capable of CW or swept-frequency signal generation are also available in the form of *voltage controlled oscillators* (VCO), *Yttrium Iron Garnet* (YIG) oscillators and cavity-tuned Gunn oscillators, to name a few. Cavity-tuned Gunn oscillators are compact, stable and reasonably pure microwave signal generators with output power levels in the tens of milliwatts. In addition, these oscillators are activated using a dc input voltage to the Gunn diode which is housed in a cavity (i.e., a resonator) whose dimensions are such that a

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

specific frequency can resonate within it while providing a relatively high quality factor (i.e., a relatively pure, single frequency, signal). Subsequently, the generated microwave signal is coupled from the cavity into a waveguide through a small aperture. There are many manufacturers producing any of the above mentioned types of oscillators (and more). As examples and for more detailed specification of various oscillators one may consult with the Agilent Technologies Test and Measurement Catalog* and Millitech Millimeter Wave Products Catalog†.

Network Analyzers. A network analyzer is an extremely accurate piece of laboratory equipment for calibrated measurement of phase and magnitude of microwave signals (reflected and transmitted) measured at some desired measurement reference plane. They are wide band and digitally controlled and provide for many important and desirable measurement features. They are also extensively used for characterizing the scattering parameters (e.g., reflection and transmission properties) of a microwave component or device. Consequently, they are relatively expensive. HP8510 and HP8720 series are examples of these vector network analyzers. For detailed specifications of this piece of equipment the reader may consult with the Agilent Technologies Test and Measurement Catalog.

Waveguides. Waveguides are hollow tubes with rectangular or circular cross sections. Waveguides are in the family of transmission lines and allow for propagation of microwave signals without producing much signal attenuation. Waveguides support *transverse electric* (TE) or *transverse magnetic* (TM) fields and cannot support TEM fields. This is due to the fact that waveguides are single conductor transmission lines and the TE and TM modes result from the solution of Maxwell's equations given the boundary conditions inside a waveguide. Waveguides have a cutoff frequency below which no signal can propagate. Therefore, a waveguide functions as a high pass filter. The mode with the lowest cutoff frequency is known as the dominant mode. This is the mode at which waveguides operate when transferring a microwave signal from one point to another. Additionally, to avoid the propagation of higher-order modes (i.e., other than the dominant mode) an upper frequency bound is associated with a waveguide. This upper bound is at the cutoff frequency of the first higher-order mode that can be generated. Waveguides are also dispersive which means the phase velocity becomes a frequency dependent parameter. Therefore, to avoid excessive dispersion, the operating frequency band of a waveguide is usually 1.25 times the cutoff frequency of the dominant mode and extends to the cutoff frequency of the first higher-order mode. For example, an *X*-band rectangular

*Agilent Technologies Test and Measurement Catalog, 2001, Palo Alto, CA.
† Millitech Millimeter Wave Products Catalog, 1995, South Deerfield, MA.

Copyright © Marcel Dekker, Inc. All rights reserved.

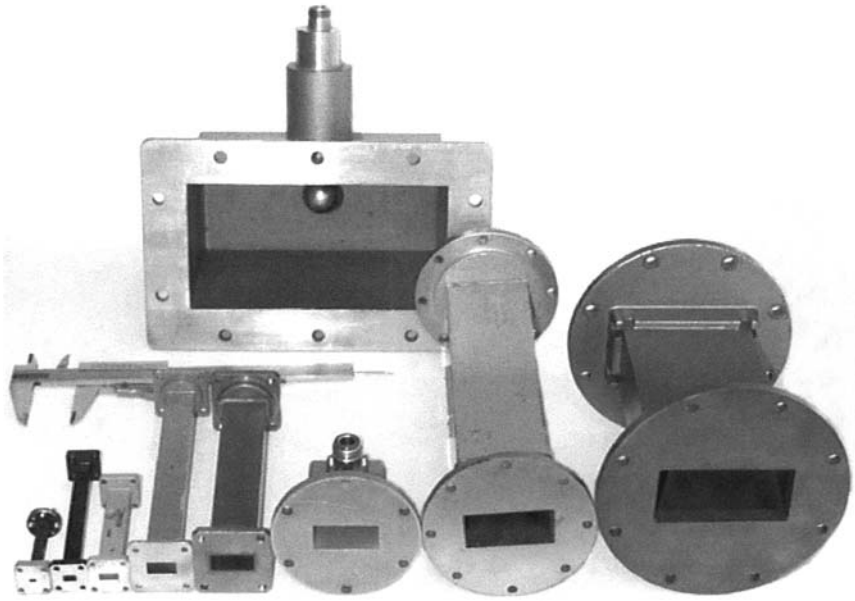


FIGURE 9.3 Rectangular waveguides with square and circular flanges covering a frequency range of about 1–75 GHz.

waveguide has a cutoff frequency of 6.557 GHz and an operating frequency range of 8.2–12.4 GHz. At this frequency band the waveguide cross-section is 2.286 cm by 1.016 cm. Waveguide sections can be attached to other sections or waveguide-based microwave components and devices using a flange. Figure 9.3 shows several rectangular waveguides (with square and circular flanges or coaxial adapters) covering a frequency range of about 1 GHz (the waveguide with the largest cross-section) to 75 GHz (the waveguide with the smallest cross-section). For reference, the caliper shows one inch.

Antennas. Antennas provide for a relatively efficient transfer of power from a transmission line to free space. They also shape or focus a microwave beam radiated from them. The degree by which this focusing is achieved and the efficiency by which the transfer of power takes place is a function of the geometry and the relative size of an antenna. One of the more commonly used microwave antennas is a horn antenna. A horn antenna is a waveguide whose walls have been flared out and extended. Figure 9.4 shows several horn antennas (pyramidal) and a conical horn antenna all fed by a rectangular waveguide. The antennas shown in Figure 9.4 cover a frequency range of about 8–75 GHz.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

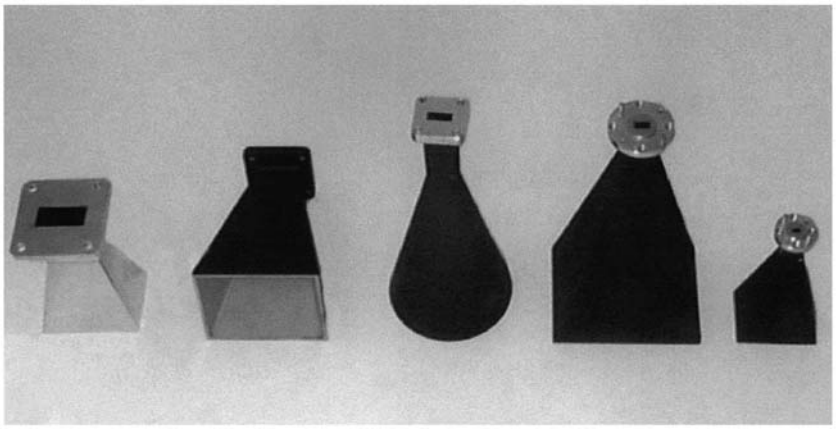


FIGURE 9.4 Horn antennas covering a frequency range of about 8–75 GHz.

Matched Loads. A matched load has the same impedance as that of the characteristic impedance of a transmission line. Therefore, when a transmission line is connected to a matched load, no reflection at the connection occurs, and all of the incident (on the matched load) signal is absorbed by the matched load and none is reflected back. In any microwave circuit no port of a device should be left open (i.e., unattached to a port of another component). Therefore, when such a situation arises, the unused port is connected to a matched load. Figure 9.5 shows the commonly used circuit symbol for a matched load.

Directional Couplers. Directional couplers are used to divide a microwave signal into two (or more) signals with different magnitudes and phases. Directional couplers may be constructed using stripline technology or by joining two waveguides using a common wall. When two directional couplers are used in tandem, it allows two signals traveling in opposite directions to be separated from one another (i.e., dual-directional coupler). One of the primary applications of directional couplers is to keep part of an incident signal as reference for comparing it with a reflected signal from an object under test, such as in mixer

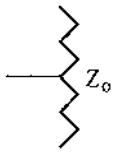


FIGURE 9.5 Microwave circuit schematic of a matched load.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

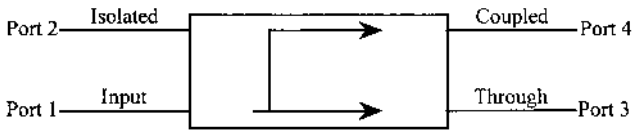


FIGURE 9.6 A microwave circuit schematic of a directional coupler.

configurations and in network analysis. Figure 9.6 shows a circuit symbol for a directional coupler. A signal fed into port 1 appears primarily at port 3 (known as the through port), a smaller portion of it appears at port 4 (known as the coupled port) and none appears at port 2 (known as the isolated port). The signal levels at the through and coupled ports are determined by the construction of a given directional coupler. If equal portions appear at ports 3 and 4 then the coupler is known as a 3-dB coupler (i.e., divides the input power at port 1 into two equal halves). The lengths of the transmission lines to ports 3 and 4 also determine the signal phase appearing at these ports. Thus, the signals at these two ports may have equal phases or may have a certain phase difference (e.g., 90°) with respect to each other. A matched load is usually connected to port 2 which prevents any reflections (or a signal appearing from port 1 at this port) to reenter the directional coupler. If port 2 is not connected to a matched load then a reflected signal from port 3 can appear at port 2. In this way one can monitor the properties of a reflected signal from port 3.

Circulators. In the microwave regime and inside a transmission line (or an antenna) two signals can travel in opposite directions. Within the transmission line they form a standing wave pattern (interference pattern). An example of a situation like this is when an incident and a reflected signal coexist in a transmission line. However, these two signals can be separated. Circulators have three ports and a signal entering a given port appears at only one of the two remaining ports and not the other. Figure 9.7 shows the circuit symbol for a

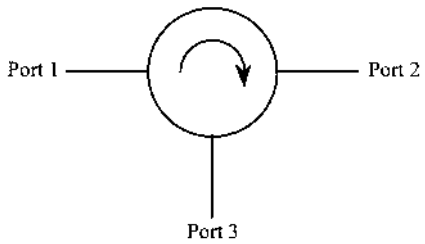


FIGURE 9.7 Microwave circuit schematic of a circulator.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

circulator. A signal fed into port 1 follows the arrow and appears *only* at port 2 and none appears at port 3 (isolated port). Likewise, a signal fed in port 2 appears at port 3 *only*, and a signal at port 3 appears at port 1 *only.*

Isolator. Isolators are used to prevent reflections or unwanted signals from appearing at a port in a microwave circuit. Therefore, isolators allow a signal to travel in one direction only. There are several approaches for designing an isolator. However, an isolator may be constructed from a circulator and a matched load, as shown in Figure 9.8a. A signal fed into port 1 appears at port 2. However, a signal fed into port 2 (such as a reflected signal) appears at port 3 only which is matched, and the signal is completely (in an ideal situation) absorbed by the matched load. Thus, an isolator is usually denoted by the circuit symbol shown in Figure 9.8b.

Detectors. A detector is a nonlinear device incorporating a solid state device such as a Schottky barrier diode. The V–I characteristic of the diode is used to perform rectification (amplitude demodulation) and signal detection. When operating in the latter mode, it is commonly referred to as a diode detector.

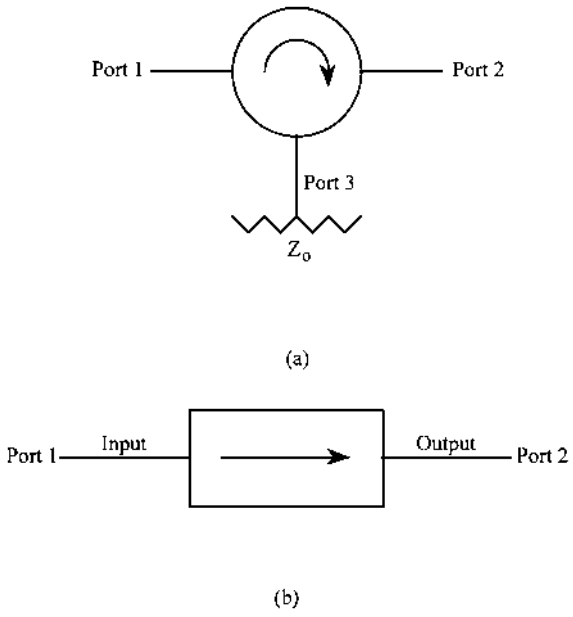


FIGURE 9.8 (a) Microwave circuit schematic of an isolator made of a circulator and a matched load and (b) common microwave circuit schematic of an isolator.

Copyright © Marcel Dekker, Inc. All rights reserved.

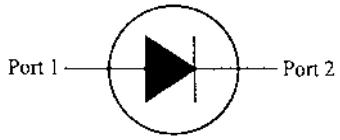


FIGURE 9.9 A microwave circuit schematic of a detector.

Essentially, such a detector produces a dc voltage output proportional to the input microwave signal power. Such detectors come in a variety of packages, operating frequency ranges and detection sensitivity. Figure 9.9 shows a common circuit symbol for a detector.

Mixers. Mixers incorporate a nonlinear solid state device, as well. Mixers are essentially an extension of a diode detector and are used for a variety of applications such as modulation, demodulation (i.e., up and down conversion) and phase detection. Figure 9.10 shows the circuit symbol for a mixer with its RF (radio frequency) port, LO (local oscillator) port and IF (intermediate frequency) port. The RF and LO signals are combined together (usually inside the mixer package) and appear at the input of the nonlinear device used. If the RF and LO frequencies are not the same, then the mixer output (IF port) will primarily consist of signals at frequencies equal to the addition of the two frequencies and the difference between the two frequencies (other frequency components, i.e., intermodulation components may also be present). Subsequently, by using a filter (usually inside the mixer package) the desired signal (usually the difference frequency component) is obtained. When the RF and LO signals are at the same frequency, then the IF signal will be a dc (zero frequency) signal proportional to the relative phase difference between the RF and the LO signals appear at the IF port. When operating in this mode, the mixer functions as a phase detector.

To demonstrate how one can assemble several of the above mentioned components and devices for a useful application consider a typical microwave circuit for a Doppler radar, as shown in Figure 9.11. The microwave oscillator

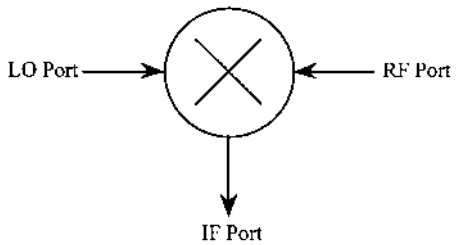


FIGURE 9.10 Microwave circuit schematic of a mixer.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

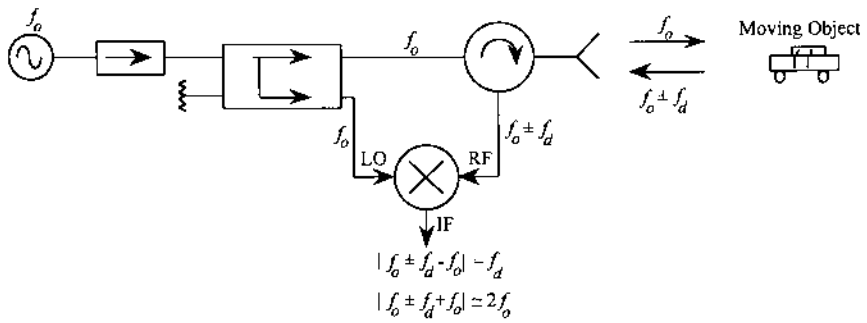


FIGURE 9.11 Microwave circuit diagram of a Doppler radar.

produces a CW signal at a frequency of f_0, and this signal is fed into a directional coupler through an isolator. The isolator is used to prevent any unwanted reflections from entering the oscillator. Such reflections may damage the oscillator or interfere with its proper operation. A portion of the signal appears at the coupled port of the directional coupler while the majority of the signal appears at the through port of the directional coupler. The latter signal then enters a circulator which reroutes it to the horn antenna. This signal is subsequently radiated by the horn antenna. This signal travels to a moving object (i.e., a car) and a portion of it gets reflected or scattered in the direction of the antenna. Since the object is moving the frequency of the reflected signal shifts up or down (depending on the relative direction of motion of the moving object, e.g., towards or away from the antenna) by an amount known as the Doppler shift, (f_d). The reflected signal is then picked up by the antenna and enters the circulator which reroutes it to the RF port of the mixer. At the mixer output (IF port) two (primary) signals will appear, one at a frequency of ($2f_o \pm f_d$) and one at a frequency of f_d. The Doppler frequency shift is directly related to the relative velocity between the antenna and the moving object. Therefore, to determine the relative velocity of the object, it is desired to extract f_d from the mixer output. The oscillator produces a signal in the GHz frequency range, while the Doppler frequency shift is in the hundreds of Hz or a few kHz range. Therefore, if the mixer output is fed into a low pass filter the Doppler frequency shift can be obtained, and hence the relative velocity can be easily calculated (and in some cases a speeding ticket may be issued!). With some simple modifications, this circuit may also be used for NDT purposes if one is interested in determining the phase of a reflected signal from an object under test. In this case there is no relative velocity between the object and the antenna. Thus, the output of the filter will be a dc voltage ($f_d = 0$) proportional to the relative phase difference between the signal at the LO port (reference signal) and the signal at the RF port (reflected signal). In this way an

Copyright © Marcel Dekker, Inc. All rights reserved.

indication of this relative phase difference can be obtained. Once calibrated, the output can be the absolute phase difference between the two signals which may then be related to the physical and dimensional properties of the object under test, as will be seen later.

9.4 MICROWAVE SENSORS/TECHNIQUES

Microwave NDE relies on the basic principle that electromagnetic fields are changed whenever they interact with a material structure. In general, the observed changes depend on both the material properties of the structure and its geometry, as well as on the frequency and polarization of the electromagnetic fields generated by the sensor. Various types of microwave sensors can be used, the choice depending on which material or geometric parameter (or parameters) is to be measured. Many times the parameter of interest is not measured directly but, rather, is inferred from the measurements. This process requires a model that relates the measured and desired quantities, and the best measurement is the one that produces the most accurate results using the simplest model. Most of the microwave sensors used in practice fall into one of five categories. These sensor types are briefly described in the following sections. The reader is referred to (20) for a more detailed discussion of industrial microwave sensors.

9.4.1 Transmission Sensors

The most straightforward sensor construction is the transmission sensor (Figure 9.12). It consists of a transmitter, a receiver, and usually a pair of horn antennas. The material to be measured is placed between the horns, which causes the amplitude and phase of the wave passing from the transmitter to the receiver to be changed. These changes depend on the material properties, the thickness of the material, the frequency, and the alignment of the material with respect to the horns. In measuring the dielectric constant of the material, uncertainties in the thickness of the material are a source of error. If the material is not a uniform solid sheet, constant layer thickness (and alignment) can be achieved with a stream-forming unit or by forcing the material through a dielectric tube between the antennas. Another source of error is the reflection at the interface between the material and the surrounding medium. This error can be reduced in part by calibration.

The advantages of transmission sensors are their simplicity and generality. Their major limitations are the need for a relatively large amount of material to achieve sufficient sensitivity and the need to reduce the influence of interface reflections.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

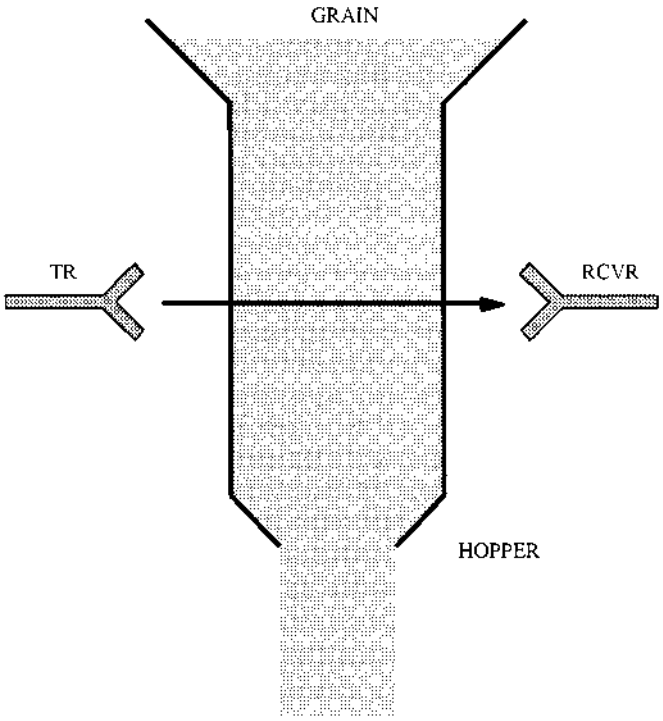


FIGURE 9.12 Microwave transmission sensor used to measure moisture content in grain. (From Ref. 24, Copyright 1998 © The American Society for Nondestructive Testing, Inc. Reprinted with permission from *TONE*: Volume 1, Sensing for Materials Characterization, Processing, and Manufacturing.)

9.4.2 Reflection and Radar Sensors

It is often more convenient to measure the signal reflected from an object rather than the transmitted signal (of course, transmission is not an option if the object is conducting (not penetrable by microwaves) or when one does not have access to both of its sides). Measuring material properties is possible by studying the magnitude and phase of the reflection coefficient, either with a contacting sensor or from a distance. An example of a contacting sensor is an open-ended coaxial transmission line (or waveguide) that has been configured as a resonator and which is pressed tightly against the object under test (Figure 9.13). Such sensors can also be used in a noncontacting fashion by carefully maintaining a small spacing between the sensor and the object under test, known as standoff distance or liftoff. Another example of a noncontacting sensor is a single broadband

Copyright © Marcel Dekker, Inc. All rights reserved.

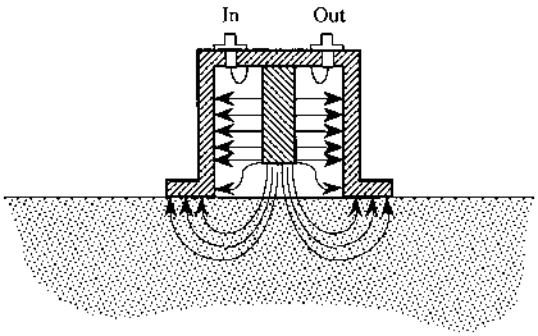


FIGURE 9.13 Coaxial resonator sensor used to sense near-surface defects. The electric fields are shown penetrating the surface to a depth determined by the cross-sectional size of the sensor. (From Ref. 20; reprinted with permission.)

microwave horn antenna that is located a relatively large distance from the object under test. Such sensors typically use short-pulse or swept-frequency signals, and we generally refer to them as radar sensors. They can measure the delay time, amplitude, phase, or change in frequency of the reflected signal. Horn antennas can also be used with continuous-wave signals at close distances by using bridge (nulling) techniques.

A typical application of reflection sensors is to evaluate the thicknesses and material properties of the layers in laminated materials by studying the reflection coefficient as a function of frequency or angle of incidence. An application example of radar sensors is in tankers and other places where the liquid is flammable or the surface may be covered by foam (Figure 9.14).

9.4.3 Resonator Sensors

Resonators compose the third major class of microwave sensors. Typically, such resonators are formed when reflecting discontinuities are introduced in a transmission line or waveguide so that a wave can bounce back and forth many times between these discontinuities before losing its energy (Figure 9.13). At certain excitation frequencies, the oppositely traveling waves will combine to form a standing-wave pattern. These resonant frequencies depend on the size of the sensor in wavelengths and are therefore affected by the dielectric constant or permeability of the object under test, which is located such that it can interact with the waves traversing back and forth in the resonator. Since the object interacts with each wave many times before the wave decays to $1/e$ of its initial

Copyright © Marcel Dekker, Inc. All rights reserved.

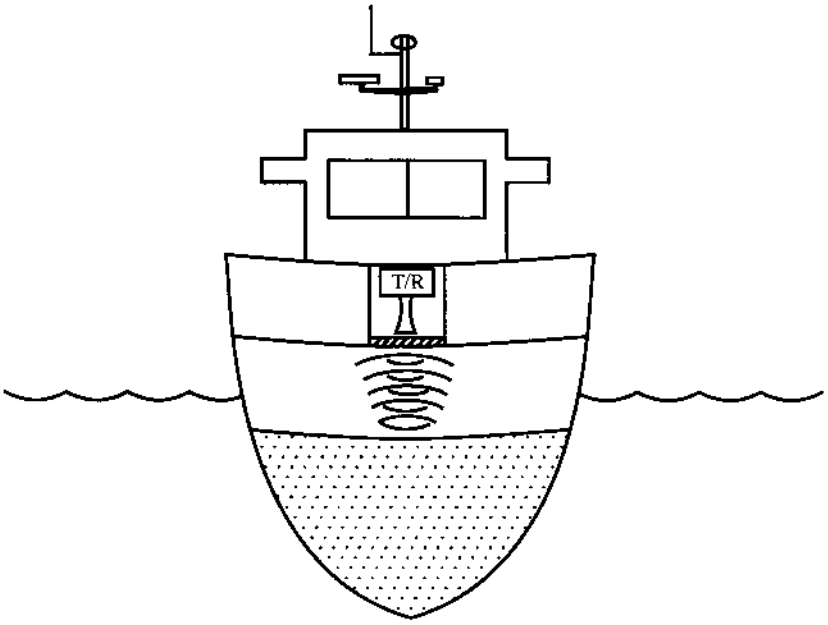


FIGURE 9.14 A fuel-level sensing radar on a ship. Placing the radar behind a hermetic dielectric window eliminates the danger otherwise caused by electronic equipment in contact with explosive gases. (From Ref. 20; reprinted with permission.)

value (Q/π times), the effect of a small or low-loss object is magnified. Hence, such sensors offer high measurement sensitivity.

A resonator sensor can be constructed in many ways, so it can be designed to fit the particular requirements of an application, depending on the size, shape, and material properties of the object under test. Therefore, the resonator measurement principle is quite versatile, but resonator sensors tend to be specialized and can seldom be used directly for applications other than those for which they are intended. They are best suited for the measurement of small, thin, or low-loss objects, and for surfaces of large objects.

9.4.4 Radiometer Sensors

A microwave radiometer sensor is an antenna (and an associated sensitive receiver) that is sensitive enough to detect the black-body thermal radiation from an object in a microwave frequency band (Figure 9.15). It is called a *passive* sensor (the other sensors discussed so far are *active* in that they rely on the use of

Copyright © Marcel Dekker, Inc. All rights reserved.

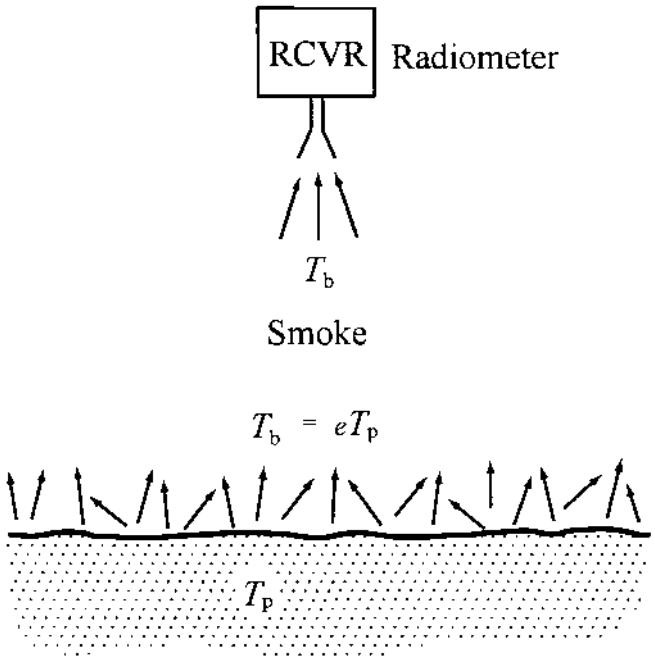


FIGURE 9.15 A microwave radiometer measures thermal noise radiated by an object at temperature T_p (e is the emissivity of the object and T_b is the brightness temperature). Microwaves are unaffected, for example, by smoke between the object and the antenna. (From Ref. 20; reprinted with permission.)

an external source to illuminate the object under test) because it "listens" to the noise radiation transmitted by an object. This noise radiation is dependent on the physical temperature of the object and its emissivity. Emissivity is another word for the power transmission coefficient of the object's surface and therefore depends on both surface roughness and the characteristic wave impedance, $\eta = \sqrt{\mu/\varepsilon}$, of the object. Radiometers can be used to produce thermographical images of objects, either by turning the antenna in different directions or by electronically scanning the antenna beam using a phased array. Another possibility is to use an array of near-field probe-like antennas which scan in raster fashion over the surface of the object.

In industry, radiometers can be used for remote temperature measurements, as, for example, in ovens, kilns, and other places where use of conventional contacting temperature sensors or infrared radiometers is impossible because of high temperature, smoke, or water vapor. Another field of potential application is

Copyright © Marcel Dekker, Inc. All rights reserved.

in medicine. Research in this area has concentrated on the detection of hot spots that might indicate cancerous tumors or other abnormalities (21). Microwave thermography has also been used to monitor changes in temperature during microwave hyperthermia, e.g., in the treatment of tumors.

9.4.5 Imaging Sensors

A special group of sensors are those which are used for imaging the surface or interior of objects, or for detecting concealed objects. These active imagers are divided into holographic sensors and tomographic sensors, depending on whether they measure reflected or transmitted radiation, respectively. Holographic sensors are similar to radars, but they differ in that the reflection is measured from several directions and both amplitude and phase are measured. Given such data, one can reconstruct a source image using holographic back-projection algorithms (22). In microwave tomography, on the other hand, the object is illuminated with one antenna and the resulting complex scattered field is measured in all directions on the other side (Figure 9.16). Complicated tomographic algorithms exist for reconstructing an image of the object (23), but work to develop NDE applications has been limited.

9.4.6 General Remarks on Sensors

Clearly, a wide variety of microwave sensors exist, which facilitates their tailoring to the application at hand. These sensors allow one to measure the amplitude and/or phase of the electromagnetic wave fields reflected by and/or transmitted through an object under test. These measured quantities are often quantitatively related to material properties or object geometry through the use of models. In other cases one is only interested in monitoring relative changes in object properties and therefore models are not required. For example, noncontacting

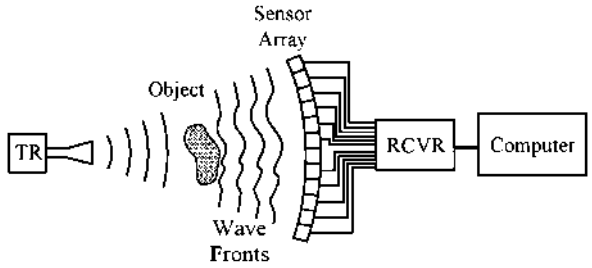


FIGURE 9.16 Microwave tomography using a sensor array. (From Ref. 20; reprinted with permission.)

Copyright © Marcel Dekker, Inc. All rights reserved.

microwave sensors are often used for process monitoring in an industrial environment (24).

Another facet of the versatility of microwave sensors is that they can sometimes be designed to exploit some particular scattering feature of an anomaly in the object under test. An example of such special behavior is electromagnetic mode conversion by surface cracks in metals or knots in timber. In this case, the microwave sensor can take advantage of the different properties of the scattered field (i.e., their modal properties) to discriminate against superfluous background signals and improve the dynamic range of the measurement.

9.5 APPLICATIONS

As outlined in the Introduction, there are a great many realized and potential applications for microwave NDE. In this section we discuss a few of these applications in more detail in order to better illustrate and explain the techniques involved, and also to demonstrate specific uses.

9.5.1 Dielectric Material Characterization Using Filled Waveguides*

One of the most common applications in microwave NDE is measuring the dielectric properties of solids, powders, and liquids using a completely-filled short-circuited waveguide. The adjective "completely-filled" refers to the situation where the dielectric sample fills up the cross-section of the waveguide. As was discussed earlier, a waveguide is a hollow cylindrical pipe whose cross-section is typically rectangular or circular (other cross-sections are also possible, but are less common). The useful bandwidth of a waveguide depends on its cross-sectional dimensions. For example, the operating frequency range for an X-band waveguide (22.86 mm $\times$ 10.16 mm) is 8.2–12.4 GHz.

First, we present the results of using such a technique for measuring the dielectric properties of carbon-loaded rubber and its constituents (25). These results also demonstrate the potential of using microwave dielectric characterization for detecting the presence of curatives in this material. Second, we discuss using this technique to determine distributed porosity in polymer plastics (26).

Although there are a number of well-established techniques available for measuring the dielectric constant of materials (27), the filled-waveguide tech-

*Portions reprinted, with permission, from IEEE Transactions on Microwave Theory and Techniques; Vol. MTT-42, No. 1, pp. 18–24, January 1994 (25) (Copyright © 1994 IEEE), and from Materials Evaluation; Vol. 53, pp. 404–408, March 1995 (26) (Copyright © 1995 by The American Society for Nondestructive Testing, Inc.).

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

nique is an attractive approach since it possesses the following desirable characteristics:

1. It is able to measure low- and high-loss dielectric materials with good accuracy.
2. It facilitates varying the sample thickness to achieve improved measurement accuracy.
3. It permits measurements to be performed in a broad range of frequency bands.
4. It is inexpensive and simple to use, only requiring microwave and electronic equipment that is readily available, making it easy to reproduce.
5. It is readily adaptable to measuring powders and fluids, as well as solids.

Measurement Procedure

Of the many microwave approaches for measuring the dielectric properties of materials, a dual-arm waveguide bridge (28) or a partially filled waveguide (29) are potential candidates. However, the former technique suffers from requiring a relatively large number of microwave components that must possess very good performance characteristics, and the latter technique is only suitable for sheet materials. The well known completely-filled short-circuited waveguide technique (30–31) is a better choice because it provides all of the desired measurement characteristics mentioned above. This method is based on measuring the complex reflection coefficient, Γ, whose magnitude and phase are then used to extract the dielectric constant of the sample. Hence, the completely-filled short-circuited-waveguide is an example of a reflection sensor.

Figure 9.17 shows a photograph of an apparatus for measuring the dielectric properties of rubber samples and their constituents (25). Five different waveguide bands were used in obtaining the data to be presented: 3.95–5.85 GHz, 5.85–8.2 GHz, 8.2–12.4 GHz, 12.4–18 GHz, and 18–26.5 GHz (the standard designations for these waveguides are WR-187, WR-137, WR-90, WR-62, and WR-42, respectively). The oscillator (not shown) generates a microwave signal at the desired frequency, which is then fed through a precision attenuator and a slotted waveguide. Usually, an isolator (not shown) is placed in between the adapter and the attenuator. The standing-wave characteristics inside the waveguide (which are related to the complex reflection coefficient) were measured using a sliding electric-field probe connected to a detector and a sensitive voltmeter. This type of slotted-waveguide equipment is commonly found in many undergraduate electromagnetics (high-frequency) laboratories. Precisely machined sample holders connected to the slotted waveguide were used to accommodate solid samples of different thicknesses (dimensional precision is

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

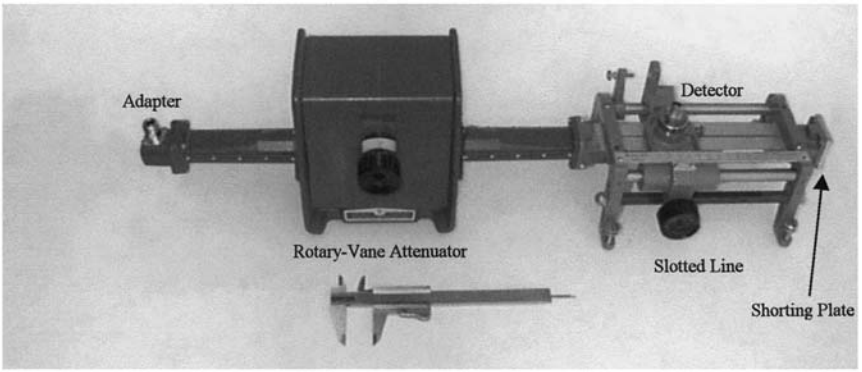


FIGURE 9.17 Picture of a completely-filled short-circuited slotted waveguide apparatus for dielectric property measurement (photograph reprinted with permission of Agilent Technologies, Inc.).

important in order to minimize air gaps between the samples and the waveguide walls). These sample holders were subsequently terminated by a shorting plate. For liquids and powders, the vertical sample holder was filled completely and a piece of clear (microwave transparent) tape was used to hold the materials in place.

As mentioned earlier, the standing-wave ratio is measured by sliding the detector along the slotted waveguide and recording the maximum and minimum voltages. The ratio of these two voltages gives the standing-wave ratio. In doing so, the operating point of the diode detector (on its V-I characteristic curve) is not the same when measuring the maximum and the minimum voltages. This causes measurement error associated with determining SWR. However, using a rotary-vane precision attenuator allows keeping this operating point constant while recording SWR (26,29–31).

The relationship between the complex reflection coefficient, Γ, and the complex relative dielectric constant of the sample, $\varepsilon_r = \varepsilon_r' - j\varepsilon_r''$, is given by

$$\varepsilon_r = \frac{\left(\frac{X}{kL}\right)^2 + \left(\frac{\lambda_g}{2a}\right)^2}{1 + \left(\frac{\lambda_g}{2a}\right)} \tag{30}$$

where X is a solution of the following transcendental equation

$$\frac{\tan X}{X} = \frac{1 - \Gamma}{jkL(1 + \Gamma)} \tag{31}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

and k (given by $2\pi/\lambda_g$) is the wavenumber inside the waveguide. The parameter λ_g is the wavelength inside the waveguide, L is the sample thickness, and a is the broad dimension of the waveguide. It should be noted again that, since a waveguide is made from one contiguous conductor (as opposed to a coaxial transmission line which has two separate conductors), a waveguide cannot support TEM waves. Rather, its propagating modes are classified as transverse electric (TE) and transverse magnetic (TM). These modes have components of magnetic field and electric field along the waveguide axis, respectively, whereas TEM waves do not have any electromagnetic field components along the propagation axis. However, TE and TM waves in a waveguide can be thought of as TEM waves that are propagating at an angle to the waveguide axis and bouncing back and forth off the waveguide walls. Thus, the wavelength of TE and TM waves measured along the waveguide axis, λ_g, is larger than the *free-space wavelength*, λ_0. For example, a 10-GHz signal would have $\lambda_0 = 3$ cm, whereas λ_g would be almost 4 cm.

It is known that Eq. (30) has an infinite set of complex roots (31). Thus, to find the correct root, either one must have an estimate of the dielectric constant of the material filling the waveguide, or else one must make two or more measurements using different sample thicknesses. The latter procedure was followed in this example.

The reflection coefficient, as a function of dielectric constant and sample thickness, can be expressed as

$$\Gamma(\varepsilon_r, L) = \frac{\sqrt{Y} - j\tanh(kL\sqrt{Y})}{\sqrt{Y} + j\tanh(kL\sqrt{Y})} \tag{32}$$

where Y is given by

$$Y = \left[1 + \left(\frac{\lambda_g}{2a}\right)^2\right]\varepsilon_r - \left(\frac{\lambda_g}{2a}\right)^2 \tag{33}$$

The two measured parameters are the SWR, and the *position of a standing-wave minimum*, $z_{\min}$, in the slotted waveguide. The magnitude of the reflection coefficient is related to SWR by Eq. (15), and the *phase of the reflection coefficient*, ϕ is given by

$$\phi = m\pi - 2kz_{\min} \tag{34}$$

where m is an odd integer that can be determined from the known waveguide dimensions and the frequency. Thus, these measurements lead to an array of values, $\Gamma(\varepsilon_r, L_n)$, where n denotes the number of samples with different thicknesses. The unknown dielectric constant, ε_r, is then determined by finding the best fit to Eq. (32).

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

To illustrate a typical set of measurement data, the null position and standing-wave ratio for two rubber samples (whose specific properties will be mentioned later) were measured using this technique and the results are shown in Figures. 9.18 and 9.19. The discrete points shown in these figures represent typical data obtained at 5 GHz for two rubber samples having various thicknesses. These values were used to calculate the dielectric constant of the rubber samples using a best fit to Eq. (32). The lines in the figures represent the results obtained from Eq. (32) using the calculated value of the dielectric constant. It can be seen that the quality of the fit is quite good. It is important to note that using more than one sample thickness yields a more accurate estimate of the dielectric constant. It is also advantageous to use sample thicknesses that correspond to the neighborhood around the standing-wave phase transition (minimum SWR) that is apparent in Figures. 9.18 and 9.19. The thickness corresponding to this phase transition can be calculated either from prior approximate knowledge of the dielectric constant, or from preliminary measurements using random thicknesses.

Measurement Accuracy

The accuracy of this measurement technique is well established (32–33). The accuracy and sensitivity can be improved by measuring multiple sample thicknesses around the phase transition region and by direct fitting of the measurement results to obtain the dielectric constant. This improved accuracy allows one, for

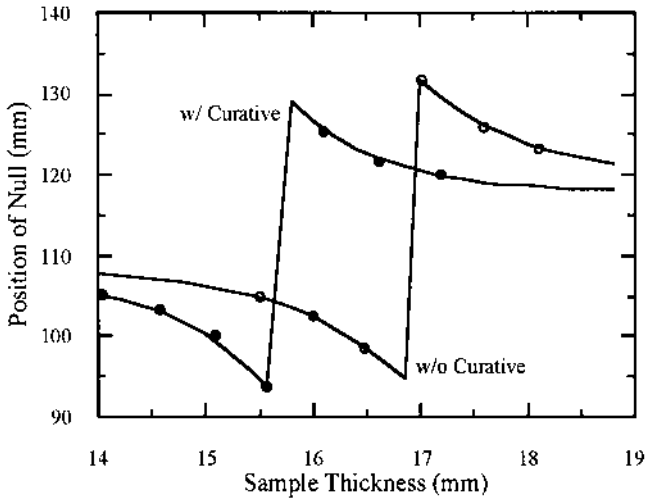


FIGURE 9.18 Null position in the slotted waveguide at 5 GHz as a function of sample thickness. (From Ref. 25; reprinted with permission.)

Copyright © Marcel Dekker, Inc. All rights reserved.

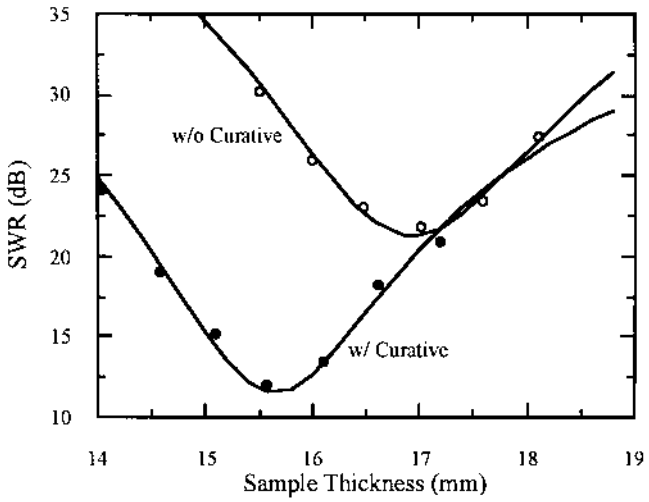


FIGURE 9.19 SWR at 5 GHz as a function of sample thickness. (From Ref. 25; reprinted with permission.)

example, to consistently detect the weak effects of curatives in uncured rubber samples, particularly when these materials have low losses (25).

A commonly used set of measurement parameters consists of the SWR and null position for various sample thicknesses. The accuracy of such measurements is highest if one uses a rotary-vane attenuator and a precision slotted line (see Figure 9.17). If the SWR is large, the following correction factor can be applied (SWR expressed in dB):

$$(\mathrm{SWR}_c)_{\mathrm{dB}} = -20\log_{10}\left[10^{\frac{-(\mathrm{SWR})_{\mathrm{dB}}}{20}} - 10^{\frac{-(\mathrm{SWR}_0)_{\mathrm{dB}}}{20}}\right] \tag{35}$$

where a lossless waveguide and a perfect short-circuit termination have been assumed. Here, the subscript c on SWR, indicates the corrected value, no subscript indicates the measured value, and the subscript 0 indicates the value for the short-circuited waveguide without any material present, as shown in Fig. 9.17. This equation shows that there is a limitation in measuring low-loss materials when SWR is nearly equal to SWR_0. In this case the accuracy can be improved by using thicker samples.

Special considerations apply when measuring low- or high-loss materials. For low-loss materials, the real part of the dielectric constant can be determined precisely from the multiple-thickness null-location information, but the imaginary part cannot be deduced with reasonable accuracy from these data. Thus, for this case the SWR measurement is very important, as it (SWR) is the sole indication

Copyright © Marcel Dekker, Inc. All rights reserved.

of losses. For high-loss materials, a different problem exists. In this case, the null-location curve quickly degenerates to a line with a certain slope as a function of sample thickness. Thus, the precise value of the dielectric constant cannot be extracted reliably from the null-location measurements alone and SWR measurements are also needed. Measurements should be performed close to the first SWR minimum, because beyond that the SWR oscillates negligibly as a function of sample thickness (due to the losses in the material).

It is possible to estimate the measurement uncertainty associated with the process described above by assuming typical measurement errors and calculating the resulting total error in the derived dielectric constant. For example, suppose we assume the following measurement errors: a measurement resolution of the rotary-vane attenuator equal to ± 0.25 dB, a position error of the indicated standing-wave minimum equal to ± 0.05 mm, and a sample thickness error of ± 0.1 mm. The error due to frequency instability can be neglected since modern oscillators are very stable. The dielectric constant, and subsequently the percentage differences for the real part of the dielectric constant and the loss tangent ($\tan\delta = \varepsilon_r''/\varepsilon_r'$), can be calculated using these assumed uncertainties, resulting in the maximum error (worst case) for a given measurement. In this case, the errors for all the rubber samples discussed below were calculated to be less than 1% for the real part of dielectric constant and less than 3.5% for the loss tangent. Note again that the accuracy degrades if measurement points are not in the vicinity of the phase transition and an insufficient number of samples is used.

Dielectric Characterization of Rubber Compounds and Their Constituents

A representative and important example of microwave dielectric characterization is contained in a published study of the microwave behavior of the dielectric properties of various rubber stocks (25). This type of information could be important in such practical applications as controlling rubber mixing processes and in the development of new materials with specified properties (e.g., microwave absorbers). In this study, the dielectric properties of rubber constituents such as ethylene-propylene-diene monomer (EPDM) elastomer, mineral filler, oil, zinc oxide, and curatives were first measured. Then, the influence of carbon-black volume percentage on the dielectric properties of rubber was investigated. The ability of microwaves to detect the presence of curatives in uncured rubber and the role of frequency in this detection was also demonstrated. These results indicate that microwave material-characterization techniques may also be used for cure monitoring, which is an important application.

Rubber-Compound Constituents. To fully understand the dielectric behavior of various rubber compounds, one must first have a complete knowledge of the dielectric characteristics of their individual constituents, namely EPDM,

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

TABLE 9.2 Volume Content (%) and measured X-band Dielectric Properties for Common Rubber-Compound Constituents

Constituent	Volume %	ε'_r	tan δ
EPDM	45.5	2.1	0.008
Oil	25.85	2.2	0.014
Carbon black	17.0	–	–
Mineral filler	9.3	2.3	0.017
Curatives	1.8	3.1	0.032
Zinc oxide	0.36	3.9	0.28

Source: Ref. 25; reprinted with permission.

mineral filler, zinc oxide, oil and curatives. Table 9.2 lists the volume content (by percentage) and measured dielectric properties of these constituents for the typical rubber compound used in this example. This table shows the mean values of ε'_r and loss tangent of the constituents in X-band (8.2–12.4 GHz), using the filled-waveguide technique described previously. Measurements in other frequency bands produced similar values. Due to the very high conductivity of carbon black, its dielectric properties were not measured. The accuracy of these measurements (listed in Table 9.2) is only about 5% for ε'_r and 10% for the loss tangent, since only two arbitrary thicknesses were used to produce this particular set of measurements.

The results indicate that the dielectric constants of EPDM, oil and mineral filler are, for all practical purposes, equal. Zinc oxide and curatives have slightly higher dielectric constants, but they occupy small volume percentages in the total composition of a typical rubber compound. If only physical mixing is involved (not considering any chemical reaction), the contribution of curatives to the rubber-compound dielectric constant is expected to be negligible. However, it has been observed that the chemical reaction triggered by curatives in uncured rubber samples does cause a detectable change in the dielectric properties of rubber even at ambient temperature. This important phenomenon will be discussed further later on in the section on curative detection.

Cured Rubber. The effects of frequency and carbon-black volume content (%) on the dielectric properties of cured rubber were also determined using the filled-waveguide technique. Carbon black, depending on the application of the rubber (i.e., hoses, belts or tires) adds strength and increases its DC conductivity. Measurements were performed on 14 rubber samples having varying degrees of carbon-black concentration. The carbon blacks used for this study were aggregates of particles ranging in diameter from 5–125 nanometers.

Copyright © Marcel Dekker, Inc. All rights reserved.

The rubber samples were prepared starting with the basic formulation given in Table 9.2 and then the volume percentages of carbon black were varied between 8.4%–35.6%. This was done in such a way that any increase in the amount of carbon-black percentage was compensated by removing an equal amount of EPDM and keeping the percentages of all other constituents constant (same percentages as shown in Table 9.2), the goal being to isolate the influence of carbon black on rubber-compound dielectric constant. Table 9.3 lists the carbon-black volume content for the measured rubber samples (sample 5 is the reference sample). These samples were prepared in sheet form with thicknesses of 0.5, 1, 2, 3 and 5 mm and their dielectric constants were measured at 5, 7, 10, 16, and 24 GHz.

Figures 9.20 and 9.21 show the results of these measurements for the real and imaginary parts of the relative dielectric constant, ε_r' and ε_r'', respectively, at 5 and 24 GHz (the results for other frequencies fall in between these two extremes). These results show that the dielectric constant of rubber increases as the carbon-black volume percentage increases. This is to be expected due to the relatively high conductivity of carbon black. Clearly, dielectric constant variations are more pronounced at 5 GHz (larger slope), which makes this frequency (or lower) a better choice for measuring dielectric constant variations due to carbon black.

It has been reported (34) that the rolling effect associated with the production of rubber sheets causes anisotropy in the dielectric properties of the material. To minimize this effect, the samples used in this example were prepared

TABLE 9.3 Carbon-Black Volume Percentages for the Measured Samples

Sample No.	Carbon Black Volume %
1	8.40
2	10.13
3	12.13
4	14.41
5 (reference)	17.00
6	18.02
7	19.07
8	20.14
9	21.23
10	22.34
11	25.29
12	28.58
13	32.05
14	35.61

Source: Ref. 25; reprinted with permission.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

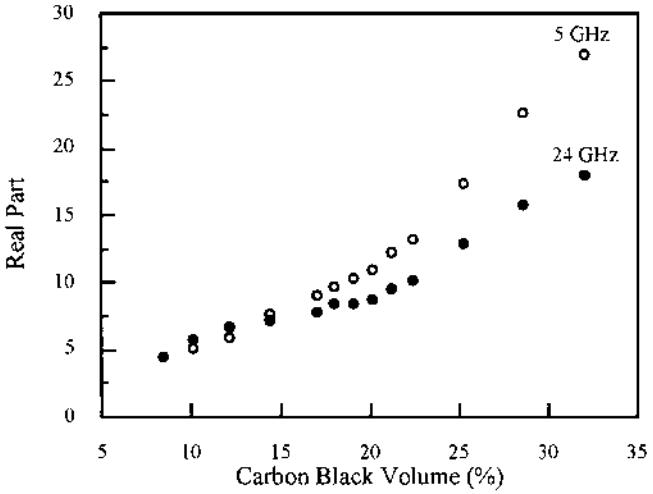


FIGURE 9.20 Real part of the dielectric constant (ε_r') for all cured rubber compounds at 5 and 24 GHz. (From Ref. 25; reprinted with permission.)

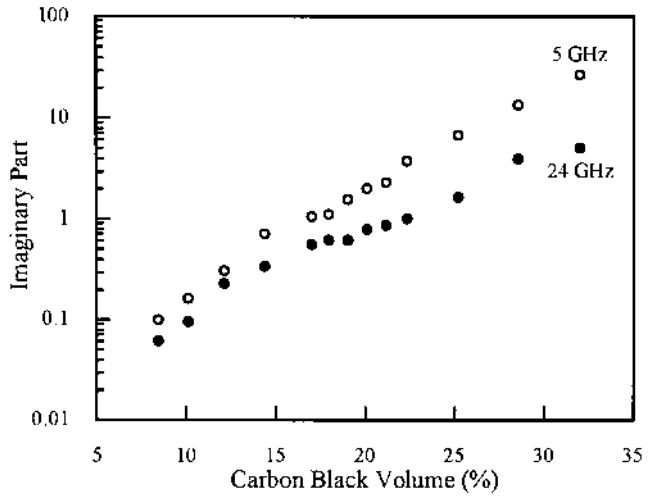


FIGURE 9.21 Imaginary part of the dielectric constant (ε_r'') for all cured rubber compounds at 5 and 24 GHz. (From Ref.25; reprinted with permission.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

by performing the rolling process in many different directions. To check the success of this procedure, the dielectric properties of several rubber samples arbitrarily oriented with respect to the electric field vector in the waveguide were also measured. The results indicated that there were no systematic deviations in dielectric constant with orientation (within the measurement errors), which confirms the absence of dielectric anisotropy.

Curative Detection. The presence of curatives was also checked in rubber loaded with carbon black. The goal of these experiments was to investigate the possibility of using microwave dielectric characterization to detect the presence (or absence) of these curatives prior to the product-producing stage and curing. All of these measurements were performed at ambient temperature.

Uncured versions of samples no. 4, 5, 7, 9, 11, and 14 (refer to Table 9.3) were prepared with and without curatives. These samples were chosen to represent the entire range of carbon-black percentages used in practical applications. Because the differences in dielectric constant were expected to be small, sample thicknesses were chosen near the phase-transition values (determined by experimentation) in order to maximize these differences.

Figures 9.22 and 9.23 show the real and imaginary parts of the relative dielectric constant, ε_r' and ε_r'', respectively, of the uncured samples at 5 GHz, with and without curative and as functions of carbon-black volume percentage. From Figure 9.22 it is apparent that the addition of curatives does not affect the real part of the dielectric constant in a uniform fashion for all of the carbon-black volume percentages tested. However, in spite of the low losses exhibited by curatives themselves (see Table 9.2, and compared to carbon-black-loaded rubber) the imaginary part of the dielectric constant of these samples tends to increase with the addition of curatives for all carbon-black volume percentages. Furthermore, the differences in ε_r'' caused by curatives are readily detected, as is apparent in Figure 9.23. This relatively large effect is thought to be due to a chemical reaction triggered by the curatives at room temperature. It is well known in the rubber industry that a rubber compound containing curatives will cross-link (cure) in due time even at room temperature. It is apparent from these data that the initial stage of curing can be detected using microwave measurements. Another observation is that, after the carbon-black percentage increases beyond 25%, the difference between the samples with and without curatives becomes smaller. It may be hypothesized that, at these high carbon-black percentages, the effect of the chemical reaction associated with curatives is masked by the overwhelming presence of unlinked, conductive, carbon black.

Concluding Remarks The example discussed above shows that the completely-filled waveguide approach is very useful for conducting extensive dielectric constant evaluations. The measurement apparatus is relatively simple,

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

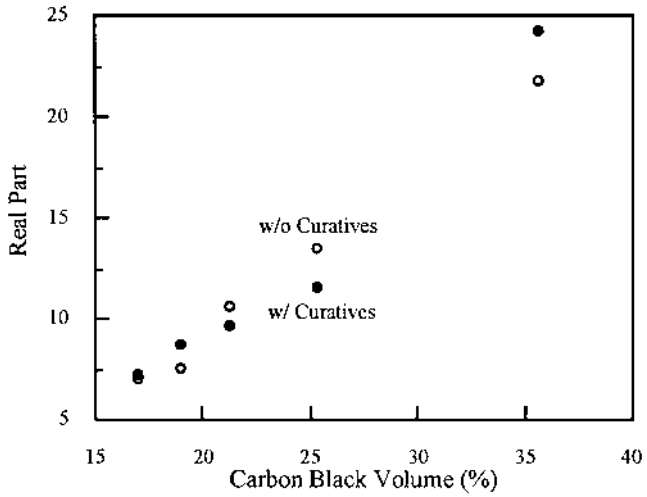


FIGURE 9.22 Real part of the dielectric constant (ε_r') at 5 GHz for uncured rubber compounds with and without the addition of curative. (From Ref. 25; reprinted with permission.)

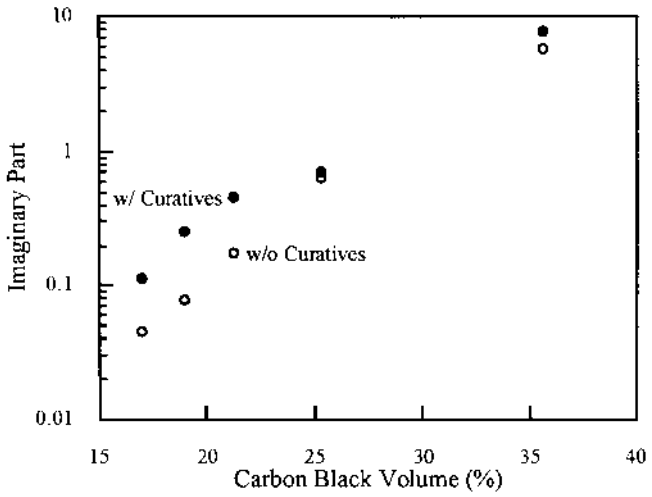


FIGURE 9.23 Imaginary part of the dielectric constant (ε_r'') at 5 GHz for uncured rubber compounds with and without the addition of curative. (From Ref. 25; reprinted with permission.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

inexpensive, and is realizable for a large range of frequency bands. Moreover, this technique offers very good repeatability and measurement accuracy, particularly after some refinements of the measurement procedure.

Other chemically reactive materials besides rubber have also been examined for curing effects using microwave dielectric characterization. Among them are resin binders and concrete-based materials (35–38). Potential applications include determining cure state and temporal behavior during the curing cycle.

Porosity Evaluation in Polymer Composites

Determining porosity level (air content) in cured polymers, ceramics, and composite materials is another important practical issue. In cured polymers, the presence of porosity reduces mechanical performance as a result of stress concentrations. Also, localized porosity can be particularly damaging to the joint strength of adhesively bonded components. In ceramics, the relative density is an important processing parameter, and again the ceramic is extremely sensitive to stress concentration (lowered density). If not fully densified, a ceramic will be weak and have low stiffness. In composites, porosity can occur within the matrix material, which will affect the mechanical performance in the same way it does in bulk materials. In addition, porosity often concentrates at specific locations in composite materials (either between plies or at the fiber/matrix interface) and can dramatically lower flexural and shear performance. Increases in porosity during use (i.e., when the material is loaded by applied forces) may precede macroscopic damage and possibly indicate the presence of delamination. Hence, a technique capable of detecting and accurately determining porosity levels in materials is very desirable.

The capability of microwave techniques for detecting and evaluating porosity has been demonstrated (26). In this study, samples of polymer-microballoon-filled epoxy resin with 0%, 48.9%, 58.7% and 68.5% air-volume fractions were carefully prepared to simulate various porosity levels. The size of the uniformly distributed air-filled microballoons ranged between 15–200 micrometers in diameter, with a bulk density of 0.009 gm/cm^3. The samples were shaped to fit tightly inside rectangular waveguides and the measurements were conducted at frequencies of 8.2, 10, 12, 14, 16, and 18 GHz, again using the completely-filled short-circuited waveguide technique described previously. The thicknesses of these samples were reduced as the experiments proceeded to provide the multiple sample thicknesses required for obtaining good accuracy in these measurements. A major goal of this study was to provide a basis for extrapolating its results to the use of microwave techniques in evaluating more complicated materials such as polymer matrix composites.

Tables 9.4 and 9.5 show the results for the relative permittivity, ε_r', and loss factor, ε_r'', respectively. We see that there is relatively little variation as a function of frequency within the measurement accuracy, which is expected since these

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

TABLE 9.4 ε'_r for Polymer-Microballoon-Filled Epoxy.

	Air Content			
Frequency (GHz)	*0%*	*48.9%*	*58.7%*	*68.5%*
8.2	2.80	1.87	1.69	1.48
10	2.87	1.84	1.63	1.46
12	2.83	1.88	1.70	1.47
14	2.87	1.83	1.70	1.50
16	2.84	1.82	1.58	1.47
18	2.84	1.84	1.67	1.47

Source: Ref. 26; reprinted with permission.

samples have low dielectric constants and loss factors. However, there is a clear distinction among different air-volume contents. The differences in porosity between the samples produces a nearly linear change in the values of ε'_r and ε''_r. However, one must be careful not to generalize this trend to composites whose dielectric properties differ from those used here. For example, in ceramics with higher relative permittivities ($\varepsilon'_r = 6$–20) one would expect a greater sensitivity in dielectric properties to changes in porosity level.

Based on the measurement results obtained at 10 GHz, a 1% change in porosity level translates to about a 1.1% change in ε'_r and a 3% change in ε''_r. Conversely, this means that, if the relative permittivity and the loss factor can be measured to within these percentages, then it should be possible to detect a 1% change in porosity for this dielectric material. Additional measurements and analyses indicate that it is indeed reasonable to expect the ability to detect porosity variations of 1–2% in materials with similar dielectric constants. This conclusion is essentially independent of frequency. These encouraging results

TABLE 9.5 ε''_r for Polymer-Microballoon-Filled Epoxy

	Air Content			
Frequency (GHz)	*0%*	*48.9%*	*58.7%*	*68.5%*
8.2	0.086	0.032	0.020	0.013
10	0.086	0.034	0.023	0.015
12	0.082	0.033	0.022	0.014
14	0.083	0.033	0.027	0.022
16	0.077	0.026	0.021	0.019
18	0.068	0.027	0.025	0.014

Source: Ref. 26; reprinted with permission.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

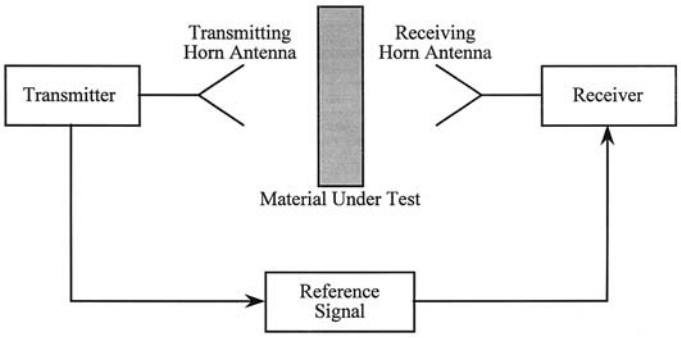


FIGURE 9.24 Experimental arrangement for microwave transmission.

suggest that microwave dielectric-property characterization can be used for studying porosity variations in other plastics, ceramics, and composite mixtures.

It should be mentioned that, for low-permittivity and low-loss dielectric materials, the cavity-resonator technique may provide better accuracies than the filled-waveguide technique. However, a drawback of the cavity technique is that meticulous care must be taken in preparing the samples (usually sphere-shaped) for insertion into the cavity (39).

9.5.2 Dielectric Material Characterization Using Transmission Measurements*

The previous sections discussed the use of a reflection sensor to measure dielectric material properties. Now we consider another method, namely, the transmission method. In this method, a sample is placed between a transmitter and a receiver, and the dielectric properties of the sample are calculated from measurements of the signal that passes through the sample. The sample can either be enclosed in a waveguide, or it can be located in free space. The free-space case is discussed here.

Measurement Procedure

The microwave transmission method is particularly useful for measuring the dielectric properties and thicknesses of dielectric slabs (40–41). A typical experimental arrangement is shown in Figure 9.24. The transmitter provides a signal (for example, in the range 8–12 GHz). This signal is then fed to the

*Portions of this section are adapted, with permission from Plenum Press, from the Proceedings of Review of Progress in Quantitative Nondestructive Evaluation, Vol. 10B, pp. 1431–1436, 1991 (42).

Copyright © Marcel Dekker, Inc. All rights reserved.

transmitting horn antenna. After passing through the sample under test, the signal is received by another horn antenna and fed into the receiver. By comparing the measured amplitudes and phases of signals in the absence and presence of the test sample, we can determine the complex dielectric constant or thickness of the sample, provided the effects of reflections at the air–material interfaces are small. In some cases, a reference signal may be provided to the receiver for comparison purposes (as in the case of the Doppler radar example), as shown in Figure 9.24.

It should be noted that the theory to be presented next assumes that the surface of the dielectric sample is perpendicular to the axis of the horns and that the illuminating wave is plane. Thus, slab tilt should be minimized and the horns should be far enough away from the slab (or a dielectric lens should be used) to ensure that the wavefront illuminating the slab is as planar as possible in order to reduce errors in calculating the dielectric constant.

A plane wave propagating in the z-direction within a slab (where z is perpendicular to the surface of the slab) has a complex propagation constant given by

$$\alpha + j\beta = j\frac{2\pi\sqrt{\varepsilon_r}}{\lambda_0} = jk_0\sqrt{\varepsilon_r' - j\varepsilon_r''} \tag{36}$$

where α is the *attenuation constant* in nepers/m and β is the *phase constant* in radians/m. In general, these parameters depend on the orientation of the incident electric field vector in the plane of the slab. If α and β can be determined from measurements, then the real and imaginary parts of the dielectric constants can be calculated from the equations (42)

$$\varepsilon_r' = \frac{(\beta^2 - \alpha^2)\lambda_0^2}{(2\pi)^2} \tag{37}$$

and

$$\varepsilon_r'' = \frac{2\alpha\beta\lambda_0^2}{(2\pi)^2} \tag{38}$$

For lossy media, the depth of penetration into the medium is defined as the distance from the surface (in the medium) where the power carried by the penetrating wave has decayed to $1/e$ of its value at the surface. Hence the penetration depth, δ_p, in a dielectric material (analogous to skin depth in conductors, except that skin depth is defined in terms of field amplitude rather than power) is

$$\delta_p = \frac{1}{2\alpha}(\text{m}) \tag{39}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

It can be shown that the complex transmission coefficient for a planar dielectric slab of thickness, d, is given by (43)

$$\tau = e^{-(\alpha+j\beta)d} \frac{1}{1 + \dfrac{\rho^2}{1-\rho^2}\left(1 - e^{-2(\alpha+j\beta)d}\right)} \tag{40}$$

where

$$\rho = \frac{1-\sqrt{\varepsilon_r}}{1+\sqrt{\varepsilon_r}} \tag{41}$$

is the reflection coefficient at the air–material interface. The left-hand exponential factor in Eq. (40) represents propagation within the slab material, while the right-hand fractional factor contains the effects of all reflections at the air–material interfaces. If ρ is not small, then this second factor cannot be neglected in the transmission coefficient unless the thickness of the slab is small compared to a wavelength in the material. This latter condition was satisfied for the experimental results presented below.

Results

As an example of the use of the transmission method, we cite some data taken for Kevlar composite (42). The samples were constructed of unidirectional Kevlar fibers bonded in an epoxy matrix (250°F cure). The Kevlar material used in these measurements was made of a synthetic aromatic polymer base and the plys (layers) were highly anisotropic, both mechanically and, as we shall see, electromagnetically. The ply thickness was 0.005 in.

The results for the measured dielectric constant (average value in the X-band frequency range) are shown in Table 9.6. These results are for the incident electric field aligned with the fibers (vertical polarization), and for the incident electric field aligned perpendicular to the fibers (horizontal polarization). At least 10 independent measurements were averaged to produce the results shown in the tables. Even with this averaging, the accuracy of this measurement technique is not as good as that obtainable using the completely-filled short-circuited waveguide technique. Based on these measurements, the dielectric constant is considered to be constant over the measurement frequency band, and hence its average value (measured at several frequencies in this band) is reported in Table 9.6.

The measured dielectric constant of Kevlar at 10 GHz was compared to that reported in (44), and good agreement was obtained. Note that the values obtained for vertical polarization are somewhat larger than those obtained for horizontal polarization. This behavior can be explained by the fact that the fibers in the Kevlar sample were parallel to the vertically polarized incident electric field, and

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

TABLE 9.6 Average Dielectric Constant for Kevlar—Vertical and Horizontal Polarizations

Frequency (GHz)	Complex Dielectric Constant Vertical Polarization	Complex Dielectric Constant Horizontal Polarization
8.2–12.4	4.53 − j0.50	3.50 − j0.42

thus there is more interaction with the fibers in this case. If the fibers had been conducting, such as the graphite fibers used in graphite/epoxy composites, there would have been even more of a difference between the two polarization cases.

9.5.3 Inspection of Layered Dielectric Composites Using Reflection Measurements*

One of the more active areas of research and development in microwave NDE techniques has been in the inspection of multilayered dielectric composite structures backed by either a conducting plate, or free space, or an infinite half-space of a dielectric material. These techniques divide naturally into near-field microwave approaches, in which an open-ended rectangular waveguide or an open-ended coaxial probe is employed (45–55), and far-field (uniform plane-wave) approaches (4,34,41,56–58). In this section we discuss the latter approach.

Introduction

Materials composed of layered dielectric slabs bonded together adhesively are used widely today in many types of structural components. A disbond between any two slabs caused by the absence of adhesive or a delamination (a pancake-like void) can cause a major failure in the component. In many application areas, such as in the aerospace and construction industries, these materials are often backed by a conductor and this limits the microwave inspectability to only one side of the material. Clearly, it is desirable to have a noncontacting inspection technique for detecting and evaluating such delaminations or disbonds. In this section we present the basic theory of a technique that utilizes the phase characteristics of the reflection coefficient of a microwave signal that partly passes through the layered dielectric composite and is reflected by the backing conducting plate. This technique can also be used to monitor the thickness

*Portions reprinted, with permission, from IEEE Transactions on Instrumentation and Measurements; Vol. IM-39, No. 6, pp. 1059–1063, December 1990 (57) (Copyright © 1990 IEEE).

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

uniformity of adhesive layers, which is useful, for example, for manufacturing process-control purposes.

The technique discussed here is applicable to lossy materials in general, provided the loss does not prevent reasonable signal penetration into the material. However, signal phase variation depends primarily on the real part of the relative dielectric constant, whereas the attenuation of the signal as it travels through the material depends primarily on the imaginary part of the relative dielectric constant. For the purpose of our discussions, we ignore amplitude variations and limit our considerations to phase measurements only and, hence, to lossless materials (slabs with $\varepsilon_r'' = 0$). This measurement technique is applicable to structures made of any number of layers but, to be specific, we limit our model derivation to three layers (two dielectric slabs and one delamination or disbond layer).

Theory

Consider a uniform plane wave incident normally on a multilayer dielectric medium backed by a conducting plate. The *incident electric field in air*, E_0^i, is partially transmitted into the layered medium and, after being totally reflected by the conducting plate, it emerges as a *reflected electric field*, E_0^r. The ratio of these two signals gives the reflection coefficient of the layered structure. The difference in phase between the reflection coefficients observed in the presence and absence of a disbond (or a defect in general) is related to the *disbond thickness*, the frequency, f, and the relative dielectric constant, ε_r, of the disbond layer. If f and ε_r are known, then the disbond thickness can be calculated from the measured phase difference.

Figure 9.25 illustrates a layered structure consisting of two dielectric slabs separated by a disbond layer or delamination, all backed by a conducting plate. Calculation of the reflection coefficient for such a stratified medium involves deriving the forward- and backward-traveling electric and magnetic field components in each layer, given an incident field, and application of the appropriate boundary conditions at each interface. In general, the propagation constant, $\alpha + j\beta$, in each layer is a complex quantity that describes the phase and attenuation characteristics of the electromagnetic wave as it propagates inside the given layer. As mentioned earlier α is the attenuation constant (neper/m) and β is the phase constant (rad/m). The derivation procedure outlined in this section applies equally well to lossy media, as well as to structures with more than three layers.

Referring to Figure 9.25, the fields in the nth layer (assuming $\alpha = 0$) are given by the following expressions (superscripts i and r denote incident and reflected fields outside the layered structure, respectively, while superscripts $+$ and $-$ represent forward- and backward-traveling waves inside a layer, respec-

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

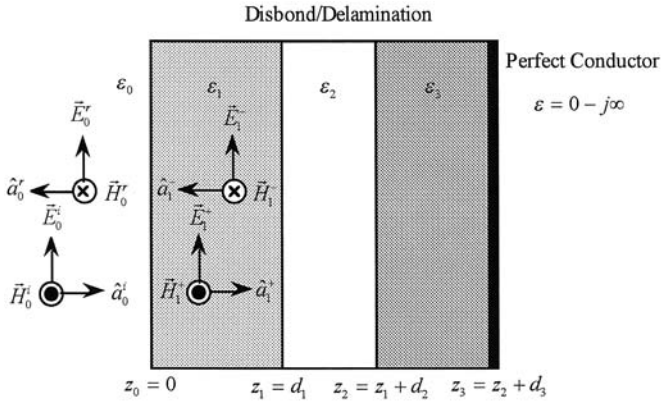


FIGURE 9.25 Incident and reflected fields in a stratified medium consisting of two dielectric layers (ε_{r1} and ε_{r3}) and a disbond or delamination layer (ε_{r2}), all backed by a conducting plate (not to scale). (From Ref. 57; reprinted with permission.)

tively, n denotes the layer number, arrows indicate vector quantities, and the symbol ^ indicates unit vectors):

$$\vec{E}_n = \hat{x}\left(E_n^+ e^{-j\beta_n z} + E_n^- e^{+j\beta_n z}\right) \tag{42}$$

$$\vec{H}_n = \hat{y}\frac{1}{\eta_n}\left(E_n^+ e^{-j\beta_n z} - E_n^- e^{+j\beta_n z}\right) \tag{43}$$

where

$$\beta_n = \frac{2\pi}{\lambda_0}\sqrt{\varepsilon_{rn}} \tag{44}$$

and

$$\eta_n = \sqrt{\frac{\mu_0}{\varepsilon_0 \varepsilon_{rn}}} \tag{45}$$

Here, λ_0 is the wavelength in free space, μ_0 and ε_0 are the permeability and permittivity of free space, respectively, ε_{rn} is the relative dielectric constant of the nth layer ($n = 1, 2, 3$), and η_n is the intrinsic impedance of the nth layer.

The boundary conditions in this plane-wave case are simple to state: the total tangential electric and magnetic fields must be continuous at each dielectric–dielectric interface and the total tangential electric field must be zero on the

Copyright © Marcel Dekker, Inc. All rights reserved.

surface of the perfect conductor. Starting at the conductor interface $z_3 = d_1 + d_2 + d_3$, we have

$$E_3^+ e^{-j\beta_3 z_3} + E_3^- e^{+j\beta_3 z_3} = 0 \tag{46}$$

or

$$E_3^- = -E_3^+ e^{-j2\beta_3 z_3} \tag{47}$$

Substituting Eq. (47) into Eq. (42) for layer 3 gives:

$$\vec{E}_3 = \hat{x}\left(E_3^+ e^{-j\beta_3 z} - E_3^+ e^{+j\beta_3 (z-2z_3)}\right) \tag{48}$$

$$\vec{H}_3 = \hat{y}\frac{1}{\eta_3}\left(E_3^+ e^{-j\beta_3 z} + E_3^+ e^{+j\beta_3 (z-2z_3)}\right) \tag{49}$$

Using this result and progressively applying the continuity boundary conditions at $z_2 = d_1 + d_2$, $z = d_1$, and $z_0 = 0$, we can derive the *input reflection coefficient*, $\Gamma_{in} = E_0^r / E_0^i$, at the surface of the layered structure. The result is

$$\Gamma_{in} = \frac{(\eta_1/\eta_0)D - 1}{(\eta_1/\eta_0)D + 1} \tag{50}$$

where

$$D = \frac{1 + Ce^{-j2\beta_1 d_1}}{1 - Ce^{-j2\beta_1 d_1}} \tag{51}$$

$$C = \frac{(\eta_2/\eta_1)B - 1}{(\eta_2/\eta_1)B + 1} \tag{52}$$

$$B = \frac{1 + Ae^{-j2\beta_2 d_2}}{1 - Ae^{-j2\beta_2 d_2}} \tag{53}$$

and

$$A = \frac{(\eta_3/\eta_2)j\tan\beta_3 d_3 - 1}{(\eta_3/\eta_2)j\tan\beta_3 d_3 + 1} \tag{54}$$

The phase of the reflection coefficient is given by the usual formula

$$\phi = \tan^{-1}\left[\frac{\mathrm{Im}(\Gamma_{in})}{\mathrm{Re}(\Gamma_{in})}\right] \tag{55}$$

Eqs. (50) and (55) provide a model that allows us to relate measured reflection-coefficient data to material parameters. For example, if we calibrate using a layered material having no disbond, the measured phase difference between the disbond and no-disbond cases can be used with Eq. (55) to detect the presence of

Copyright © Marcel Dekker, Inc. All rights reserved.

a disbond layer and characterize it (assuming the other layers are the same in the two cases).

Experimental Apparatus

The derivation given in the preceding section assumes that the fields inside all the dielectric layers are uniform plane waves. The practical implication of this assumption is that, if measured data are to be used with this model, the layered structure must be placed in the far-field of the antennas that are used in the experiment. A typical measurement apparatus is shown in Figure 9.26. A microwave oscillator is used to generate a *continuous-wave* (CW) signal at a given frequency. An isolator is placed between the oscillator and the rest of the system to prevent reflections from reaching the oscillator and detuning or damaging it. The signal is then split into test and reference signals. The reference signal becomes the input signal to the reference channel of a microwave network analyzer. The test signal is fed through another isolator to prevent reflections from corrupting the reference signal, and then to a small horn antenna from which it radiates and illuminates the multilayered medium under test. The signal then propagates through the medium and is reflected by the conducting plate, after which it is received by another small horn antenna nearby and subsequently fed into the test channel of the network analyzer. It is also possible to use a single horn antenna for both transmitting and receiving by using a circulator or a dual-directional coupler. A potential drawback of this arrangement is decreased isolation between transmitted and received signals, but, depending on the system characteristics, this may or may not be an issue. The network analyzer

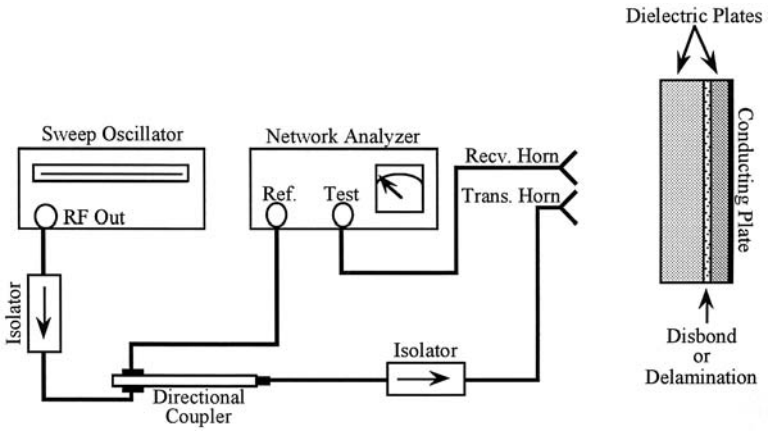


FIGURE 9.26 Experimental arrangement for microwave reflection measurements. (From Ref. 57; reprinted with permission.)

Copyright © Marcel Dekker, Inc. All rights reserved.

compares the amplitude and the phase of the reflected test signal with those of the reference signal and displays the resulting relative values.

It is important to note that the phase difference is measured with respect to the front surface of the layered structure. Thus, when making measurements of this type, one typically calibrates the phase measurements by first measuring a structure that does not contain any disbond, and then locates the flawed structure such that its front surface is in exactly the same position as that of the calibration specimen. This may seem tedious at first, but, using a suitable jig and practicing a few times, an operator quickly learns how this setup procedure can be accomplished effectively. It should be noted that, when conducting near-field experiments, such a calibration may *not* be necessary. This is due to the fact that the known distance between the sensor (antenna or open-ended transmission line) and the specimen under test (known as the standoff distance or liftoff) is often incorporated explicitly in the model used to interpret the measurements.

Theoretical Results

Variations in the *reflection-coefficient phase difference*, $\Delta\phi$, as a function of disbond thickness, d_2, for four different cases with $\varepsilon_{r1} = \varepsilon_{r3} = 3.5$ and $f = 10$ GHz are shown in Figure 9.27. Curves (a) and (b) show the results for a layered structure composed of two 50 mm-thick dielectric plates separated by air

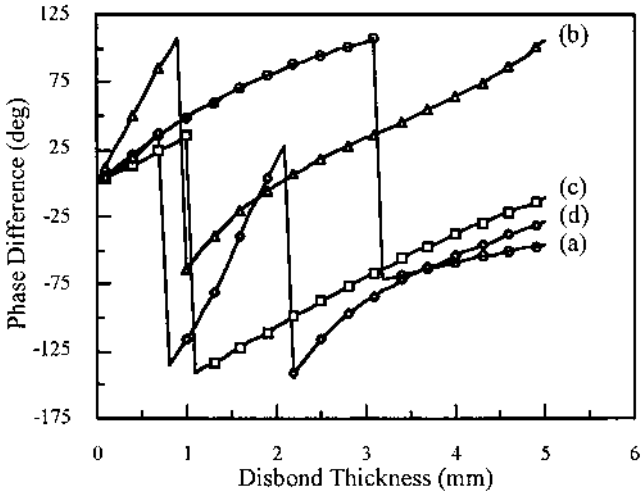


FIGURE 9.27 Reflection-coefficient phase difference as a function of disbond thickness for (a) $d_1 = d_3 = 50$ mm, $\varepsilon_{r2} = 1$, (b) $d_1 = d_3 = 50$ mm, $\varepsilon_{r2} = 6.5$, (c) $d_1 = d_3 = 25$ mm, $\varepsilon_{r2} = 1$, (d) $d_1 = d_3 = 25$ mm, $\varepsilon_{r2} = 6.5$. (From Ref. 57; reprinted with permission.)

Copyright © Marcel Dekker, Inc. All rights reserved.

($\varepsilon_{r2} = 1$) and epoxy ($\varepsilon_{r2} = 6.5$), respectively. In both cases $\Delta\phi$ increases as the thickness of the second layer increases. The sharp 180° phase changes (called phase wrapping) are caused by the *arctan* function in Eq. (55); the period of this discontinuous behavior decreases when the disbond dielectric constant increases. Decreasing the thicknesses of the dielectric plates to 25 mm (curves (c) and (d)) produces faster $\Delta\phi$ variations compared to the first two cases.

Figure 9.28 shows the computed results for variations in $\Delta\phi$ at 10 GHz as a function of the dielectric constant of the second layer when $\varepsilon_{r1} = \varepsilon_{r3} = 3.5$ and $d_2 = 1$ mm. Curves (a) and (b) were computed for $d_1 = d_3 = 50$ mm and $d_1 = d_3 = 25$ mm, respectively. For the thicker dielectric plates (case (a)) the 180° phase reversal occurs at a higher value of disbond dielectric constant compared to that of the thinner plates (case (b)). For these same two cases at 5 GHz (curves (c) and (d)), the 180° phase reversal occurs for much higher values of disbond dielectric constant, and $\Delta\phi$ variations at this frequency are not as pronounced when comparing thicker and thinner dielectric plates. This observation indicates that the sensitivity of $\Delta\phi$ to disbond dielectric constant is higher at higher frequencies.

From these results it is apparent that, due to the repetitive nature of $\Delta\phi$ (180° phase wrap) versus disbond thickness, the disbond thickness cannot be

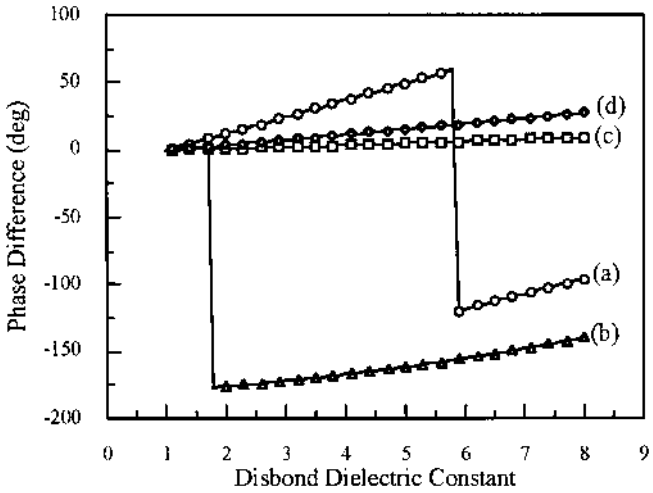


FIGURE 9.28 Reflection-coefficient phase difference as a function of disbond dielectric constant for (a) $d_1 = d_3 = 50$ mm, $f = 10$ GHz, (b) $d_1 = d_3 = 25$ mm, $f = 10$ GHz, (c) $d_1 = d_3 = 50$ mm, $f = 5$ GHz, (d) $d_1 = d_3 = 25$ mm, $f = 5$ GHz. (From Ref. 57; reprinted with permission.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

determined uniquely. There are at least three remedies for this ambiguity problem:

1. If a priori knowledge of the relative range of disbond thicknesses that are possible is available, it can be used to resolve this ambiguity.
2. It is possible to reduce the ambiguity by a judicious choice for the operating frequency.
3. Using more that one operating frequency could also be used to resolve this ambiguity, but this increases the complexity of the inspection.

Experimental Results

The validity of these theoretical findings have been tested experimentally using a synthetic rubber plate of thickness $d_1 = 25.45$ mm ($\varepsilon_{r1} = 3.32$) and a Plexiglas^{TM} plate of thickness $d_3 = 26.2$ mm ($\varepsilon_{r3} = 2.53$) separated by a variable air gap and all backed by a metal plate (57). Using the experimental setup discussed above, $\Delta\phi$ was measured at 10 GHz as a function of the air-gap thickness between these two plates. Figure 9.29 shows the comparison between theoretical and experimental results. The vertical lines through the circles indicate the measurement standard deviations; clearly the agreement between theory and measurement is very good. Thus, the measurements clearly indicate the presence of a disbond, and inversion of the data should yield accurate values for the air-gap thicknesses.

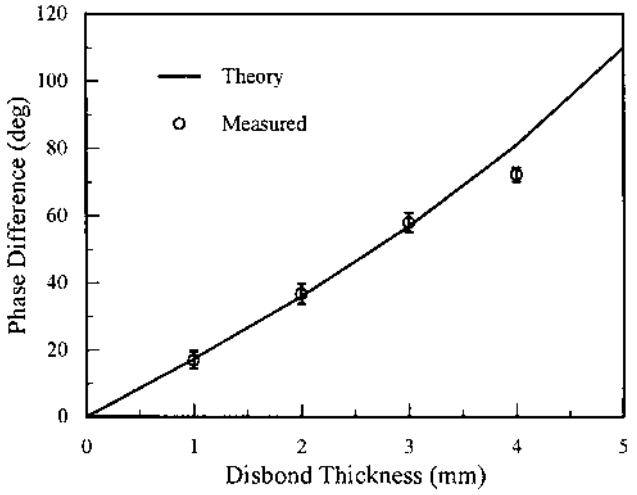


FIGURE 9.29 Comparison of theory and experiment at 10 GHz for a simulated composite material made from synthetic-rubber and Plexiglas plates separated by a variable air gap (disbond). (From Ref. 57; reprinted with permission.)

Copyright © Marcel Dekker, Inc. All rights reserved.

It is important to note that in practice a disbond is usually less than a millimeter thick. From Figure 9.29 it is apparent that there are about 20° of phase difference for 1 mm of disbond. This translates to a resolution of 50 μm, which is considered very good.

The results of a second experiment using two identical Plexiglas plates are shown in Figure 9.30. In this case the disbond thickness was increased to include a 180° phase wrap in the data. Again, the agreement between theory and experiment is very good.

These results illustrate the potential of this technique for detecting and evaluating disbonds. The agreement between the theoretical and measured results can be improved even more by conducting more independent measurements and by having better knowledge of the dielectric slab permittivities. If the disbond thickness and frequency are such that a 180° transition in the $\Delta\phi$-versus-disbond-thickness curve just happens to occur, a slight change in operating frequency will alleviate this problem. Also, this same technique can be used to accurately determine the thickness of a dielectric slab backed by a conducting plate (41). All that is needed in the model derivation is to change the number of layers to one.

9.5.4 Microwave Inspection Using Near-Field Probes

During recent years, progress in microwave NDE has focused on developing novel inspection techniques and increasing its realm of applications. One such technique involves using various microwave probes or sensors whose near-field

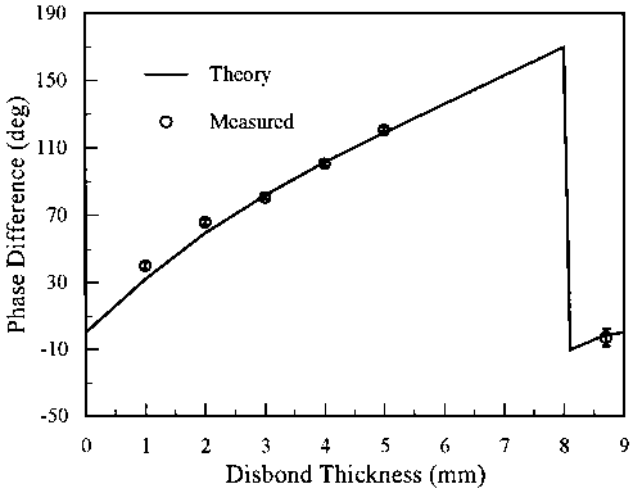


FIGURE 9.30 Comparison of theory and experiment at 10 GHz for a simulated composite material made from two identical Plexiglas plates separated by a variable air gap (disbond). (From Ref. 57; reprinted with permission.)

Copyright © Marcel Dekker, Inc. All rights reserved.

(rather than far-field) properties form the basis of the inspection process. The multilayer dielectric-composite inspection methodology presented in the previous section is an example of a far-field technique. In this case the material under test is placed sufficiently far from a microwave antenna such that the fields incident on the material are planar and there are no field variations transverse to the propagation direction. On the other hand, this is no longer true in the near-field of a sensor and the field properties (magnitude and phase) are dependent upon all spatial variables. In particular, the field properties vary nonlinearly as functions of distance from the sensor (standoff distance).

Three basic regions can be identified that surround a microwave radiator or source (59). The region closest to the radiator is known as the *reactive near-field* region in which reactive (quasistatic) fields predominate. The next region further away from the radiator is known as the *radiating near-field* region in which both radiating and reactive fields are present, but here the radiating fields begin to predominate. The third region (furthest away from the radiator) is known as the *far-field* region in which only radiating (traveling-wave) fields exist. At great enough distances from the source the fields in this region have planar phase fronts, and thus they are called uniform plane waves. They are also referred to as transverse electromagnetic (TEM) waves since the electric- and magnetic-field vectors are orthogonal to each other and to the direction of propagation.

It is instructive to compare far-field and near-field microwave NDE techniques in more detail. Far-field techniques are characterized by the following features:

1. The specimen under test is located in the far-field region of an antenna.
2. The antenna and the specimen are not in physical contact.
3. The antenna should be an efficient radiator (receiver) and, therefore, the number of suitable antenna designs is limited.
4. The analytical formulations that describe the interactions of plane waves with material media (such as described earlier in this chapter) are relatively straightforward and easy to evaluate.
5. The number of controllable measurement parameters available for optimizing the inspection process are relatively limited.
6. The sensitivity to structural parameters such as thickness variations and very thin disbonds in complex stratified dielectric composites is relatively low.
7. The spatial (lateral) resolution is relatively coarse and predominantly determined by the size of the operating wavelength (and the signal-processing techniques applied).

On the other hand, the characteristics of near-field techniques are exemplified by the following features:

1. The specimen under test is located in the near-field region of a probe or sensor.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

2. The probe and the specimen may or may not be in contact.
3. A wide variety of useful near-field probes are available, for example:

   Open-ended rectangular and circular waveguides
   Open-ended coaxial lines
   Short monopole antennas
   Microstrip patches
   Open-cavity resonators.

4. The analytical formulations that describe the interaction of near fields with material media are relatively complex and require extensive computer resources for their evaluation.
5. Additional measurement parameters are available for optimizing the inspection process, for example, the standoff distance (liftoff) between the probe and the specimen under test.
6. Since the probe fields are more concentrated, a high level of measurement sensitivity to slight variations in dielectric or geometrical parameters can be obtained.
7. The spatial (lateral) resolution can be quite high, since in the near-field region this resolution is primarily determined by the physical dimensions of the probe and not by the operating wavelength.
8. Relative phase shifts in the reflection or transmission coefficients are more readily measured than in the far-field case.
9. A near-field microwave NDE system is usually simpler, and more compact and portable, than a far-field system.

In both the far-field and near-field cases one-sided (reflection) and two-sided (transmission) measurements are possible.

Figure 9.31 shows the general schematic diagram of a near-field inspection system that can be used for the nondestructive evaluation of many different types of materials. This figure illustrates a painted (coated) steel specimen being examined (scanned) for the presence of a finite-sized patch of rust under the paint. Such systems have been used in the past few years to inspect a wide variety of materials and composites. Applications have included detecting steel reinforcing bars in concrete, localized defects and flat-bottomed holes in thick glass-reinforced polymers, disbonds, delamination and impact damage in sandwich composites, variations in resin binder level in low-density fiberglass, and localized porosity and rust under paint and thick composite-laminate coatings (26,36,60–64). The scanning system provides the capability for obtaining a two-dimensional raster scan (C scan) of the specimen under test. Consequently, defects may be viewed as two-dimensional images and, because of the high spatial resolution provided by near-field techniques, a good estimate of the extent (area) of a defect, as well as other attributes, may be obtained quickly without the

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

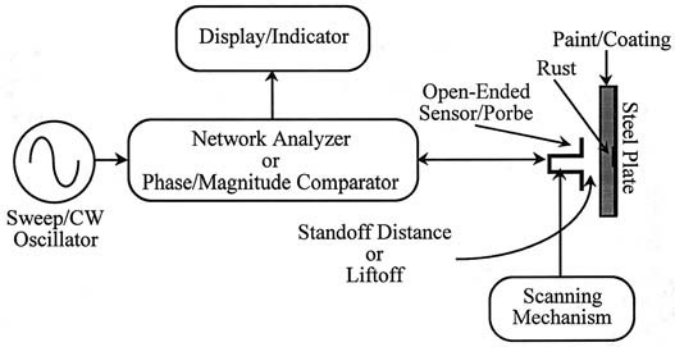


FIGURE 9.31 Schematic diagram of a near-field inspection system, shown scanning a painted steel plate for the presence of a patch of rust under the paint.

need for complicated signal-processing algorithms. This is one of the more powerful capabilities of microwave near-field inspection.

A complete discussion of the electromagnetic field properties in the near-field region of several common probes, and the analysis of the interaction of such fields with material media for NDE purposes, is outside of the scope of this chapter. However, it is important that the reader have an intuitive understanding of the basic mechanisms involved in near-field measurements. The aim of the following brief discussion is to provide this understanding.

Near-field measurements, such as the one depicted in Figure 9.31, often make use of open-cavity resonator probes. The term "open" refers to the fact that one end of the cavity, which could be constructed from a section of waveguide or coaxial line, is open in the sense that the fields in the cavity are free to extend outside the cavity proper and interact with the material under test. By properly selecting the frequency of operation (which determines the size of the resonator) and the standoff distance, the resonator can be operated at a sensitive portion of its resonant characteristic when interacting with a nondefective region of the specimen under test. Then, when a defective region comes into the "field of view" of the resonator, the resonant response of the probe changes because of the different dielectric properties of the defect. This change can be monitored by measuring the magnitude and/or phase of the signal reflected from the resonator or, equivalently, by measuring the properties of the corresponding standing wave in the waveguide or transmission line connected to the resonator. Either of these measurements can be optimized for sensitivity and the presence of typical defects can easily be detected. The data obtained in this way can also be used (with appropriate models) to quantify the dielectric and geometrical properties of the defective region.

Copyright © Marcel Dekker, Inc. All rights reserved.

Examples of Near-Field Microwave Imaging

This section provides a few examples of the inspection and imaging processes discussed above. The reader should consult the references for more detail. In particular, (61) is recommended for a thorough discussion of standoff-distance analysis and a clear demonstration of the use of this parameter to one's advantage.

The images to be presented in this section were produced by placing the specimen under test on a two-dimensional computer-controlled scanning table. Then, an open-ended rectangular-waveguide probe, operating at a selected frequency and standoff distance, was placed above the specimen. A two-dimensional data array was then acquired by moving the specimen in a raster fashion underneath the fixed probe (using the table) and measuring the DC voltage values that are proportional to the phase and/or magnitude of the reflection coefficient at the open end (aperture) of the waveguide. This two-dimensional matrix of voltage values was then normalized and divided into several levels. Each level was assigned a gray scale, thereby creating a gray-scale image of the specimen. The images shown here were only used to detect the presence of anomalies and to obtain good estimates of their spatial extent. Further analysis to determine more quantitative information about the dielectric properties of the defective regions could be performed by analyzing the measured data, but a discussion of this is outside of the scope of this chapter.

Moisture Permeation Detection. A 25.4-mm-thick low-density and low-permittivity foam sample, 44 cm square, was fabricated and 0.5 cc of water was injected near one of its surfaces. Since water has a very high dielectric constant compared to foam, it makes a good high-contrast defect for demonstrating the imaging capabilities of the system. An open-ended rectangular waveguide probe, operating at a frequency of 24 GHz and a standoff distance of 1 mm, was used to obtain a near-field image of this small volume of water and its permeation. The waveguide aperture in this case was 10.67 mm $\times$ 4.32 mm. The probe was placed on the side of the foam sample opposite to that in which the water was injected. Figure 9.32 shows the resulting 80-mm $\times$ 80-mm image of this sample (with data taken at 2-mm steps in both directions). The presence of an anomaly (which is known to be water) in the middle of the image is clearly evident. In addition, the spatial extent of the permeation of this small volume of water is also clearly shown (the darker region), but, of course, this apparent extent is somewhat larger than the true extent because the electric fields surrounding the probe spread out as the distance away from the waveguide aperture increases. The darker gray areas are produced by interaction of the probe with areas in which the waveguide aperture is exposed partially to water and partially to dry foam and also by the fact that the water density varies due to permeation. The uniform medium gray indicates the area of the foam sample that does not contain detectable water.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

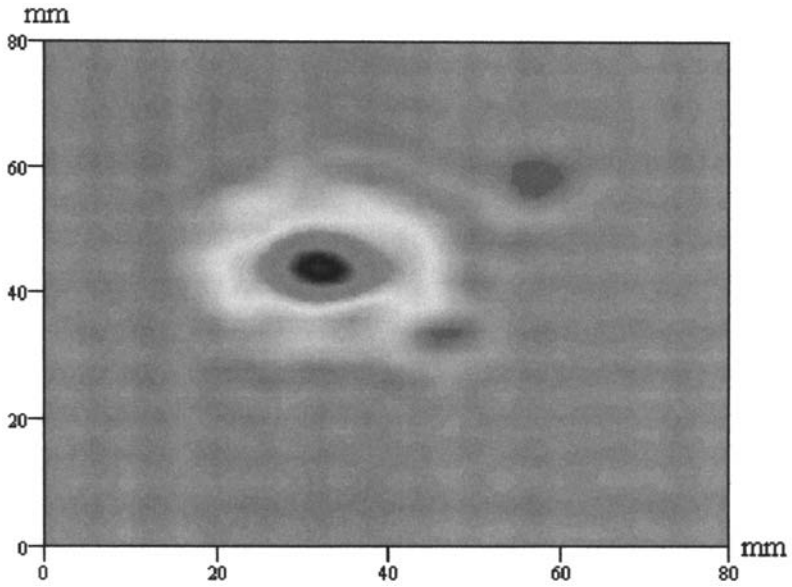


FIGURE 9.32 Near-field image of a 25.4-mm-thick foam sample with 0.5 cc of water injected in the sample surface opposite to the one being scanned. The image was taken at a frequency of 24 GHz and a standoff distance of 1 mm.

Impact Damage Detection in Honeycomb Composites. In another experiment to test the ability of the same near-field system to detect impact damage, a sample was made from a 16.5-cm-square and 25.4-mm-thick section of honeycomb with two 0.5-mm-thick Kevlar reinforced epoxy skins adhesively attached to its surfaces. A pool ball, weighing at 200 gm, was dropped on one of its surfaces from a distance of 193 cm, creating a slight amount of impact damage. Initially the maximum surface depression at the point of impact was measured to be 0.05 mm. However, as in most impact-damage cases, this surface depression decreased over time to a value that is difficult to both see and measure. Nevertheless, it is believed that damage remains in most impact-damaged composites and that this contributes to the eventual failure of the specimen while under load. This residual damage, while not readily detectable by the human eye, is detectable by microwaves. Possible reasons for this microwave detectability are (a) the skin may become deformed slightly and cause slight variations in standoff distance as a sensor scans the specimen, (b) the adhesion between the skin and the honeycomb may be lost, resulting in a thin disbond, (c) the impact may cause microscopic porosity around the Kevlar strands in the skin, in particular around the cross-wave regions, and (d) the honeycomb itself may be

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

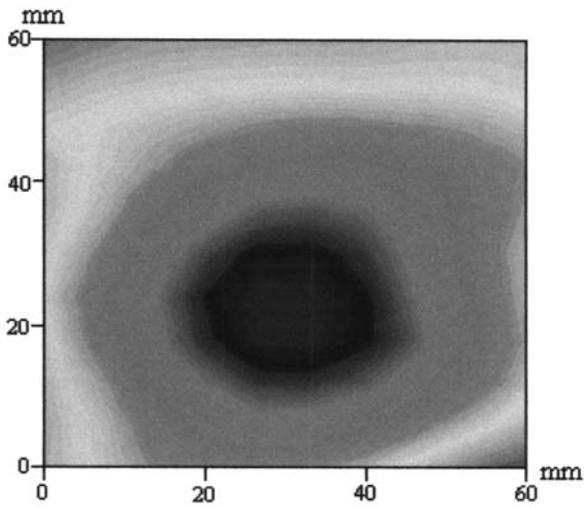


FIGURE 9.33 Near-field image of impact damage in a 25.4-mm-thick honeycomb sample with 0.5-mm-thick Kevlar reinforced epoxy skin. The image was taken at a frequency of 24 GHz and a standoff distance of 2 mm.

physically damaged (however so slightly). These various types of damage may either cause a change in the dielectric properties or the geometry of the honeycomb sample. In general, a combination of all of these effects is likely to occur.

A 60-mm × 60-mm image of this impact-damaged honeycomb produced using the rectangular open-ended waveguide probe at a frequency of 24 GHz and a standoff distance of 2 mm is shown in Figure 9.33. The dark region (the amplitude represented by this color is not the same as that in Figure 9.32) indicates the damaged area and its extent. The medium gray-color transition areas again result partly from the finite size of the waveguide aperture and partly from the actual progression of the damage in the specimen. The reader is referred to (62) and (65) for a more detailed discussion of impact-damage detection.

Rust Under Paint.* Detection of rust or corrosion under paint and composite laminate coatings in various industrial and military environments is an important concern in many applications. When rust is detected reliably in its early stages of development, savings of millions of dollars in maintenance costs can be realized because of damage minimization and reduction of repaint cycles.

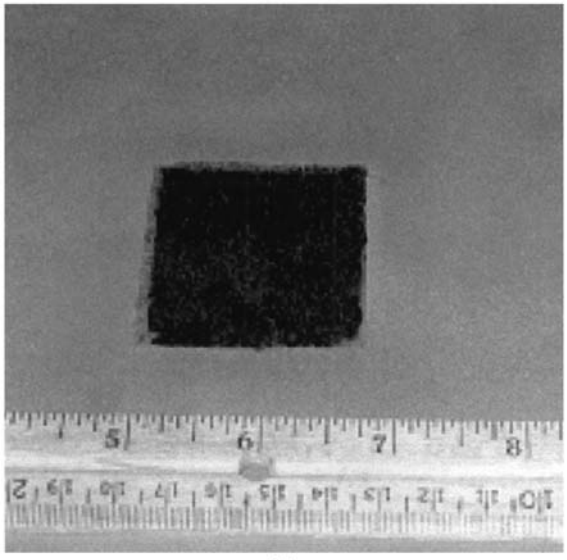
*Portions of this section are reproduced, with permission from Springer-Verlag, from Research in Nondestructive Evaluation, Vol. 9, pp. 201–212, 1997 (64).

Copyright © Marcel Dekker, Inc. All rights reserved.

Since the presence of rust or corrosion may be considered as an additional new thin dielectric layer under the paint or coating, microwave NDE methods are very well suited to inspecting for this type of damage.

The following experimental results illustrate the practical utility of near-field microwave imaging in this case. Figure 9.34a shows a photograph of a steel specimen having a 40-mm × 40-mm area of rust. This specimen was produced from a relatively flat piece of steel which was covered by a naturally occurring thin layer of rust. The 40-mm × 40-mm area was masked by a piece of tape and the remaining exposed surface was sand blasted to remove the rust. The average thickness of the rust layer was measured (using a micrometer) to be approximately 0.08 mm. Subsequently, this specimen was painted as uniformly as possible with common spray paint to a thickness of 0.6 mm. This thickness was achieved after ten painting cycles. Microwave near-field images of the rust specimen were taken after each painting cycle.

Figure 9.34b shows a 80-mm × 80-mm image of the rust patch under a paint thickness of 0.292 mm, obtained at a frequency of 12 GHz and a standoff distance of 3 mm (data were taken in steps of 2 mm in both directions). The rusted area is clearly visible. Even though the waveguide aperture in this case was



(a)

FIGURE 9.34a A 40-mm by 40-mm area of rust on a steel plate. (From Ref. 64; reprinted with permission.)

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

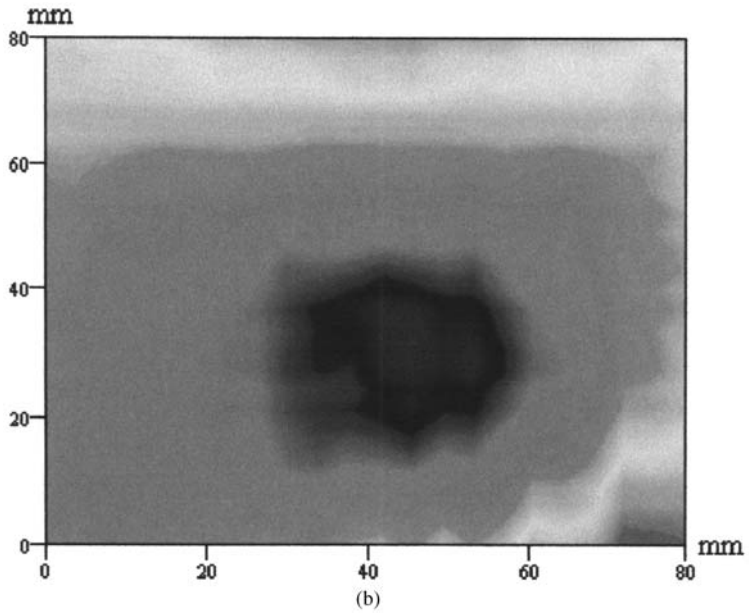


FIGURE 9.34b Near-field image of the rust specimen under a paint thickness of 0.292 mm. The image was taken at a frequency of 12 GHz and at a standoff distance of 3 mm.

19.05 mm × 9.525 mm, the spatial extent of rusted area seen in the image closely matches the actual area of the rust (Figure 9.34a).

Another image was produced for the same paint thickness, but this time using a frequency of 31 GHz and a standoff distance of 3 mm (data were again taken in 2-mm steps size in both directions). The results are shown in Figure 9.34c. Once again the rusted region is clearly visible. At this higher frequency the size of the waveguide aperture is about 9 times smaller than the aperture size used at 12 GHz and thus this image shows more detail in the rusted region. Besides having better spatial resolution, this smaller probe is more sensitive to standoff-distance variations.

The examples provided in this section are strong indications of the potential usefulness of near-field microwave inspection and imaging techniques for nondestructive testing purposes. As mentioned earlier, these systems can be designed and built to be operator-friendly, handheld, battery-operated, and portable. In many practical applications two-dimensional images similar to those shown here may not be necessary. A linear scan through a defective area should suffice when trying to detect the defect. However, quantifying the severity and extent of the defect may require additional data and signal processing.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

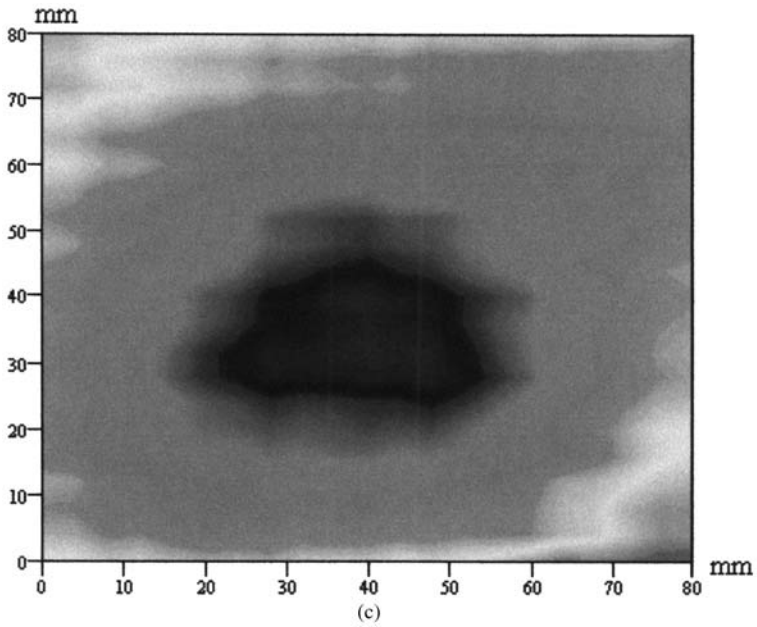


FIGURE 9.34c Near-field image of the rust specimen under a paint thickness of 0.292 mm. The image was taken at a frequency of 31 GHz and at a standoff distance of 3 mm.

9.5.5 Summary and Future Trends

Microwave NDE can be defined as the inspection and characterization of materials and structures using high-frequency electromagnetic energy. This well-developed branch of NDE has its own technical subject matter, applications, and particular research methods, as well as a characteristic sensor and instrumentation base. Well-known physical theories of the propagation and guiding of electromagnetic fields, and the interaction of such fields with matter (such as the theory of dielectric mixtures), are used to optimize measurements and to relate measured data to physical parameters of interest. A wide range of applications exist, the most common involving inspection of dielectric materials to assess variations in dielectric constant, loss factor, moisture content, physical structure, and dimensions. In some cases, these measurements are directly relatable to other physical and chemical properties such as porosity, cure state, corrosion, etc. Anomalies in the surfaces of conducting materials can also be detected.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Microwave sensing for nondestructive evaluation and quality control has been in use for over 45 years. During that period a wide array of application-specific microwave-NDE measurement techniques and hardware has been developed and put into use. In many cases this development has been facilitated by the availability of related technologies whose progress has been fueled mainly by needs in the radar and wireless communication fields. Recent trends in these fields have been toward miniaturizing and ruggedizing hardware, improving hardware performance, and integrating system functions. These trends suggest that in the future it will become more feasible and cost-effective to integrate microwave monitoring systems directly into manufacturing lines.

In view of the trend towards miniaturization, we can expect an increase in the use of sensor arrays to permit rapid inspection of larger areas or to achieve higher spatial resolution. In addition, microwave sensors will be increasingly operated at millimeter-wave frequencies and combined with other types of sensors to provide greater sensitivity and discrimination. Microwave NDE (like other NDE techniques) will also benefit from the current activities aimed at combining sensors with digital signal-processing and microcontroller hardware to produce "smart" sensors that are, for example, self-calibrating or adaptive to their environment. The increased use of wireless communication techniques to control the operation of, or acquire data from, an array of hard-to-access sensors is another likely possibility. Finally, microwave NDE is still a current field of active research (66–69), and such work will lead to continuous improvements in measurement techniques and data interpretation.

PROBLEMS

1. Calculate the wavelength, the propagation constant, and the intrinsic impedance of nonmagnetic media with the following relative dielectric constants of (a) 9, (b) $9 - j0.001$, (c) $9 - j0.01$ and (d) $9 - j1$ at a frequency of 5 GHz and 10 GHz. Comment on the results.
2. In problem 1, the wavelength becomes a complex quantity in lossy dielectric media. What do the real and the imaginary parts of a complex wavelength mean as far as the wave properties are concerned when traveling in these media?
3. Calculate the skin depth for a nonmagnetic conductor with a conductivity of 5×10^7 (S/m) at 10 MHz, 100 MHz, 1 GHz, 10 GHz and 100 GHz. If this conductor were to be magnetic, what should its permeability be so that at 100 MHz it would have the same skin depth as in the previous case at 10 GHz?

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

4. When a wave travels in a lossless medium the intrinsic impedance of the medium that relates the magnetic field intensity to the electric field intensity is a real quantity and hence the two fields are said to be in time phase with each other. However, when the medium is lossy, then its intrinsic impedance is a complex quantity that produces a phase lag/lead between the electric and magnetic field intensities. Can you draw a parallel between this situation and a low frequency circuit? If so, what would the fields be represented by in the circuit and how can the time phase lag/lead be described in the circuit (hint: think of circuit elements that may cause such a phase lag/lead). Calculate the relationship between the electric field and magnetic field intensities at a frequency of 10 GHz for media (nonmagnetic) with relative dielectric constants of (a) 16, (b) 16 − j0.04 and (c) 16 − j4.
5. Using the measurement accuracy analysis outlined in the chapter for measuring the dielectric properties of materials using a completely filled waveguide technique, explain the limitation that may be put on the measurements if the precise attenuator has dynamic ranges of 30 dB, 40 dB, 50 dB and 60 dB. Which of these attenuators would be more desirable to use? Explain your reason.
6. A dielectric slab is backed by a conducting plate. If the slab is nonmagnetic with a relative dielectric constant of ε_r and a thickness of d, derive the expression for the effective reflection coefficient (similar to that done in the chapter). If $\varepsilon_r = 6$ and you are to be able to distinguish between two such slabs with respective thicknesses of 8 mm and 8.5 mm, what frequency in the 8.2 GHz–12.4 GHz (X-band) would you use?
7. Repeat problem 6 for $\varepsilon_r = 6 - j0.9$. Compare the results with those obtained in Problem 6 and comment on the results.
8. Derive the expression for the effective reflection coefficient of an n-layered medium backed by a conducting plate. Each layer can have any thickness and dielectric constant (nonmagnetic layers).
9. Derive the expression for the effective reflection coefficient of a two-layered material backed by free space.
10. Derive the expression for the effective transmission coefficient of a two-layered material backed by freespace. Transmission coefficient is defined as the ratio of the transmitted electric field intensity to that of the incident electric field intensity.
11. Suppose you have a steel plate covered with 2 mm of paint with a relative dielectric constant of $\varepsilon_r = 3.5 - j0.05$. Over time a thin layer of rust develops between the paint and the steel plate. If the rust has a dielectric constant of $\varepsilon_r = 8 - j0.5$, plot the phase difference between the rusted and no-rust cases as a function of rust thickness between 0 mm and 1 mm (for every 0.05 mm interval) at frequencies of 10 GHz and 30 GHz. Comment on the results. If you are to use a network analyzer that provides for a phase

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

measurement of 0.1 degrees, how accurately can you determine the rust thickness?

SYMBOLS

a	Broad dimension of rectangular waveguide
a_s	Radius of a small sphere
$\hat{a}$	Unit vectors in an arbitrary direction
A, B, C, D	Intermediate variables
α, β	Real and imaginary parts of the propagation constant, k
C_0	Effective resonator capacitance
$\vec{d}$	Vector distance between charges
d_n	Thickness of the nth layer
$\vec{D}$	Electric flux-density vector
δ	Skin depth
δ_p	Penetration depth (power)
$\vec{E}$	Electric-field vector
$\vec{E}_t$	Transverse electric-field vector
$\vec{E}_{\text{inc}}, \vec{H}_{\text{inc}}$	Incident electric and magnetic fields
$\vec{E}_{\text{sc}}, \vec{H}_{\text{sc}}$	Scattered electric and magnetic fields
E_0^i, E_0^r	Incident and reflected electric fields in air
E_n^+, E_n^-	Electric fields of forward- and backward-traveling waves in nth layer
ε	Dielectric constant
ε_0	Permittivity of vacuum
ε^*	Effective complex dielectric constant
ε'	Real part of ε^*, the absolute permittivity
ε''	Negative imaginary part of ε^*, the absolute loss factor
ε_r	Relative complex dielectric constant with respect to the permittivity of vacuum
ε_r'	Real part of ε_r, the relative permittivity
ε_r''	Negative imaginary part of ε_r, the relative loss factor (loss factor)
ε_{rel}	Dielectric constant relative to any host medium
$\Delta\varepsilon'$	Change in the real part of the dielectric constant
f	Frequency
f_0	Resonant frequency
Δf	Frequency deviation from resonance
ϕ	Phase of reflection coefficient
$\Delta\phi$	Reflection-coefficient phase difference
Γ, Γ_0	Complex reflection coefficient
Γ_{in}	Input reflection coefficient at the surface of a layered structure
$\vec{H}_t$	Transverse magnetic-field vector

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

η	Characteristic impedance
η_0	Characteristic impedance of free space
j	Unit imaginary number
k	Propagation constant (wave number)
k_0	Free-space wave number
K, M	Empirical resonator calibration factors
L, L_n	Sample (plate) thickness
λ	Wavelength
λ_0	Free-space wavelength
λ_g	Guide wavelength
m, n	Integers
$\vec{m}$	Vector magnetic-dipole moment
μ	Permeability
$\hat{n}$	Unit vector
ω	Angular frequency
$\vec{p}$	Vector electric-dipole moment
$\vec{P}$	Electric polarization vector
q	Charge
Q, Q_0	Resonator quality factor
$\Delta(1/Q_0)$	Change in reciprocal resonator quality factor
r	Distance between point of observation and a scatterer
r_0	Resonator input resistance at resonance
R, G, L, C	Resistance, conductance, inductance, capacitance
ρ	Reflection coefficient at an air–material interface
S_{ki}	Scattering parameter
SWR	Standing-wave ratio
σ	Conductivity
σ_e	Equivalent conductivity
t	Time
$\tan\delta$	Loss tangent
τ	Complex transmission coefficient
v_c	Phase velocity
$V_i^{\pm}$	Equivalent voltage coefficients for incident and reflected waves in the ith transmission line or waveguide
Δv	Differential volume
x, y, z	Cartesian coordinates
$\hat{x}, \hat{y}$	Unit vectors in the x and y directions
X, Y	Solution variables
χ_e	Electric susceptability
Y_{in}	Input admittance
$z_{\min}$	Position of a standing-wave minimum

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

REFERENCES

1. RJ Botsco, RW Cribbs, RJ King, RC McMaster. Nondestructive Testing Handbook, 2nd ed., vol. 4. American Society for Nondestructive Testing, 1986, sec. 18.
2. A Kraszewski. Microwave Aquametry. New York: IEEE Press, 1996.
3. R Hochschild. Applications of microwaves in nondestructive testing. Nondestructive Testing **21**: 115–120, 1963.
4. TM Lavelle. Microwaves in nondestructive testing. Materials Evaluation **25**: 254–258, 1967.
5. R Wright, KHR Skutt. Electronics: Circuits and Devices. New York: Ronald Press, 1965.
6. RE Collin. Foundations for Microwave Engineering. New York: McGraw-Hill Book Co., 1966, pp 47–51.
7. AR von Hippel. Dielectric Materials and Applications. Cambridge, MA: MIT Press, 1954, pp 3–18.
8. CA Balanis. Advanced Engineering Electromagnetics. New York: John Wiley & Sons, 1989, pp 42–51.
9. CA Balanis. Advanced Engineering Electromagnetics. New York: John Wiley & Sons, 1989, pp 72–85.
10. JD Kraus. Electromagnetics. 4th ed. New York: McGraw-Hill Book Co., 1992, p 804.
11. PS Neelakanta. Handbook of Electromagnetic Materials. Boca Raton, FL: CRC Press, 1995, pp 1–56.
12. RE Collin. Foundations for Microwave Engineering. New York: McGraw-Hill Book Co., 1966, Chapter 2.
13. R Bracewell. The Fourier Transform and Its Applications. New York: McGraw-Hill Book Co., 1965, Chapter 9.
14. RE Collin. Foundations for Microwave Engineering. New York: McGraw-Hill Book Co., 1966, pp 34–38.
15. R Zoughi, S Bakhtiari. Microwave nondestructive detection and evaluation of disbonding and delamination in layered-dielectric slabs. IEEE Trans. Instrum. Meas. IM-39: 1059–1063, 1990.
16. RE Collin. Foundations for Microwave Engineering. New York: McGraw-Hill Book Co., 1966, pp 170–179.
17. JD Jackson. Classical Electrodynamics. 2nd ed. New York: John Wiley & Sons, 1975, pp 411–418.
18. RE Collin. Foundations for Microwave Engineering. New York: McGraw-Hill Book Co., 1966, pp 313–321.
19. DM Pozar. Microwave Engineering. 2nd ed. New York: John Wiley & Sons, 1998.
20. E Nyfors, P Vainikainen. Industrial Microwave Sensors. Norwood, MA: Artech House, Inc., 1989.
21. B Bocquet, JC van de Velde, A Mamouni, Y leroy, G Giaux, J Delannoy, D Delvalee. Microwave radiometric imaging at 3 GHz for the exploration of breast tumors. IEEE Trans. Microwave Theory Tech. MTT-38: 791–793, June 1990.
22. MS Hawley, A Broquetas, L Jofre, JC Bolomey, G Gaboriaud. Microwave imaging of tissue blood content changes. Journal of Biomedical Engineering. **13**: 197–202, 1991.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

23. RA Williams, MS Beck, eds. Process Tomography; Principles, Techniques, and Applications. Oxford, UK: Butterworth-Heinemann, 1995.
24. AJ Bahr, DG Watters. Microwave dielectric sensing methods. In: G Birnbaum and BA Auld, eds. Sensing for materials characterization, processing, and manufacturing. ASNT Topics on Nondestructive Evaluation Series, Columbus, OH **1**: 319–334, 1998.
25. S Ganchev, N Qaddoumi, D Brandenburg, S Bakhtiari, R Zoughi, J Bhattacharyya. Microwave diagnosis of rubber compounds. IEEE Trans. Microwave Theory Tech. MTT-42: 18–24, 1994.
26. S Gray, S Ganchev, N Qaddoumi, G Beauregard, D Radford, R Zoughi. Porosity level estimation in polymer composites using microwaves. Materials Evaluation **53**: 404–408, 1995.
27. HE Bussey. Measurement of RF properties of materials. A survey. Proc. IEEE **55**: 1046–1053, 1967.
28. MT Hallikainen, FT Ulaby, MC Dobson, MA El-Rayes, L-K Wu. Microwave dielectric behavior of wet soil—Part I: Empirical models and experimental observations. IEEE Trans. Geosci. Remote Sensing GE-23: 25–34, 1985.
29. MA Tsankov. Comments on dielectric measurements of sheet materials. IEEE Trans. on Instrum. Meas. IM-23: 248–249, 1974.
30. S Roberts, A von Hippel. A new method for measuring dielectric constants and loss in the range of centimeter waves. J. Appl. Phys. **17**: 610–616, 1946.
31. HM Altschuler. Dielectric constant. In: M Sucher, J Fox, eds. Handbook of Microwave Measurements, Vol. II. Brookline, NY: Polytechnic Press, 1963, Chapter 9.
32. MA Tsankov. Measurable values of permittivity and loss tangent of dielectrics and ferrites for waveguide microwave methods. Bulg. J. Phys. **8**: 403–414, 1981.
33. H Chao. An uncertainty analysis for the measurements of microwave conductivity and dielectric constant by the short-circuited line method. IEEE Trans. Instrum. Meas. IM-35: 36–41, 1986.
34. O Hashimoto, Y Shimizu. Reflecting characteristics of anisotropic rubber sheets and measurement of complex permittivity tensor. IEEE Trans. Microwave Theory Tech. MTT-34: 1202–1207, November 1986.
35. J Jow, MC Hawley, M Finzel, J Asmussen, Jr., H-H Lin, B Manring. Microwave processing and diagnosis of chemically reacting materials in a single-mode cavity applicator. IEEE Trans. Microwave Theory Tech. MTT-35: 1435–1443, 1987.
36. N Qaddoumi, S Ganchev, R Zoughi. Microwave diagnosis of low density glass fibers with resin binder. Research in Nondestructive Evaluation **8**: 177–188, 1996.
37. K Bois, R Mirshahi, R Zoughi. Dielectric mixing models for cement based materials. Review of Progress in Quantitative Nondestructive Evaluation **16A**: 657–663, 1997.
38. KJ Bois, AD Benally, PS Nowak, R Zoughi. Microwave nondestructive determination of sand to cement (s/c) ratio in mortar. Research in Nondestructive Evaluation **9**: 227–238, 1997.
39. J Baker-Jarvis, C Jones, B Biddle, M Janezic, RG Geyer, JH Grosvenor, Jr, CM Weil. Dielectric and magnetic measurements: a survey of nondestructive, quasi-nondestructive, and process-control techniques. Research in Nondestructive Evaluation **7**: 117–136, August 1995.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

40. DK Ghodaonkar, VV Varadan, VK Varadan. A free-space method for measurement of dielectric constants and loss tangents at microwave frequencies. IEEE Trans. Instrum. Meas. IM-37: 789–793, June, 1989.
41. R Zoughi, M Lujan. Nondestructive microwave thickness measurements of dielectric slabs. Materials Evaluation **48**: 1100–1105, 1990.
42. R Zoughi, B Zonnefeld. Permittivity characteristics of Kevlar, carbon composites, E-glass, and rubber (33% carbon) at X-band (8–12 GHz). Review of Progress in Quantitative Nondestructive Evaluation **10B**: 1431–1436, 1991.
43. RE Collin. Field Theory of Guided Waves. 2nd ed. Piscataway, NJ: IEEE Press, 1991, pp 192–199.
44. Data manual for Kevlar aramid 49. Wilmington, DE: Du Pont de Nemours & Co., 47, 1986.
45. RT Compton. The admittance of aperture antennas radiating into lossy media. Ph.D. dissertation, Ohio State University, Columbus, OH, 1964.
46. J Galejs. Antennas in Inhomogeneous Media. Elmstord, NY: Pergamon Press, 1969.
47. S Bakhtiari, S Ganchev, R Zoughi. Open-ended rectangular waveguide for nondestructive thickness measurement and variation detection of lossy dielectric slabs backed by a conducting plate. IEEE Trans. Instrum. Meas. IM-42: 19–24, February, 1993.
48. S Bakhtiari, S Ganchev, N Qaddoumi, R Zoughi. Microwave non-contact examination of disbond and thickness variation in stratified composite media. IEEE Trans. Microwave Theory Tech. MTT-42: 389–395, March, 1994.
49. S Bakhtiari, S Ganchev, R Zoughi. A generalized formulation for admittance of an open-ended rectangular waveguide radiating into stratified dielectrics. Research in Nondestructive Evaluation **7**: 75–87, 1995.
50. S Ganchev, N Qaddoumi, E Ranu, R Zoughi. Microwave detection optimization of disbond in layered dielectrics with varying thicknesses. IEEE Trans. Instrum. Meas. IM-44: 326–328, April, 1995.
51. N Qaddoumi, R Zoughi, GW Carriveau. Microwave detection and depth determination of disbonds in low-permittivity and low-loss thick sandwich composites. Research in Nondestructive Evaluation **8**: 51–63, 1996.
52. S Gray, R Zoughi. Dielectric sheet thickness variation and disbond detection in multilayered composites using an extremely sensitive microwave approach. Materials Evaluation **55**: 42–48, 1997.
53. V Theodoridis, T Sphicopoulos, F Gardiol. The reflection from an open-ended rectangular waveguide terminated by a layered dielectric medium. IEEE Trans. Microwave Theory Tech. MTT-33: 359–366, May 1985.
54. FE Gardiol, et. al. The reflection of open-ended circular waveguide—application to nondestructive measurement of materials. Reviews of Infrared and Millimeter Waves, KJ Button, ed., vol. 1. New York: Plenum Press, 1983, pp 325–364.
55. S Ganchev, N Qaddoumi, S Bakhtiari, R Zoughi. Calibration and measurement of dielectric properties of finite thickness composite sheets with open-ended coaxial sensors. IEEE Trans. Instrum. Meas. IM-44: 1023–1029, December, 1995.
56. RJ Botsco. Nondestructive testing of plastics with microwaves. Materials Evaluation **27**: 25A–32A, June 1969.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

57. R Zoughi, S Bakhtiari. Microwave nondestructive detection and evaluation of disbonding and delamination in layered-dielectric slabs. IEEE Trans. Instrum. Meas. IM-39: 1059–1063, December, 1990.
58. R Zoughi, S Bakhtiari. Microwave nondestructive detection and evaluation of voids in layered dielectric slabs. Research in Nondestructive Evaluation **2**: 195–205, 1990.
59. CA Balanis. Antenna Theory, Analysis and Design. New York: John Wiley and Sons, 1977.
60. R Zoughi, GL Cone, and P Nowak. Microwave nondestructive detection of rebars in concrete slabs. Materials Evaluation **49**: 1385–1388, November, 1991.
61. N Qaddoumi, SI Ganchev, G Carriveau, R Zoughi. Microwave imaging of thick composites with defects. Materials Evaluation **53**: 926–929, August, 1995.
62. SI Ganchev, RR Runser, N Qaddoumi, E Ranu, GW Carriveau. Microwave nondestructive evaluation of thick sandwich composites. Materials Evaluation **53**: 463–467, April 1995.
63. S Bakhtiari, N Gopalsami, AC Raptis. Characterization of delamination and disbonding in stratified dielectric composites by millimeter wave imaging. Materials Evaluation **53**: 468–471, April 1995.
64. N Qaddoumi, A Shroyer, R Zoughi. Microwave detection of rust under paint and composite laminates. Research in Nondestructive Evaluation **9**: 201–212, 1997.
65. DW Radford, BW Barber, S Ganchev, R Zoughi. Microwave nondestructive evaluation of glass fiber/epoxy composites subjected to impact fatigue. Proceedings of SPIE Symposium, Advanced Microwave and Millimeter Wave Detectors Conference, San Diego CA. **2275**: 21–26, July 24–29, 1994.
66. JK Bolomey, N Joachimowicz. Dielectric metrology via microwave tomography: present and future. Materials Research Society Symposium Proceedings **347**: 259–268, 1994.
67. C Huber, H Abiri, S Ganchev, R Zoughi. Modeling of surface hairline crack detection in metals under coatings using open-ended rectangular waveguides. IEEE Trans. Microwave Theory Tech. MTT-45: 2049–2057, November 1997.
68. KJ Bois, H Campbell, AD Benally, PS Nowak, R Zoughi. Microwave noninvasive detection of grout in masonry. Masonry Journal **16**: 49–54, June 1998.
69. KJ Bois, AD Benally, PS Nowak, R Zoughi. Cure-state monitoring and water-to-cement ratio determination of fresh portland cement based materials using near-field microwave techniques. IEEE Trans. Instrum. Meas. IM-47: 628–637, June 1998.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

10

Optical Methods

Donald D. Duncan, John L. Champion, Kevin C. Baldwin, and David W. Blodgett
Johns Hopkins University, Laurel, Maryland

10.1 INTRODUCTION

10.1.1 Optical Techniques Overview

In this chapter, we deal with optical techniques in *nondestructive evaluation* (NDE). This field falls within the general category of using light to discern certain properties of matter. Figure 10.1 illustrates this concept of using light in this manner. This illustration also serves to describe the general approach of the chapter. We begin with a description of the properties and behavior of light, for instance, how light propagates though space as well as a dielectric medium. Incidentally, the use of the term light in no way restricts our attention to the visible spectrum. Optical engineers refer to electromagnetic energy in general as "light."

The next step is the interaction of light with matter, i.e., the matter that we wish to obtain some information about. This interaction could be at the macroscopic level, where the wavelength is small compared to the structure being probed (macroscopic) or comparable to the structure (microscopic). In some cases it can be both simultaneously. For example we may be interested in the color and shape of an object. The first property is primarily related to the

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

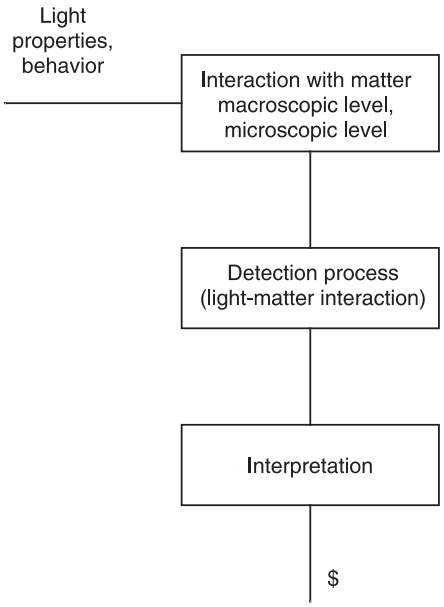


FIGURE 10.1 Illustration of NDE concept.

molecular and microstructural properties of the object (wavelength level), while features of the shape may be many wavelengths.

Next comes the detection process. This itself is a light–matter interaction. For instance in a photodetector, due to the photoelectric effect, an incident photon produces a free electron. Traditionally, many optical NDE techniques involved photographic films. This role in many applications has been supplanted by solid state detectors, e.g., *charge-coupled device* (CCD) cameras. Nevertheless, film still plays an important role. A discussion of film is a convenient way to introduce a number of NDE concepts. Interestingly, the detection process itself plays an important role because it's nonlinear. Specifically, physical detectors are sensitive to the square of the electric field. As such they perform a "mixing" operation that is critical to a number of important NDE schemes.

The final step in our NDE process is that of interpretation. One may assume that once the data are acquired, the job is finished. Wrong! In fact, the interpretation phase is probably the most important part of the problem. At this point in the whole process we attempt to answer such questions as: What does this measurement tell me about the object being probed? Do the data address the problem at hand? Is the precision of the measurement appropriate? What is the

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

accuracy of the measurement? What is the sensitivity of the measurement? The answers to all of these questions flow out of the interpretation process.

The specific properties of light that we make use of may be first order, for example, intensity, wavelength, or polarization. Or they may be second order properties such as the correlation of different properties of light (field strength, intensity, etc.) at two different points in space or time (or both). Properties of the object that we are interested in discerning may be morphology (shape or surface texture), motion and deformation, color, temperature, chemical composition, or pressure, to name a few. Throughout this process, we are in a sense, limited by certain properties of light. Similarly, we are limited in that there are only certain specific ways that light interacts with matter. The art, however, is in how we combine these two, and how the results are interpreted. In this chapter, no attempt will be made to give a comprehensive view of this broad interdisciplinary field, but merely to indicate what is possible and to illustrate a couple very powerful optical NDE techniques.

One may have the preconceived notion that optical NDE techniques are applied only in a laboratory or possibly a manufacturing environment. While applications abound in these situations, they have much broader venues. In fact, one may place these applications under the general category of optical remote sensing. Furthermore, the word "optical" has particular meaning to different people. For example to one concerned with human vision, optical refers to wavelengths in the range of 400–700 nm. To the optical engineer, however, optical refers more generally to any wavelengths that can be manipulated with mirrors, lenses, prisms, etc. These wavelengths exceed the visible spectrum on both the short and long side. Finally, to the electrical engineer, optical may refer to the microwave regime or longer. In fact, many of the concepts that have been developed for use in the visible have been generalized to entirely different wavelength regimes.

10.1.2 Historical Perspective

If one had to describe the field of optical nondestructive testing with one word, it would be "interdisciplinary." It represents contribution from many fields other than optics, such as electrical and mechanical engineering, signal processing, and material science. That the study of optics is one of the oldest sciences should come as no surprise. After all, vision is one of man's most important senses. Some of the earliest systematic writings on optics were by the Greek philosophers and mathematicians Empedocles (C490-430BCE) and Euclid (C300BCE). Light has long been used as a tool for measurement, but of course when these techniques were developed, there was no such thing as a laser. The invention of the laser in 1960 provided an additional tool that in many cases allowed greater precision, and certainly provided a greater convenience.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

The field of course has had contributions from many fields of engineering and science if for no other reason than as an impetus to exploit the potential of optical techniques to learn more about the environment, intrinsic material properties, and behavior of structures under stress.

That the field has had contributions from the field of electrical engineering may come as a surprise. The modern theory of optical systems makes heavy use of Fourier analysis (Goodman, 1996) and concepts that were traditionally used to characterize electrical circuits. In fact much of the modern nomenclature is borrowed from this field. Much of the underlying mathematical formalism is also shared by modern data processing techniques.

10.1.3 About this Chapter

In this chapter we discuss primarily whole-field techniques, that is, techniques in which an image is formed of the object under test. These are distinct from "point" measurements in which some characteristic of the object is probed at a single point.

One of the common themes in this chapter is interference and many modern NDE techniques are termed interferometric. That is, they rely upon the interference of light with itself. One of the earliest examples might be the use of Newton's fringes (Jenkins and White, 1957) to assess the flatness of objects. Interferometric techniques have become especially important since the advent of the laser provided a convenient high power source of coherent light. In most cases this interference takes place during the detection process as alluded to previously. Holographic and most speckle techniques rely on this coherence. The concept of holography was invented in 1949 by Dennis Gabor (Gabor, 1949) prior to the laser, but subsequently made practical using the approaches developed in the early 1960s in the U.S. by Leith and Upatnieks (Leith and Upatnieks, 1962) and in the Soviet Union by Denisyuk (Denisyuk, 1962).

Many traditional NDE techniques, however, are not usually thought of as being interferometric in nature. An example is the use of "structured" light (Gåsvik, 1995). Nevertheless, interference concepts play prominently in the processing of the resulting data. Here however, the interference is effected through nonlinear computational techniques. These processing techniques rely heavily on concepts from Fourier analysis and depend on the power provided by personal computers.

No attempt is made to give a comprehensive treatment of all the myriad whole-field optical NDE techniques, but merely to cite a few representative techniques that serve to illustrate the power and versatility of this interdisciplinary study. Because many of the techniques require a fair understanding of the behavior of light we have included a section on some of these basics. The more knowledgeable reader may proceed directly to Section 10.3 on optical

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

techniques. Throughout, an attempt has been made to minimize the mathematics (though not at the sacrifice of rigor) and to appeal to heuristic arguments. A good example of a text in this vein is *Fundamentals of Optics* (Jenkins and White, 1957). For the reader new to the field of optics, other good references are *Introduction to Modern Optics, Second Edition* (Fowles, 1975) and *Optics, Third Edition* (Hecht, 1997).

10.2 THEORY

10.2.1 Basic Properties of Light

The Wave Equation

We begin with the wave equation that describes propagation of electromagnetic energy in free space:

$$\nabla^2 u(R,t) - \frac{1}{c^2}\,\frac{\partial u(R,t)}{\partial t} = 0 \tag{1}$$

where the Laplacian is defined in Cartesian coordinates as

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{2}$$

we have used the generic variable u to denote a single component of the electric or magnetic field, and c is the speed of light in a vacuum. If we restrict our attention to a strictly monochromatic (single color) signal we can express the field as,

$$u(R,t) = \big|U(R)\big| \cos[\phi(R) - \omega t] \tag{3}$$

This is more commonly expressed as

$$\begin{aligned} u(z,t) &= \mathrm{Re}\big\{\big|U(R)\big| e^{i[\phi(R)-\omega t]}\big\} \\ &= \mathrm{Re}\big\{U(R) e^{-i\omega t}\big\} \end{aligned} \tag{4}$$

where "Re" denotes that we're taking the real part. Further, this explicit notation is usually suppressed with the understanding that it is the real part that we are talking about. With this expression for the field, the wave equation takes on the reduced form known as the Helmholtz equation:

$$\big(\nabla^2 + k^2\big) U(R) = 0 \tag{5}$$

where

$$U(R) = \big|U(R)\big| e^{i\phi(R)} \tag{6}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

and the wavenumber is defined as

$$k = \frac{2\pi}{\lambda} \tag{7}$$

There are a number of exact solutions to the wave equation, two of which we will discuss in the material that follows. For example, it can be seen that a plane wave,

$$u(z, t) = U_0 \cos(kz - \omega t) \tag{8}$$

is an exact solution. One can think of this solution in terms of a fixed instant in time or at a fixed point in space. The former is illustrated in Figure 10.2. It consists of a series of planes of constant phase that are perpendicular to the z-axis. The separation between the planes of integer phase can be determined as follows:

$$\begin{aligned} kz_n - \omega t &= n\,2\pi \\ kz_{n+1} - \omega t &= (n+1)\,2\pi \\ z_{n+1} - z_n &= \frac{2\pi}{k} = \lambda \end{aligned} \tag{9}$$

If one selects the equation for a single plane,

$$kz - \omega t = \text{Const.} \tag{10}$$

and differentiates both sides with respect to time, one obtains

$$k\,\frac{dz}{dt} - \omega = 0 \tag{11}$$

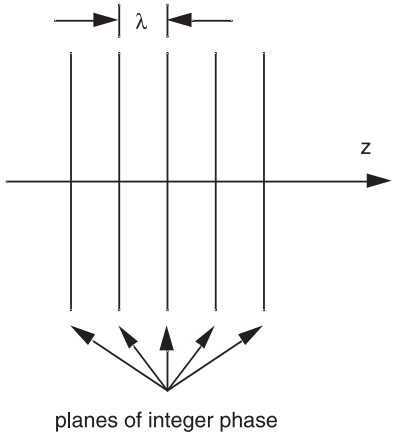


FIGURE 10.2 Illustration of a plane wave propagating along the z-axis.

Copyright © Marcel Dekker, Inc. All rights reserved.

Clearly, the derivative is a velocity, so that we have the important relationship

$$c = \frac{1}{k}\,\omega = \frac{\lambda}{2\pi}\,2\pi\nu = \lambda\nu \tag{12}$$

This result states that these planes of constant phase move along the z-axis at the velocity of light in free space, c. In dielectric media, the velocity is lower, and we describe this as follows:

$$\frac{c}{v} = n \tag{13}$$

where n is the *refractive index* of the medium.

If one were to choose a point in space and observe the field as time progresses, one would observe harmonic variation as illustrated in Figure 10.3. On the other hand if one could "ride along with" the wave it would appear static, much as a surfer perceives his wave to be more or less fixed.

In the material of this chapter, we will usually be discussing strictly monochromatic fields. We will therefore suppress the time dependence. In this notation, the general form of the plane wave is seen to be

$$U(R) = U_0\, e^{i\overline{\kappa}\cdot\overline{R}} \tag{14}$$

where $\overline{\kappa}$ is the direction of propagation and $\overline{R}$ is referred to as the *field point* (the point in space at which we wish to determine the field). The geometry is illustrated in Figure 10.4. The scalar dot product in the exponent of this equation can be rewritten

$$\overline{\kappa}\cdot\overline{R} = k\left(\hat{x}x\,\frac{\kappa_x}{k} + \hat{y}y\,\frac{\kappa_y}{k} + \hat{z}z\,\frac{\kappa_z}{k}\right) \tag{15}$$

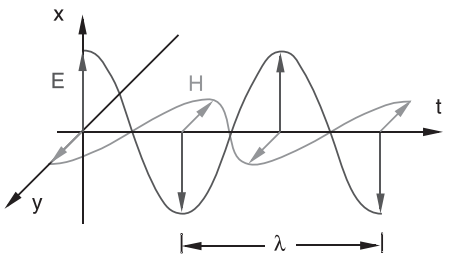


FIGURE 10.3 E and H fields for linear polarization as a function of time.

Copyright © Marcel Dekker, Inc. All rights reserved.

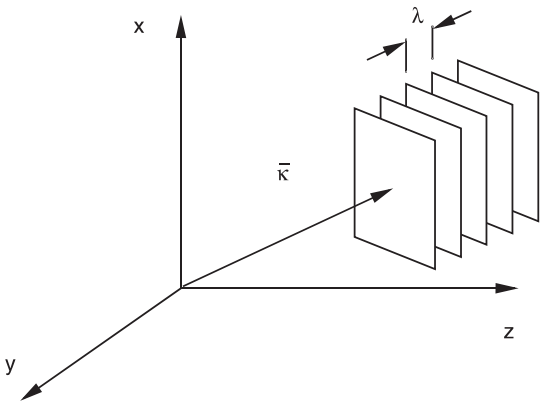


FIGURE 10.4 The wave vector.

The quotients on the right hand side of this equation are seen to be the projections of the wave vector onto the respective axes. For this reason, they are called the *direction cosines* (see Figure 10.5) and are usually denoted as follows:

$$\left(\frac{\kappa_x}{k}, \frac{\kappa_y}{k}, \frac{\kappa_z}{k}\right) = (\cos\theta_1, \cos\theta_2, \cos\theta_3) = (\alpha, \beta, \gamma)$$
$$\gamma = \sqrt{1 - \alpha^2 - \beta^2} \tag{16}$$

The direction cosines are often used to specify the direction of propagation.

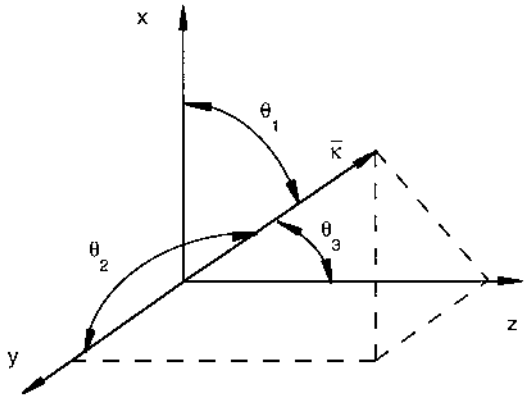


FIGURE 10.5 Direction cosines.

Copyright © Marcel Dekker, Inc. All rights reserved.

Another important exact solution to the wave equation is the spherical wave,

$$U(R) = U_0 \frac{e^{ikR}}{R} \tag{17}$$

This field consists of a series of concentric spherical phase surfaces separated one wavelength apart and with an amplitude that falls off inversely with distance. The general form of the spherical wave for a point source located at a point $\overline{R}_0$ is

$$U(R) = U_0 \frac{e^{ik|\overline{R}-\overline{R}_0|}}{|\overline{R}-\overline{R}_0|} \tag{18}$$

We've been discussing the EM field in terms of a wave, although we've hinted at the ray nature of light. The wave vector is simply the local normal to the wavefront. Of course for a plane wave, this wavefront normal is a constant. For a spherical wave, however, it is a function of location on the wavefront (see Figure 10.6). In other words, it is a point property of the field.

This description of light has been of a single component of the transverse electric field, i.e., for linear polarization. For a more complete description of the general polarization state, see Fowles (1975) or O'Neill (1963) for a discussion of the Jones vector description of light.

10.2.2 Interference

Optical detectors, the human eye for example, do not respond directly to the electromagnetic field amplitude, but instead to the optical power. Intensity is a measure of the optical power per unit area and is proportional to the time averaged squared magnitude of the field,

$$I(x, y, z) \propto \langle |u(x, y, z, t)|^2 \rangle \tag{19}$$

where the time period over which the averaging (angular brackets) is performed is long compared to the temporal field variation. In the field of optics it is customary to ignore this constant of proportionality and to define the "intensity" as the time averaged squared magnitude of the field. (The proportionality becomes an equality.)

Now suppose that we have the situation illustrated in Figure 10.7, two plane waves propagating at angles $\pm\theta$ with respect to the z-axis and striking the observation plane, $z = 0$. We will first calculate the intensity pattern using this definition of intensity and the cosine behavior of the field. Afterward, we will repeat this using the complex definition of the field. This latter method is much simpler and arrives at the same result. Each of these plane waves satisfies the

Copyright © Marcel Dekker, Inc. All rights reserved.

FIGURE 10.6 Wavefront normals and ray directions.

Copyright © Marcel Dekker, Inc. All rights reserved.

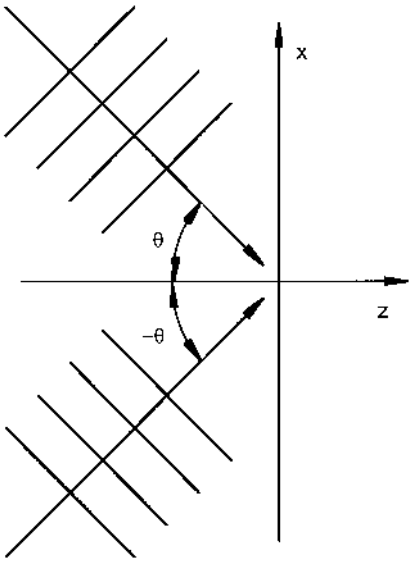


FIGURE 10.7 Two intersecting plane waves.

Helmholtz equation. Thus, due to linearity, so does their sum. The total field in the observation plane is therefore,

$$\begin{aligned} u(x,t) &= U_0 \cos\left(\overline{\kappa}_1 \cdot \overline{R} - \omega t\right) + U_0 \cos\left(\overline{\kappa}_2 \cdot \overline{R} - \omega t\right) \\ \overline{\kappa}_1 &= \hat{x}k \sin\theta + \hat{z}k \cos\theta \\ \overline{\kappa}_2 &= -\hat{x}k \sin\theta + \hat{z}k \cos\theta \\ \overline{R} &= \hat{x}x \end{aligned} \tag{20}$$

Combining Eqs. (19) and (20) yields

$$\begin{aligned} I(x,t) = {} & \langle U_0^2 \cos^2(\overline{\kappa}_1 \cdot \overline{R} - \omega t)\rangle + \langle U_0^2 \cos^2(\overline{\kappa}_2 \cdot \overline{R} - \omega t)\rangle \\ & + 2\langle U_0^2 \cos(\overline{\kappa}_1 \cdot \overline{R} - \omega t)\cos(\overline{\kappa}_2 \cdot \overline{R} - \omega t)\rangle \end{aligned} \tag{21}$$

Now we explicitly introduce the time averaging operation. For example, the first term on the right-hand side of Eq. (21) is

$$\langle U_0^2 \cos^2\left(\overline{\kappa}_1 \cdot \overline{R} - \omega t\right)\rangle = \frac{1}{T}\int_{-T/2}^{T/2} U_0^2 \cos^2\left(\overline{\kappa}_1 \cdot \overline{R} - \omega t\right)\, dt \tag{22}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

As stated earlier, the time, T, over which this integration takes place is long compared to $1/v$. It is left as a homework problem to show that by using Eq. (22), the intensity can be expressed as

$$I(x) = U_0^2[1 + \cos(2k \sin\theta)] \tag{23}$$

All of this required a number of manipulations. We contrast this approach with that using the complex representation of the fields. With this approach, the intensity in the observation plane is

$$\begin{aligned} I(x) &= \left|U_0 e^{i(kx\sin\theta - \omega t)} + U_0 e^{i(-kx\sin\theta - \omega t)}\right|^2 \\ &= \left|2U_0 \cos(kx\sin\theta)\right|^2 \\ &= 2U_0^2[1 + \cos(2kx\sin\theta)] \end{aligned} \tag{24}$$

Comparing Eqs. (23) and (24) we see that the only difference is the factor of 2. Thus we see the utility of the complex notation.

Looking more closely at Eq. (35) reveals that the intensity distribution in the observation plane is a sinusoid with period given by

$$\begin{aligned} &2kx\sin\theta|_{x=T} = 2\pi \\ &T = \frac{\lambda}{2\sin\theta} \end{aligned} \tag{25}$$

One can gain an appreciation for the sensitivity of these fringes to the angle separating the two plane waves if we look at the spatial frequency,

$$f \equiv \frac{1}{T} = \frac{2\sin\theta}{\lambda} \tag{26}$$

for the case of 10 cycles/mm. This is roughly the highest spatial frequency that can be observed easily with the naked eye. For a HeNe laser ($\lambda = 0.6328\,\mu$m), the corresponding angular separation is

$$2\theta = 2\sin^{-1}\left(\frac{\lambda f}{2}\right) = 0.73^\circ \tag{27}$$

To put this in perspective, we note that the angular subtense of the sun as seen from earth is about 0.53°.

Young's Double Slit Experiment

A classic experiment that is useful for describing in a quantitative fashion a number of interference phenomena is Young's double slit problem. Two apertures in an otherwise opaque screen (see Figure 10.8) are illuminated with a coherent source such as a laser (see Appendix A for a discussion of coherence). The fields in an observation plane some distance behind the screen are calculated by treating the apertures as secondary field sources. The field at the indicated observation

Copyright © Marcel Dekker, Inc. All rights reserved.

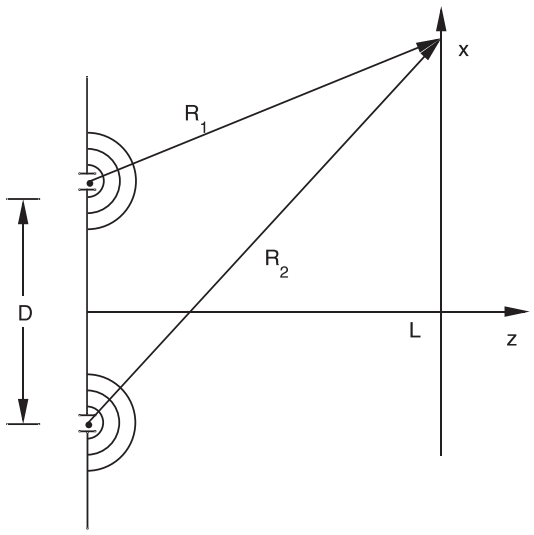


FIGURE 10.8 Young's double slit experiment.

point due to the equivalent point sources within the apertures of the opaque screen is given by the following:

$$U(x, L) = \frac{e^{ikR_1}}{R_1} + \frac{e^{ikR_2}}{R_2}$$
$$R_1 = \sqrt{(x - D/2)^2 + L^2} \tag{28}$$
$$R_2 = \sqrt{(x + D/2)^2 + L^2}$$

If we use the paraxial approximation that assumes the distances D and x are small with respect to the observation distance, L, then we can write

$$R_1 \approx L + \frac{(x - D/2)^2}{2L}$$
$$R_2 \approx L + \frac{(x + D/2)^2}{2L} \tag{29}$$

These approximations (simply the first two terms of the Taylor series expansion for the distance) are appropriate in the exponents of the spherical waves in Eq. (28). In the denominator we can make the even stronger approximation that

$$R_1 \approx R_2 \approx L \tag{30}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

The intensity that corresponds to this field is given by

$$I(x, L) = \left|U(x, L)\right|^2 = \left|\frac{e^{ikR_1}}{L} + \frac{e^{ikR_2}}{L}\right|^2$$

$$= \frac{1}{L^2}\left[2 + 2\cos k\,(R_1 - R_2)\right]$$

$$\approx \frac{1}{L^2}\left[1 + \cos\left(\frac{k\,x\,D}{L}\right)\right] \tag{31}$$

As seen in the last equation above, one observes a fringe pattern with spatial frequency

$$T = \frac{\lambda L}{D} \tag{32}$$

Actually, this is the same result that we found previously under the assumption of an angular separation between the equivalent sources:

$$\theta = \tan^{-1}\left(\frac{D}{2L}\right) \approx \frac{D}{2L} \approx \sin\theta \tag{33}$$

The Michelson Interferometer

The Young's double slit configuration is a member of a class of interferometers called *wavefront splitting*; the other general class of interferometers is termed *amplitude spitting*. A classic example of this class is the Michelson, illustrated in

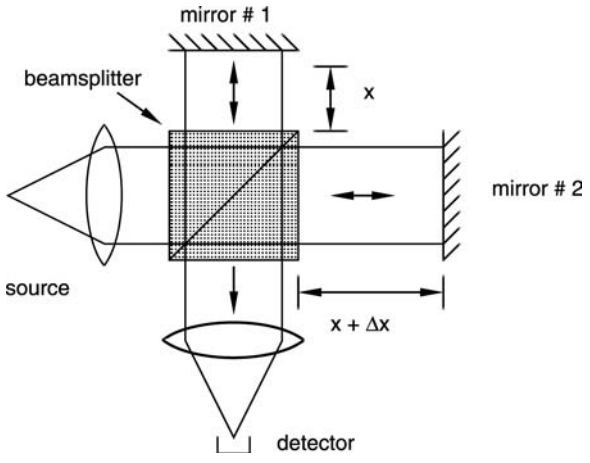


FIGURE 10.9 The Michelson interferometer.

Copyright © Marcel Dekker, Inc. All rights reserved.

Figure 10.9. As shown in this figure, a collimated beam (a plane wave) of light is incident on a beamsplitter that divides the amplitude of the incident field into two. One attenuated portion continues straight ahead, while the other portion is reflected at a right angle. Each of these beams strikes a mirror and is reflected back to the beamsplitter. The same process of splitting amplitudes is repeated with the result that the beam that has struck mirror number one is recombined with the beam that has struck mirror number two. If we assume that the beamsplitter is of a type called 50 : 50, then the amplitudes of the transmitted and reflected beams are equal. In this case the total field striking the detector is given by

$$U(x, t) = \frac{U_0}{4}\, e^{i(kz-\omega t)} + \frac{U_0}{4}\, e^{i[k(x+\Delta x)-\omega t]} \tag{34}$$

The detected intensity is

$$\begin{aligned} I(x, t) &= \left| \frac{U_0}{4}\, e^{i(2kx-\omega t)} + \frac{U_0}{4}\, e^{i[2k(x+\Delta x)-\omega t]} \right|^2 \\ &= 2\left(\frac{U_0}{4}\right)^2 (1 + \cos 2k\Delta x) \end{aligned} \tag{35}$$

The behavior of this signal as the path difference is varied is shown in Figure 10.10. As illustrated, the period of the variation is

$$\begin{aligned} 2k\Delta x|_{\Delta x=T} &= 2\pi \\ T &= \frac{\lambda}{2} \end{aligned} \tag{36}$$

One use of this type of interferometer is the measurement of distance. If one counts the number of periods in the output signal, one can infer the relative

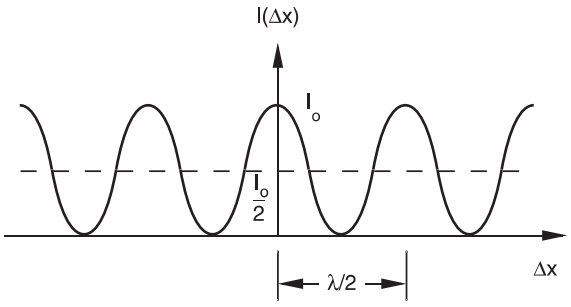


FIGURE 10.10 Detector signal as a function of path difference.

Copyright © Marcel Dekker, Inc. All rights reserved.

motion of the mirrors. For example if as Δx changes, we count n fringes, then we infer that

$$\Delta x = n\frac{\lambda}{2} \tag{37}$$

If we were to count the number of fringes in a certain period of time, we calculate

$$\begin{aligned} 2k\ \frac{\Delta x}{\Delta t} &= n\frac{2\pi}{\Delta t} \\ \frac{\Delta x}{\Delta t} &= \frac{\lambda}{2}\ \frac{n}{\Delta t} \end{aligned} \tag{38}$$

The left-hand side of this equation is clearly the velocity of the mirror, and the quotient $n/\Delta t$ on the right-hand side is a frequency (with which the fringes are counted), so that we can rewrite Eq. (38) as

$$v = \frac{\lambda}{2}\ \nu \tag{39}$$

or by rearranging,

$$\nu = \nu_D = \frac{2v}{\lambda} \tag{40}$$

This is recognized as the expression for a Doppler shift (Fowles, 1975). The conclusion that we arrive at is that the measurement of position and the measurement of velocity using this instrument are conceptually equivalent. The difference arises only in the method of processing the detector signal. One can get an appreciation for the sensitivity of this technique by considering a velocity of 3 mph (walking speed) and a wavelength of 0.6328 μm (HeNe laser). The resulting Doppler shift is 4.24 MHz!

10.2.3 Imaging Systems

The modern concept of an imaging device is that of a "black box." That is to say, the system is characterized in terms of its terminal properties, much in the way an electronic circuit is characterized in terms of its input and output terminal characteristics. In the case of an imaging system, these terminals are the input and output *pupils* (Smith, 1966). The pupils are defined as those physical apertures or their images that serve to limit the angular cone of rays that enter or leave the imaging system. As an example, consider the simple thin lens imaging system shown in Figure 10.11. Here, the finite nature of the lens serves as the physical aperture stop. It is obviously both the entrance and exit pupil because it simultaneously limits the angular cone of rays entering and exiting the system.

Now, however, consider Figure 10.12. Here the physical aperture stop is placed in front of the lens. It obviously limits the angular extent of the cone of

Copyright © Marcel Dekker, Inc. All rights reserved.

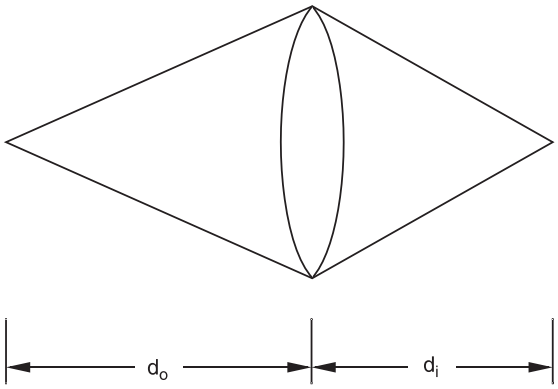


FIGURE 10.11 A simple thin lens.

rays entering the system. As such, it defines the entrance pupil of the system. However, this physical stop does not serve as the exit pupil. Rather, the image of this stop serves as the exit pupil. Again, we use the thin lens imaging rule to determine the location of this image:

$$\frac{1}{s_o}+\frac{1}{s_i}=\frac{1}{f} \tag{41}$$

In this case, however, the object distance is less than the focal length. As a result, the image distance is negative, i.e., the image is "virtual." As illustrated in Figure 10.12, this virtual image of the aperture stop serves as the exit pupil. A similar condition is observed when the physical aperture stop is placed behind the lens.

In a general imaging system, there will be some physical aperture stop. The exit pupil is the image formed of this stop by any optical elements that follow it (elements B in Figure 10.13), and the entrance pupil is the image of this stop formed by elements that precede it. From this we gather that the entrance and exit pupils are images of one another. As such, one could characterize the system in terms of either; convention has is that the exit pupil is used. These pupils can be located within the optical instrument or outside. For example when one looks through a microscope or a pair of binoculars, one must position his eye correctly so that the pupil of the eye corresponds to the exit pupil of the device. By moving one's head slightly the edges of the exit pupil can be seen. In this case the exit pupil is outside the optical instrument.

The Point-Spread Function

Simply put, the *point-spread function* (PSF) is the image of a geometrical point object. The analog in terms of an electrical circuit is the *impulse response*. A

Copyright © Marcel Dekker, Inc. All rights reserved.

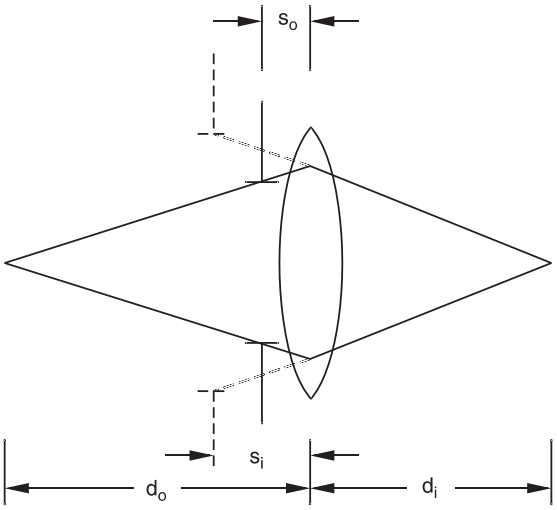


FIGURE 10.12 Thin lens with aperture stop in front.

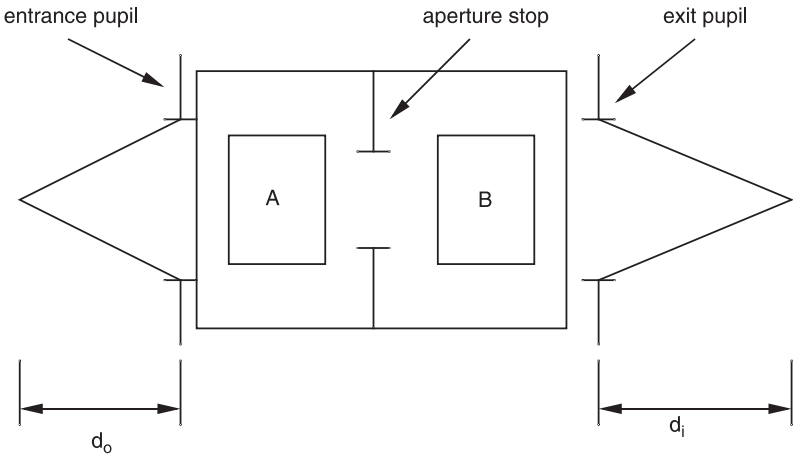


FIGURE 10.13 The "black box" concept of an imaging system.

physicist on the other hand, would refer to this as a Green's function. This nomenclature explicitly refers to the fact that the image of a point object is not of zero dimension, but due to diffraction is spread out. For such a geometrical point object, the PSF is the square modulus of the Fourier transform of the fields in the exit pupil of the imaging system (Goodman, 1996). Details of this concept are

Copyright © Marcel Dekker, Inc. All rights reserved.

unnecessary as the physics of the problem can be understood easily by resorting to the Young's double slit experiment.

Another way of looking at the relationship between the fields in the exit pupil and the image is to realize that the exit pupil represents the apparent source of the fields that converge to points in the image plane. One can imagine that the image is made up of a whole series of source pairs in the exit pupil (see Figure 10.14), each one of which produces a fringe pattern in the image plane. These source pairs are at all possible separations and azimuthal orientations in the exit pupil. The resulting image is built up of all the corresponding fringe patterns. From what we have learned from the Young's double slit experiment, we know that the smallest fringe period will be due to the source pair with the largest separation. This smallest fringe period represents the smallest spatial structure that will appear in the image. From this argument, we deduce that the PSF has the general lateral dimension of

$$PSF \sim \frac{\lambda d_i}{D} \tag{42}$$

where D is the diameter of the exit pupil. Taking account of any image magnification,

$$m = \frac{d_i}{d_o} \tag{43}$$

allows us to write

$$PSF \sim (1 + m)\lambda f^{\#} \tag{44}$$

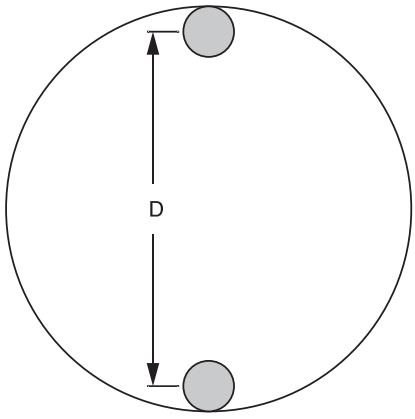


FIGURE 10.14 Illustration of subapertures (shaded) within exit pupil.

Copyright © Marcel Dekker, Inc. All rights reserved.

where $f^{\#}$ is the *f-number* of the system,

$$f^{\#} \equiv \frac{f}{D} \tag{45}$$

This result says that for a geometrical point object, the corresponding image will be spread out to a spot of approximate dimension,

$$\text{size of image of point object} \sim (1 + m)\lambda f^{\#} \tag{46}$$

Generally, then we see that imaging is a low-pass filtering operation; although the object may possess spatial structures that are very small scale, the corresponding image will have spatial structures no smaller than the PSF of the system.

We have discussed the PSF in terms of interference. The alternative view is in terms of diffraction. In this treatment, the PSF is simply the far-field pattern of the pupil (given by the Fourier transform of the pupil plane fields (Goodman, 1996).

10.2.4 Holography

What Is It and Why Should I Care?

Simply put, holography is the recording of interference patterns. You may justifiably ask yourself, "What possible use could recording interference patterns have?" As will be shown, by recording the interference of two (or more) wavefronts, it is possible to recreate both the amplitude, as well as the phase, distribution of the wavefront(s). We shall see that the phase of a wavefront is the key to optical NDE.

Absorption Holograms

When one thinks about a piece of film, they initially think of a piece of "paper" onto which a scene is imaged. After being "exposed" to the image, the film is then developed, a process in which the different shades of light that where incident on the film during the exposure become shades of gray (or color) on that piece of film. Holography uses film in a similar manner, to capture and preserve the "shades of gray" or, in this case, an irradiance distribution caused by the interference of light. By preserving an irradiance distribution, it is possible to reconstruct the wavefronts that were mixed to create that distribution.

Consider interfering two planar wavefronts, U_1 and U_2 in space (see Figure 10.7):

$$\begin{aligned} U_1(x, y, z) &= A \exp\{i(k \sin\theta_1 x + k \cos\theta_1 z - \omega t)\} \\ U_2(x, y, z) &= B \exp\{i(k \sin\theta_2 x + k \cos\theta_2 z - \omega t)\} \end{aligned} \tag{47}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

A piece of film is placed at the position, $z = 0$. The inference pattern that is "seen" by the film is

$$I(x) = \left|U_1(x, 0) + U_2(x, 0)\right|^2 = |A|^2 + |B|^2 + 2\,AB\cos[k(\sin\theta_1 - \sin\theta_2)x] \tag{48}$$

One of the differences between "conventional" film and holographic film is that holographic film is commonly used in transmission. Much like a film negative must be held up to the light to view an image, a hologram must be placed in a beam to recreate wavefronts. As with normal film negative, the transmission after development will be proportional to the irradiance distribution that the film was exposed to;

$$T(x, y, z) \propto I(x, y, z) \tag{49a}$$

or

$$T = T_0 + T_1\cos[k(\sin\theta_1 - \sin\theta_2)x] \tag{49b}$$

For the interference pattern in Eq. (48), the transmission will vary with an amplitude of T_1, about a DC offset, T_0. The transmission value, T, is normalized to a value between 0 (blocking all light) and 1 (allowing all incident light to pass). How the irradiance is mapped to transmission is a function of the film used and the period of time that the film is exposed to the irradiance pattern. This mapping is typically described by the Hurter-Driffield (H-D) curve. The typical form of the curve, shown in Figure 10.15, is a plot of the optical density, D, as a function of the logarithm of the exposure, E (which is the product of the irradiance and the duration of the exposure).

The irradiance transmission of the film is related to the optical density by

$$T = \mathrm{e}^{-D} \tag{50}$$

As seen in the figure, the response of the film is generally highly nonlinear. Of interest to holography is the T–E (transmission–exposure) curve, which is derived from the H-D curve. An example curve is shown in Figure 10.16. In order to produce a reconstructed wavefront that does not suffer from distortions, the film should be exposed for a time duration such that, any excursions of the interference pattern will be within a "linear region" of the T–E curve. By the term "linear region" is it meant that the curve can be reasonably described by the equation of a straight line. In this case, a line of slope, β, has been drawn on the plot. By staying in the linear region of the curve, it is ensured that the transmission of the film reflects the recorded interference pattern with reasonable fidelity.

By regrouping the term describing the interference pattern (Eq. (48)),

$$I(x) = C\{1 + M\cos[k(\sin\theta_1 - \sin\theta_2)x]\} \tag{51a}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

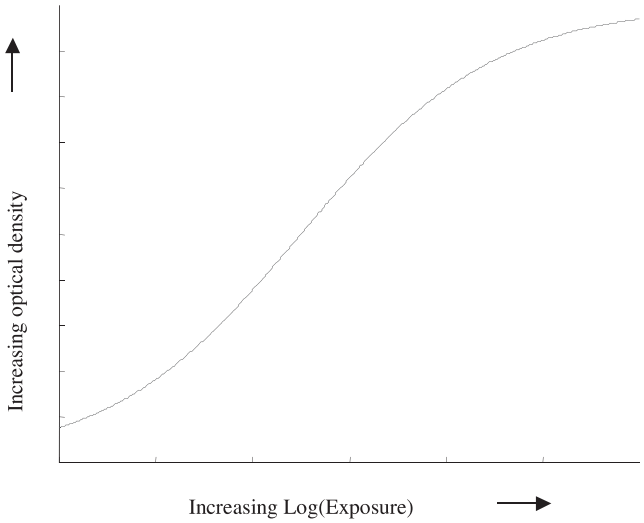


FIGURE 10.15 A typical Hurter-Driffield (H-D) curve.

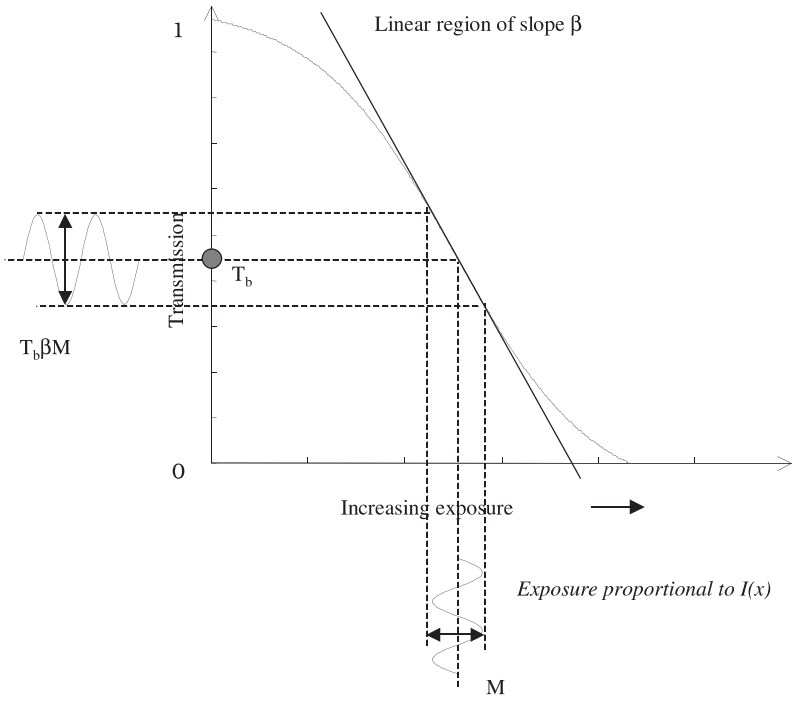


FIGURE 10.16 The T-E curve for photographic film.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

where

$$C = |A|^2 + |B|^2 \quad \text{and} \quad M = 2\frac{AB}{|A|^2 + |B|^2} \tag{51b}$$

one starts down a path which will indicate the significance of the linearity. The associated transmission pattern can be written as

$$T = T_b\{1 + \beta M \cos[k(\sin\theta_1 - \sin\theta_2)x]\} \tag{52}$$

Now, illuminate the film with one of the wavefronts used in the exposure, U_1. The wavefronts that exit the hologram are given by

$$U_1(x, 0)T = U_1(x, 0)T_b\left\{1 + \frac{\beta M}{2}\exp[ik(\sin\theta_1 - \sin\theta_2)x] + \frac{\beta M}{2}\exp[-ik(\sin\theta_1 - \sin\theta_2)x]\right\} \tag{53}$$

By expanding the expression, the significance of the terms becomes more apparent;

$$\begin{aligned} U_1 T = {} & AT_b \exp\{ik\sin\theta_1 x\} \\ & + \frac{AT_b\beta M}{2}\exp\{ik(2\sin\theta_1 - \sin\theta_2)x\} \\ & + \frac{AT_b\beta M}{2}\exp\{ik\sin\theta_2 x\} \end{aligned} \tag{54}$$

The first term is the wavefront that illuminated the hologram, scaled by the transmission of the film, the second term is the pseudoscopic image, but it is the third term that is the term of most interest to us. It is the reconstruction of the wavefront, U_2! At this point, the importance of the linearity becomes apparent. For the case shown, the amplitude of the reconstructed wavefront is scaled by β, which is a constant over the linear region of the T–E curve. Were the curve not linear, higher order terms could potentially cause the introduction of spurious reconstructions as well as nonlinear scaling of the reconstructed amplitude.

Diffraction Efficiency

Of primary importance to the holographer is the amount of optical power coupled into the reconstructed wavefront, that is, the portion of optical power in the incident beam that is channeled into the reconstructed wavefront. This is termed the *diffraction efficiency* (D.E.) and is defined as

$$\text{D.E.} \equiv \frac{\text{reconstructed image irradiance}}{\text{incident irradiance}} \tag{55}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Using the results of the previous section, the diffraction efficiency of the hologram is given by

$$\text{D.E.} = \frac{\left|\frac{AT_b\beta M}{2}\exp\{i\sin\theta_2 x\}\right|^2}{\left|A\exp\{i\sin\theta_1 x\}\right|^2} \tag{56a}$$

or

$$\text{D.E.} = \left(\frac{T_b\beta M}{2}\right)^2 \times 100\% \tag{56b}$$

If the film where perfect, such that the entire T–E curve were linear, the modulation matched exactly to the curve ($\beta M = 1$) and centered on the curve ($T_b = 1/2$), the maximum theoretical diffraction efficiency would be

$$\text{D.E.} = \left(\frac{0.5}{2}\right)^2 \times 100\% = 6.25\% \tag{57}$$

10.3 OPTICAL TECHNIQUES

10.3.1 Holographic Interferometry

Potential of Measurement Technique

In its broadest sense, holographic interferometry is a means of monitoring changes in an object by comparing optical wavefronts that have interacted with the object at different times. The type of interaction depends on the quantity to be measured. For example, if one is interested in the changes of a surface under loading, the "interaction" is as simple as a reflection of optical energy by the surface. If one where interested in how the thickness of an optically transparent material varied, the "interaction" would be the transmission of light through the object.

Holographic interferometry offers a number of advantages over other optical metrology techniques:

- It can be performed on objects difficult to measure with conventional interferometry.
- It can observe optical changes within volumes.
- Precision optics are not required (comparison is between exposures).
- Double exposure techniques are used so that accurate alignments are not required (relative motions can be removed with pulsed sources).
- Amplitude *and* phase information are stored $\Rightarrow$ time lapse or differential interferometry is possible.
- One can observe transient and steady state effects.
- Long time intervals between exposures can be used.

Copyright © Marcel Dekker, Inc. All rights reserved.

As discussed previously, holography is a technique that is used for the recording and reconstruction of wavefronts. A "snapshot" of a wavefront can be taken, analogous to a Xerox™ copy "storing" and "playing back" a simple two-dimensional image. The recorded wavefront serves as a reference against which later wavefronts can be compared. In particular, comparison will be done with wavefronts that result from placing some type of load on the object. If the wavefronts are coherent, the comparison is accomplished by interfering the two (pre- and poststressed) wavefronts together. When an object is stressed and its surface distorts, a phase difference between the recorded and the current wavefront occurs. By interfering the wavefronts, this phase difference becomes a variation in brightness, which can be easily measured. This is the essence of holographic interferometry: by comparing optical wavefronts from a pre- and poststressed object, surface distortions are mapped to changes in optical irradiance and, thus can be quantified.

As stated, the focus is upon using radiation in the optical frequency range. A very gross estimate of the sensitivity is about the order of a wavelength of light. As will be shown, the optical irradiance of the holographic interferogram can vary from completely dark to completely bright for a phase difference of half a wavelength. For a HeNe laser source, the output wavelength is 632.8 nm, leading to a value of 332 nm. Again, this is a crude estimate based upon observing only very high contrast. Typically, resolution can range from about $\lambda/2$ to $\lambda/2000$, depending upon the technique used to acquire the interferogram (Wagner, 1990).

Without needing to delve into details of the concepts and terms that were quickly introduced, one can see some of the advantages and disadvantages of using a technique that is based upon measuring the variation of an optical wavefront. The obvious advantage of the technique is that it is noncontact and noninvasive. The object is illuminated with a coherent light source, the intensity of which is well below the damage threshold of the material. Since the measurements are being made with light, remote sensing of deformations is possible. Measurements can be made in hostile environments, such as a high temperature furnace, as long as there is an optically transparent port available.

Also apparent is that, when taking measurements on the order of a fraction of a wavelength, that the surface should be isolated from vibration. When recording holograms, slight vibrations in the object can result in poor or nonexistent holograms. Like trying to take a picture of fast moving object, no one can tell what's in the photograph because of the blur caused by motion. Experiments are typically conducted on vibration isolated optical benches, where the vibrations transmitted to the table surface are greatly damped through the use of pneumatic support legs.

While this seems to be a rather stringent restriction, there are techniques that can be used to greatly lessen the sensitivity of the recording. Similar to photography, if one has a bright enough flash bulb and a fast enough shutter, a

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

fast moving object will appear to be frozen in the photograph. The same is true of holographic interferometry. By using a pulsed laser source, a bright and very short duration light pulse can be used to record a hologram, effectively freezing the wavefront in time.

Example of the Use of Wavefronts: Monitoring Ship Position on the Ocean

In order to provide insight into holographic interferometry, it is helpful to consider an analogy based upon ocean waves instead of optical waves (see Figure 10.17). Consider the wind as a source. It produces waves on the surface of the water. If the wind has been blowing at a constant speed and direction for a long time, the waves can be characterized as having their crests and troughs parallel to each other. Looking down at the surface of the ocean, the crests appear as lines, or phase fronts. The separation of the peaks (or the troughs) is the wavelength and is a function of the source characteristics (in this case, the wind velocity).

Without anything to perturb them, these wavefronts continue to propagate in their direction of travel. If there is a boat on the ocean, the presence of the ship changes the amplitude and direction of the waves that wash against its hull. Information about the boat is now carried away on those waves. If one were to look at the reflected wavefronts as a whole, information about the size and shape of the boat could be extracted from the shape of the wavefronts. If the ship were to change its position, either by moving forward or rotating, this would cause a corresponding change in the reflected waves. If a fog were to roll in and engulf the ship, one could know how the ship is moving just by monitoring these waves. In other words, one could detect and quantify the movement of the ship without having to physically measure (see) the position of the ship but by observing the changes in the reflected wavefronts.

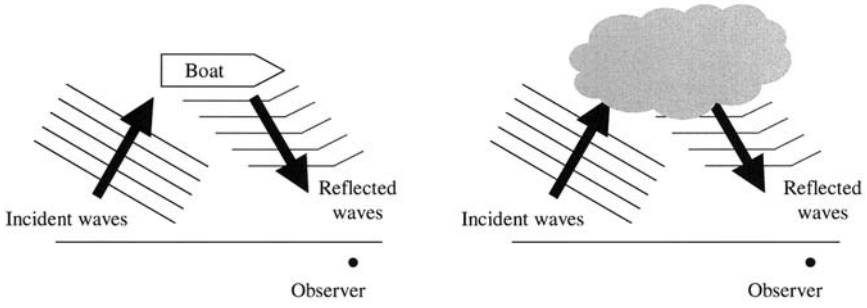


FIGURE 10.17 Observer on the shore still knows orientation and shape of boat because that information is carried on wavefronts.

Copyright © Marcel Dekker, Inc. All rights reserved.

Holographic interferometry is similar in that the changes in the optical wavefront are used to measure how an object behaves (moves) under some type of loading. In this case, the wavefronts are generated by a coherent optical source, such as a laser. When light illuminates an object, wavefronts from the source impinge upon the object like ocean waves washing on the sides of the ship. If the object is opaque, some of the incident light energy is absorbed and some is reflected. When the light wave reflects from the surface, it carries away with it information about the surface. If the surface is changed in some manner, such as by loading the object in some manner, there will be a corresponding change in the reflected optical wavefront. The change in the surface can then be quantified by comparing this wavefront to the wavefront associated with the unloaded object. Thus, how the surface deforms under the loading can be quantified without actually having to physically touch the surface. However, there is a distinct difference between extracting wavefronts information by watching the rise and fall of an ocean wave and the change in an optical wavefronts. The difference is that the waves arrive at the speed of light, 3×10^8 m/sec, making the measurement a little difficult. Observing the difference in optical wavefronts relies upon the coherent mixing, or the interference of wavefronts.

Relating Phase Difference to Surface Deformations

When comparing the optical wavefronts that have reflected from an object, the phase difference originates from a difference in the time-of-flight of the wavefronts. Figure 10.18 shows a surface that has a step in of a distance d. Initially, a planar wavefront is incident on the object, as shown in Figure 10.18a. A portion of the wavefront is reflected form the surface, while a portion of it continues along the distance d, as shown in Figure 10.18b. When the wavefront is reflected from the bottom of the step, there is a difference in the time-of-flight between the two portions of the wavefront. By interfering this wavefront with a flat wavefront, which would serve as a reference, this phase difference would cause a variation in optical intensity given by

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos(\phi_1 - \phi_2) \tag{58}$$

The difference in the times-of-flight correspond to the time it takes light to travel a distance of $2d$, twice the distance of the step. The corresponding phase difference is given by

$$\Delta\phi = \left|\vec{k}\right| 2d = \frac{4\pi}{\lambda} d \tag{59}$$

If the distance, d, where equal to a quarter of a wavelength ($d = \lambda/4$) total destructive interference would occur between the wavefronts. As shown in Figure 10.19, this variation in surface height results in a dark fringe being formed where the surface is inset.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

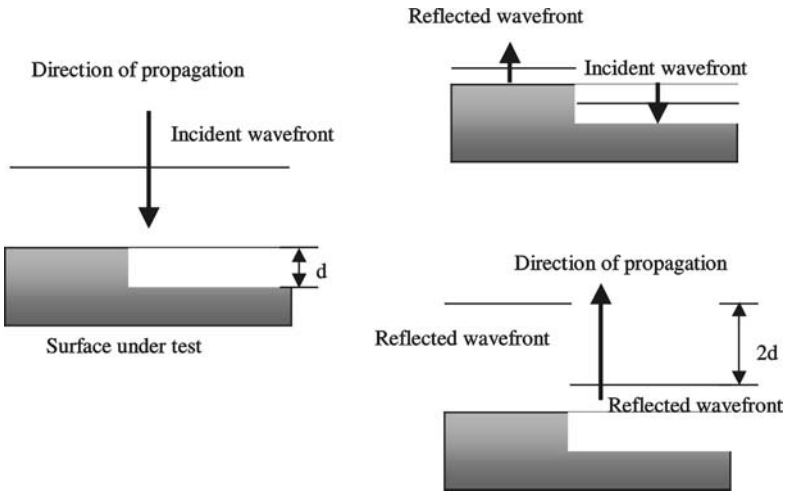


FIGURE 10.18 (a) Plane wave incident on block with step size of *d*; (b) Wavefronts reflected from block.

This technique compares the wavefront of an object to a known wavefront, in this case a planar wavefront. Differences in the surface, as compared to some fictitious flat surface, are mapped to phase differences, which are converted to differences in optical irradiance, which can be measured. This is classical interferometry, the comparison of one wavefront to another of known characteristics. This technique is commonly used in the evaluation of optical components (optical shop testing). Optical elements, such as lenses, are designed to shape wavefronts. In the case of a focusing lens, planar wavefronts are converted to spherical wavefronts that converge at the focus. It follows that how well the elements perform their function can be determined by analyzing how the optical component shapes the wavefront. Thus, one could compare the wavefront going into the lens to the wavefront following the lens to see how well it performed its task.

Some very important questions to ask at this point are "What happens when the surface has many sharp varying features many wavelength deep?" and "What happens if the surface is very rough?" Extracting data from comparing the wavefront from the surface to the planar wavefront becomes difficult, if not impossible. This is where holographic interferometry comes into play. Holographic interferometry is comparison of the object wavefronts that were taken at different times, while loading the object. The known wavefront used in classical interferometry has been replaced by a wavefront of a prestressed object. The comparison is then made between wavefronts between a pre- and a poststressed

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

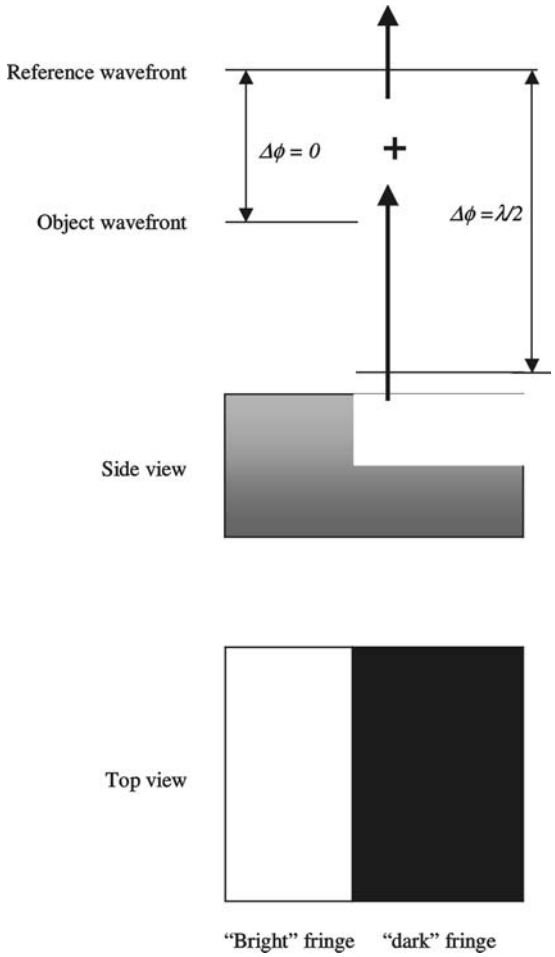


FIGURE 10.19 Interference of the reflected wavefront and a planar wavefront. Top view shows bright and dark images owing to phase difference between wavefronts.

object. By comparing these wavefronts, changes in the surface are made apparent using interferometry without having to have knowledge of the complexity of the object wavefront.

The Sensitivity Vector

One factor that is common to all interferometric measurements is the concept of a sensitivity vector. The idea is illustrated in Figure 10.20. We have an object of

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

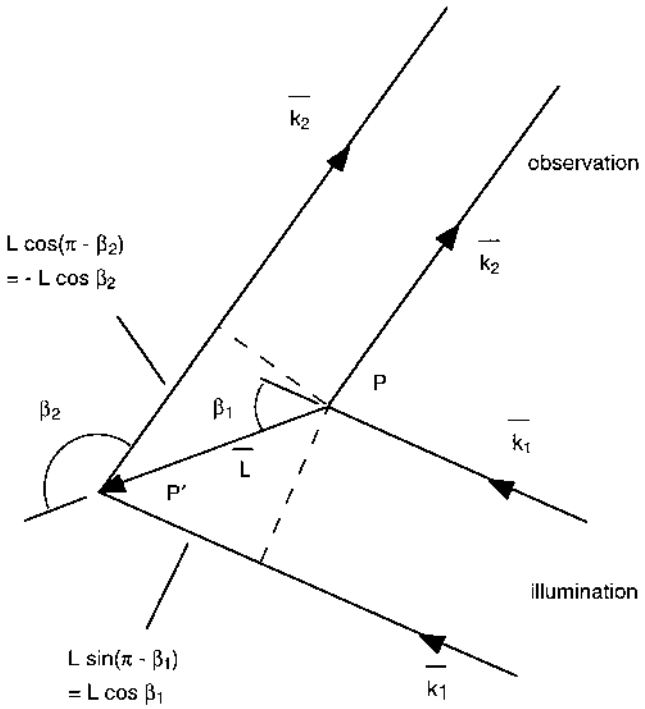


FIGURE 10.20 Trajectory of a point, P, on an object undergoing deformation.

which we may take a "snapshot" by various means prior to any movement or deformation. This serves as a reference. A second "snapshot" is taken after the experiment. In this figure we have illustrated a single point on the object before (P) and after (P') deformation. This point takes the trajectory $\overline{L}$. We wish to relate this trajectory to the measurement configuration.

From this illustration, we see that the two phase differences for the illumination and observation directions are respectively

$$\begin{aligned} \overline{k}_1 \cdot \overline{L} \\ -\overline{k}_2 \cdot \overline{L} \end{aligned} \tag{60}$$

so that the total phase difference is given by

$$\Delta\phi = \left(\overline{k}_1 - \overline{k}_2\right) \cdot \overline{L} \tag{61}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Using the definition of the scalar dot product, we can rewrite this as (see Figure 10.21)

$$\begin{aligned} \Delta\phi &= KL\cos\gamma \\ K &= |\overline{k}_1 - \overline{k}_2| \end{aligned} \tag{62}$$

This expression can be simplified somewhat through the following reasoning:

$$K = |\overline{k}_1 - \overline{k}_2| = \sqrt{k_1^2 + k_2^2 + k_1 k_2 \cos 2\Psi} \tag{63}$$

where the total angle between the illumination and observation directions is 2Ψ. Now the magnitudes of the wave vectors are identical, so that we have

$$\begin{aligned} K &= \sqrt{2k^2(1+\cos 2\Psi)} \\ &= \sqrt{4k^2\cos^2\Psi} \\ &= 2k\cos\Psi \end{aligned} \tag{64}$$

The vector $\overline{K}$ is known as the sensitivity vector because its direction with respect to the trajectory of the object point makes the measurement more or less sensitive. Further, this vector must be known to enable one to interpret the interferometric measurement.

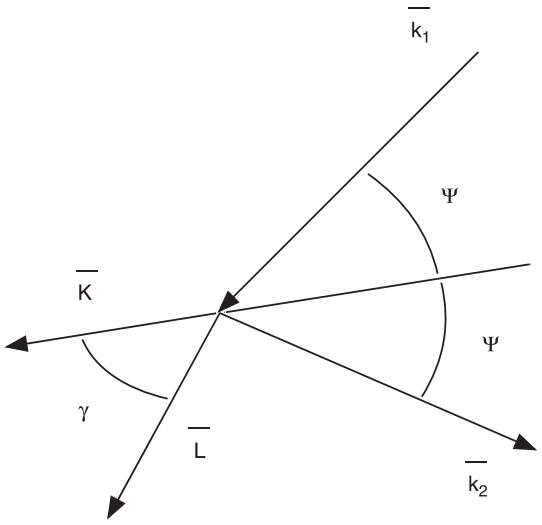


FIGURE 10.21 Illustration of the sensitivity vector.

Copyright © Marcel Dekker, Inc. All rights reserved.

The importance of the concept of a sensitivity vector can be seen when one realizes that one sees fringes for which

$$\Delta\phi = n2\pi; \quad n = 0, 1, 2, \ldots \tag{65a}$$

or

$$L\cos\gamma = \frac{n\lambda}{2\cos\Psi} \tag{65b}$$

Figure 10.22 illustrates the orientation of the sensitivity vector. See Abramson (1996) for an extensive discussion of the "holodiagram" which is a handy tool for visualizing these interference fringes.

Double Exposure Holographic Interferometry

Double exposure holographic interferometry is, as its name implies, the capturing and subsequent comparison of two wavefronts, both of which have been archived in the same hologram. Multiple independent holograms can be stored on a single film plate, as long as the exposure time of each hologram is adjusted so as not to saturate the film and drive the dynamic range to zero. Using the setup shown in Figure 10.23, a film plate is exposed to the interference of an object wavefront,

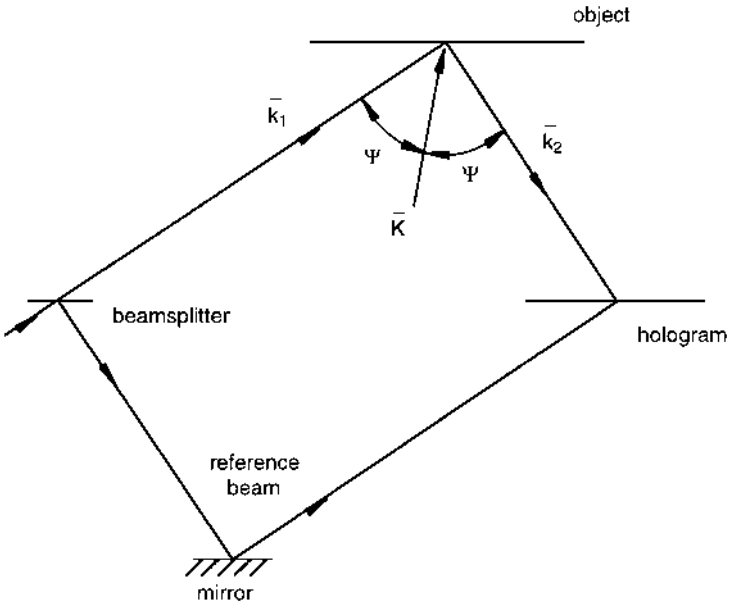


FIGURE 10.22 Relationship of the sensitivity vector to illumination and scatter directions.

Copyright © Marcel Dekker, Inc. All rights reserved.

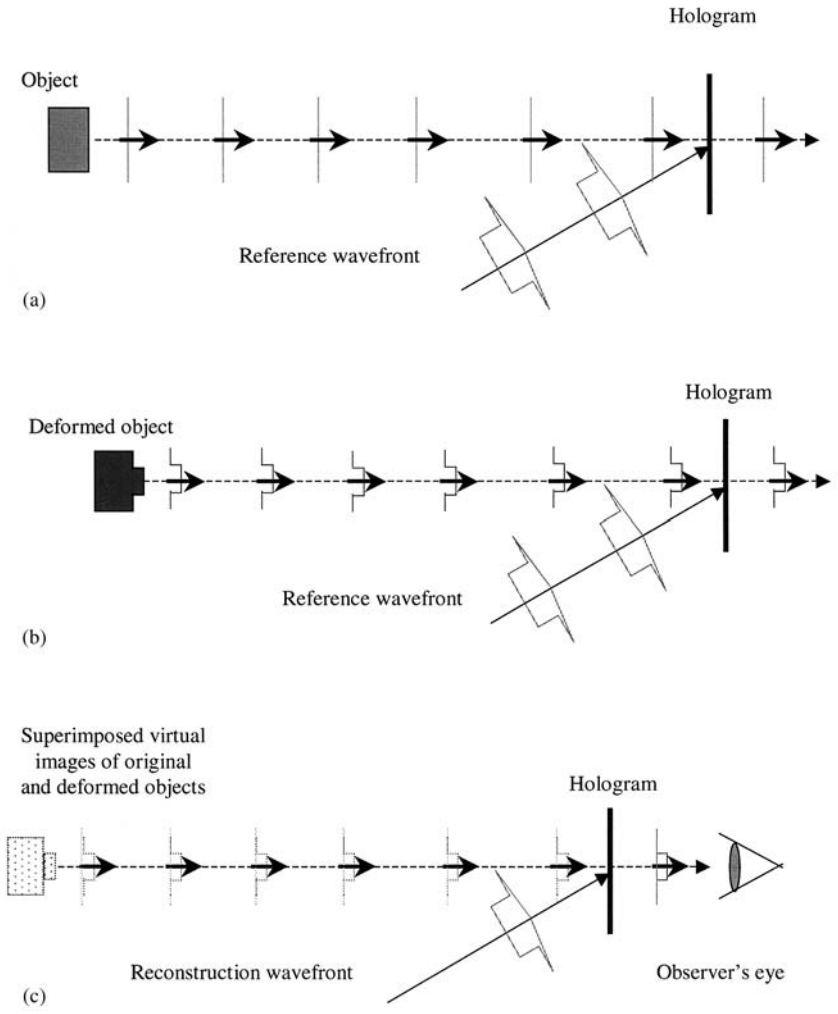


FIGURE 10.23 Double exposure holographic interferometry. (a) Recording a hologram of object; (b) recording a hologram of deformed object (film still not developed); (c) observer sees mixing of object wavefront and wavefront from deformed object, both of which are reconstructed from the same hologram.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

O_1, and a reference wavefront, R. If the hologram were to be developed at this point, it would reconstruct the object wavefront, O_1. However, instead of developing the hologram, another exposure is made. This time the object has been placed under some stress, so that the film plate records the interference of the new object wavefront, O_2, and the same reference, R. Now the film is developed and illuminated with the reference beam. What wavefronts are reconstructed? Using the short hand notation, the wavefronts leaving the hologram are

$$\begin{aligned} RT &\propto R\left|O_1+R\right|^2+R\left|O_2+R\right|^2 \\ RT &\propto R\left(\left|O_1\right|^2+\left|O_2\right|^2+2|R|^2\right) \\ &\quad +R^2(O_1^*+O_2^*)+|R|^2(O_1+O_2) \end{aligned} \tag{66}$$

Since the hologram was recorded using two separate exposures using the same reference wavefront, that reference wavefront will reconstruct both object wavefronts. Both object wavefronts will interfere and the resulting intensity distribution is given by

$$I \propto \left||R|^2(O_1+O_2)\right|^2=|R|^2(|O_1|^2+\left|O_2\right|^2+O_1O_2^*+O_1^*O_2) \tag{67}$$

If we use the definitions,

$$\begin{aligned} I_1 &= |R|^2\left|O_1\right|^2 \\ I_2 &= |R|^2\left|O_2\right|^2 \end{aligned} \tag{68}$$

then the reconstructed intensity (Eq. (102)) can be written as

$$I = I_1 + I_2 + \sqrt{I_1 I_2}(O_1O_2^* + O_1^*O_2) \tag{69}$$

Returning to phasor notation, let the original wavefront, O_1, be given by

$$O_1 = U\exp\left[i(\alpha x+\gamma z-\omega t+\phi_1)\right] \tag{70}$$

Since the object is changed so slightly under the effect of the stress, the only true difference in the waveform will be in the phase, owing to portions of the surface moving in or out;

$$O_2 = U\exp\left[i(\alpha x+\gamma z-\omega t+\phi_2)\right] \tag{71}$$

Upon substituting in these values, the intensity distribution becomes

$$I \propto |R|^2|U|^2\left|\exp[i\phi_1]+\exp[i\phi_2]\right|^2=2|R|^2|U|^2[1+\cos(\phi_1-\phi_2)] \tag{72}$$

This intensity pattern is same as that given by Eq. (93). Differences in surface heights, which occur between exposures of the hologram, are captured in the phase of the interference term. Thus, variations in the surface become variations

Copyright © Marcel Dekker, Inc. All rights reserved.

in phase, which become variations in irradiance (which can now easily be either seen with the eye or measured with an optical detector).

Holographic Contouring

Sometimes instead of knowing how an object deforms under stress, it is important to know the topography of the surface. Holography lends itself to this, making it possible to measure variations on a surface to a fraction of a wavelength of light. An example of a commonly used technique is the "dual refractive index" approach.

Consider the experimental setup shown in Figure 10.24. An object is placed inside a tank and surrounded by an atmosphere of refractive index, n_1. For the purpose of this illustrative example, the configuration shown in Figure 10.24 uses a beam splitter cube to ensure that the object wavefront is normal to both the object and the film plate.

Now, consider making an exposure with the tank filled with an atmosphere of refractive index n_1. In this case, the object wavefront that will be reconstructed, when the plate is illuminated again with the reference, can be expressed as

$$O_1 = A \exp\{i[k(L_0 + 2n_1 z_0) - \omega t]\} \tag{73}$$

where L_0 is the distance the light traveled outside of the tank. If the plate were to be left undeveloped and in the same exact position and another exposure was made, the object wavefront recorded at this time would be registered with that of the previous exposure. However, during this exposure the refractive index of the

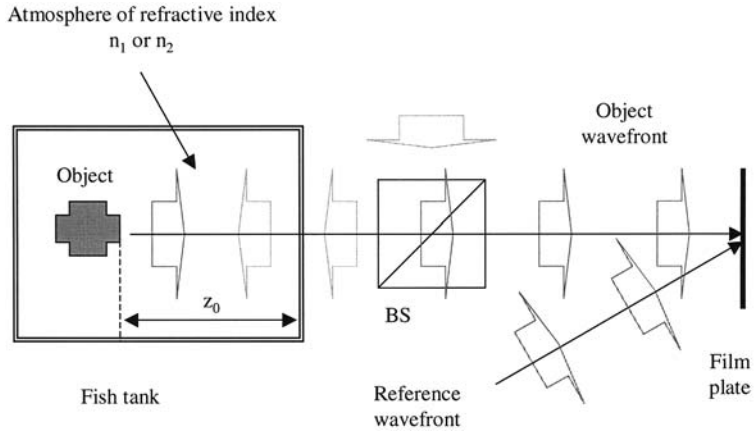


FIGURE 10.24 Experimental setup for holographic contouring using dual refractive index technique. L_0 is physical distance from object to face of tank. BS is beam splitter.

Copyright © Marcel Dekker, Inc. All rights reserved.

atmosphere instead the tank was changed (possibly by pulling a slight vacuum on the tank) to n_2. In this case, the object wavefront that would be reconstructed would be

$$O_2 = A\exp\{j[k(L_0 + 2n_2z_0) - \omega t]\} \tag{74}$$

At this point, the film plate is removed from the setup, developed, and the resulting hologram illuminated by the reference wavefront. The configuration is shown in Figure 10.25. In order to view contour fringes, the interference of the two object wavefronts, one taken with the object in refractive index n_1 and with the object in refractive index n_2, must be observed:

$$I \propto |O_1 + O_2|^2 \tag{75a}$$

or

$$I \propto 2A^2 + 2A^2\cos[2kz_0(n_1 - n_2)] \tag{75b}$$

For a bright fringe to occur, the argument of the cosine expression should be an integer multiple of 2π:

$$2kz_0(n_1 - n_2) = m2\pi \tag{76a}$$

or

$$z_0 = \frac{m\pi}{k(n_1 - n_2)} = m\frac{\lambda\pi}{2(n_1 - n_2)} \tag{76b}$$

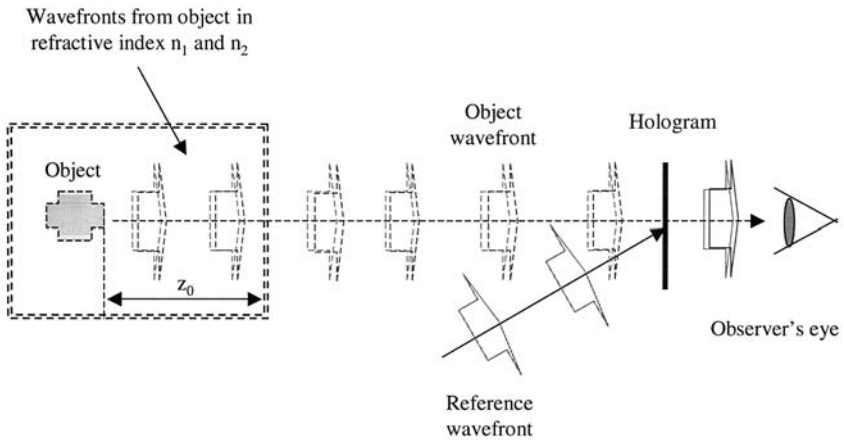


FIGURE 10.25 Configuration for ''playback'' of hologram recorded using dual refractive index technique.

Copyright © Marcel Dekker, Inc. All rights reserved.

where m is an integer. Thus, between consecutive fringes, there will be a change in the surface equal to

$$\Delta z = z_{m+1} - z_m = \frac{\lambda}{2(n_1 - n_2)} \tag{77}$$

10.3.2 Speckle Techniques

Potential of Technique

The speckle techniques that we will discuss share their conceptual foundation with the previously discussed holographic techniques in that both are interferometric methods. There are two general categories of speckle techniques:

Speckle pattern correlation interferometry (SPC interferometry)
Speckle pattern photography (SPP photography)

In each of these methods an interference pattern is derived from an optically rough surface by combining an image prior to and one after the object deformation. Depending on the specific recording and subsequent fringe observation technique, these fringes can be made sensitive to local displacements (in-plane or out-of-plane), displacement gradients, or the first derivative of these gradients. Advantages of these techniques over holographic interferometry are:

Sensitivity can be varied over a much larger range.
There are lower resolution requirements on the recording medium.
There are lower stability requirements.

The disadvantage over holographic techniques is the usually poorer fringe visibility.

As you might suspect, SPP techniques (which do not use a reference beam) are the older of the two categories. SPC interferometry is the more recent development. This approach uses a reference beam and is therefore more sensitive. We will discuss a number of specific types of SPC interferometry, principally because they can be implemented with solid state imaging devices rather than photographic film.

The Speckle Phenomenon

Before we take up the discussion of speckle techniques, we wish to discuss the speckle phenomenon itself. Consider the effect of coherently illuminating an optically rough surface (see Figure 10.26). The field at an observation point removed from this coherently illuminated surface can be considered to be due to contributions from fictitious secondary point sources on the surface, much as in Huygen's principle (Jenkins and White, 1957). As in the previous discussion, the

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

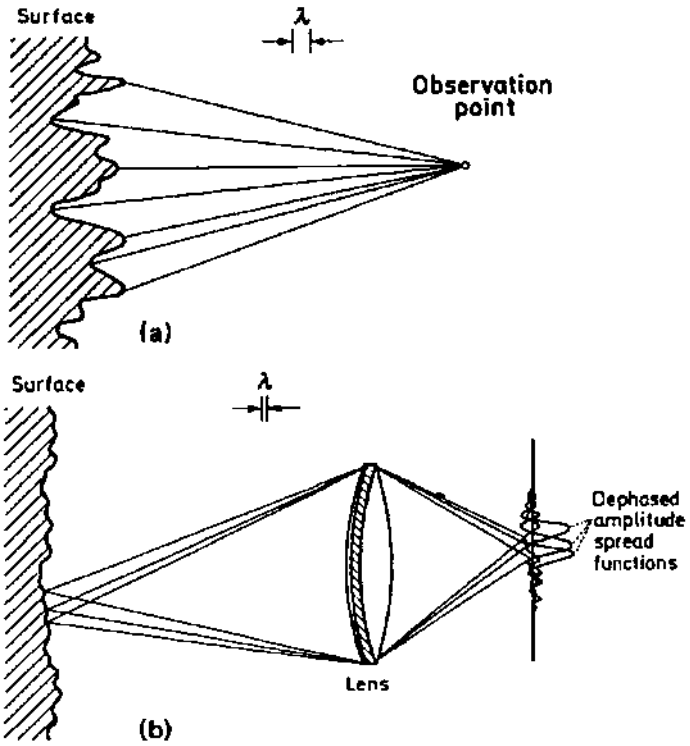


FIGURE 10.26 Origin of laser speckle. (a) objective speckle; (b) subjective speckle, from Goodman (see Dainty, 1975).

scattered field is still spatially (and temporally) coherent. The contributions for each of these secondary source points are randomly phased because of the rough surface. The effect is to produce a highly structured intensity variation in the observation plane that is referred to as laser speckle. These speckle patterns, resulting from the illumination of a surface with highly coherent light, were first observed after the invention of the laser in the early 1960s, although similar phenomena were studied as early as the nineteenth century (Dainty, 1975).

These randomly distributed bright and dark spots, known as speckles, are the result of interference between light reflected from many independent scattering areas on a rough surface. In a simple (nonimaging) free space geometry, the intensity of an objective speckle pattern in the observation plane is a result of the superposition of the complex amplitudes of the wavelets scattered from each point on the surface. For geometries where the surface is imaged onto a detector, a subjective speckle pattern is obtained which also includes the effects of diffrac-

Copyright © Marcel Dekker, Inc. All rights reserved.

tion from a finite aperture. An objective speckle pattern would be observed, for instance, if one were to replace the observation plane of Figure 10.26 with a CCD camera without a lens. If one were to place a lens on the camera, one would observe a subjective speckle pattern. The nomenclature alludes to the fact that details of the observed speckle pattern may be subject to the "interpretation" (hence "subjective") of the imaging system. In the literature, one sometimes sees these objective and subjective classifications referred to as Fresnel and Fraunhofer.

A statistical treatment of the objective speckle phenomenon treats speckle as a random walk (Goodman 1975). Under this analysis, it is found that for intensity measurements of polarized speckle the standard deviation equals the mean intensity and therefore, the contrast, defined as the ratio of the mean intensity to the standard deviation, is always equal to one. A heuristic discussion of the speckle phenomenon is contained in Appendix A.

The general SPC interferometric measurement concept is illustrated in Figure 10.27, where we have sketched a Michelson interferometer that instead of mirrors in the two legs of the interferometer, has rough surfaces. This is an optical system in which each rough surface is imaged onto the detector plane. Surface number one plays the role of a reference, while surface number two is to be

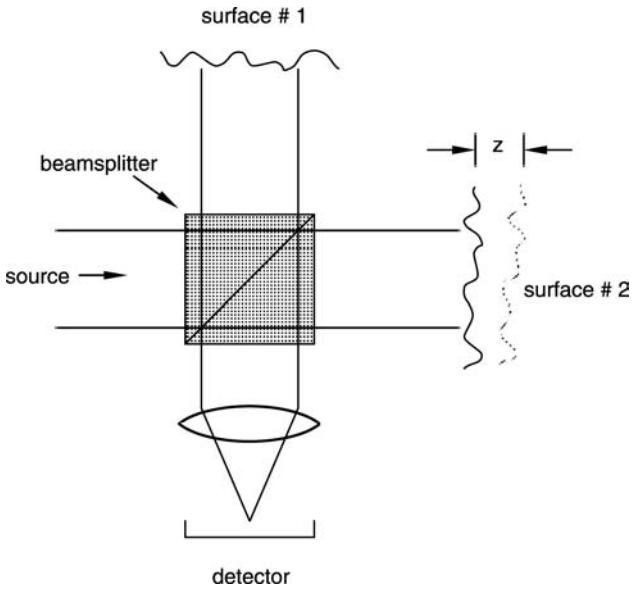


FIGURE 10.27 General configuration for performing SPC interferometry.

Copyright © Marcel Dekker, Inc. All rights reserved.

inspected. We represent the complex image fields resulting from the two rough surfaces as

$$\begin{aligned} U_1(\bar{r}) &= A_1(\bar{r})\exp\{i\phi_1(\bar{r})\} \\ U_2(\bar{r}) &= A_2(\bar{r})\exp\{i\phi_2(\bar{r})\} \end{aligned} \tag{78}$$

where $\bar{r} = (x, y)$. Because these two images are superimposed at the detector plane, the recorded intensity is

$$\begin{aligned} I_r(\bar{r}) &= |U_1(\bar{r}) + U_2(\bar{r})|^2 \\ &= A_1^2(\bar{r}) + A_2^2(\bar{r}) + 2A_1(\bar{r})A_2(\bar{r})\cos[\phi_1(\bar{r}) - \phi_2(\bar{r})] \end{aligned} \tag{79}$$

Surface number two now undergoes an out-of-plane deformation (positive when directed towards the camera) which may be a function of the transverse coordinates,

$$z(\bar{r}) = z(x, y) \tag{80}$$

after which we make a second exposure. This image is given by

$$\begin{aligned} I(\bar{r}) &= \left|A_1(\bar{r})\exp\{i\phi_1(\bar{r})\} + A_2(\bar{r})\exp\{i[\phi_2(\bar{r}) + 2kz(\bar{r})]\}\right|^2 \\ &= A_1^2(\bar{r}) + A_2^2(\bar{r}) + 2A_1(\bar{r})A_2(\bar{r})\cos[\phi_1(\bar{r}) - \phi_2(\bar{r}) - 2kz(\bar{r})] \end{aligned} \tag{81}$$

The next step is to subtract the two intensities and to square;

$$\begin{aligned} S(\bar{r}) &= [I(\bar{r}) - I_r(\bar{r})]^2 \\ &= 4A_1^2(\bar{r})A_2^2(\bar{r})\{\cos[\phi_1(\bar{r}) - \phi_2(\bar{r})] - \cos[\phi_1(\bar{r}) - \phi_2(\bar{r}) - 2kz(\bar{r})]\}^2 \end{aligned} \tag{82}$$

With a couple of simple trigonometric identities, this can be written as

$$S(\bar{r}) = 4I_1(\bar{r})I_2(\bar{r})\sin^2[\phi_1(\bar{r}) - \phi_2(\bar{r}) + kz(\bar{r})]\{1 - \cos[2kz(\bar{r})]\} \tag{83a}$$

where we have also defined the image intensities as

$$I_i(\bar{r}) \equiv A_i^2(\bar{r}) \tag{83b}$$

Now the first sinusoid on the right-hand side of Eq. (83a) is the speckle because of the phase terms due to the rough surfaces. The second term (brackets) is the fringe term that arises from the deformation. The net result of this is that one observes a fringe pattern superimposed on a speckled image of the test surface. If we inspect the series of dark fringes,

$$2kz(x, y) = 2n\pi \tag{84a}$$

we see that these are the loci of constant deformation;

$$z(x, y) = n\frac{\lambda}{2} \tag{84b}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

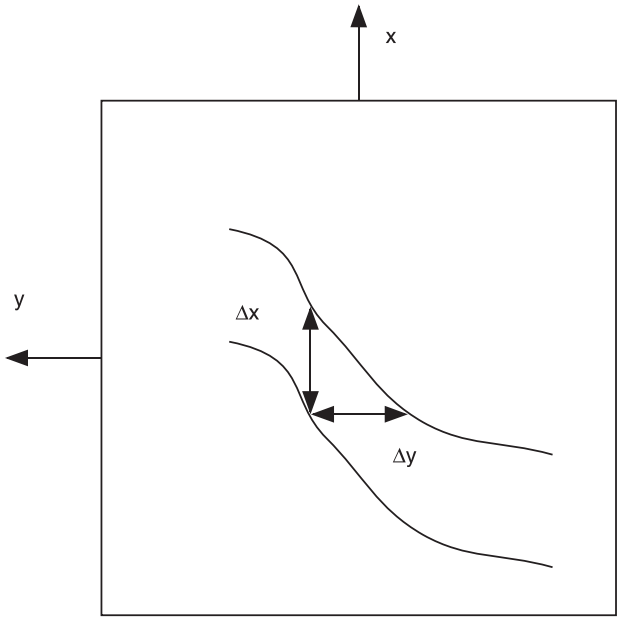


FIGURE 10.28 Conceptual illustration of SPCI fringes.

A conceptual illustration of these fringes is shown in Figure 10.28. Here, we see that the relationship between adjacent fringes is such that

$$z(x + \Delta x, y) - z(x, y) = \frac{\lambda}{2}$$
$$z(x, y + \Delta y) - z(x, y) = \frac{\lambda}{2} \tag{85}$$

By counting the number of fringes along a given path one can infer the total out-of-plane deformation along this path.

The following series of figures represent a simulation of this measurement concept. Figure 10.29a shows the speckled reference image, and Figure 10.29b the second exposure after a linear tilt in the x-direction of $\pm 4\lambda$ at the top and bottom of the image. Finally, Figure 10.29c shows the squared difference image with the fringes superimposed on the speckled image of the object. The total deformation from top to bottom is 8λ and therefore there are 16 fringes. Of course for more complex object changes these fringes are generally curved and unevenly spaced.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

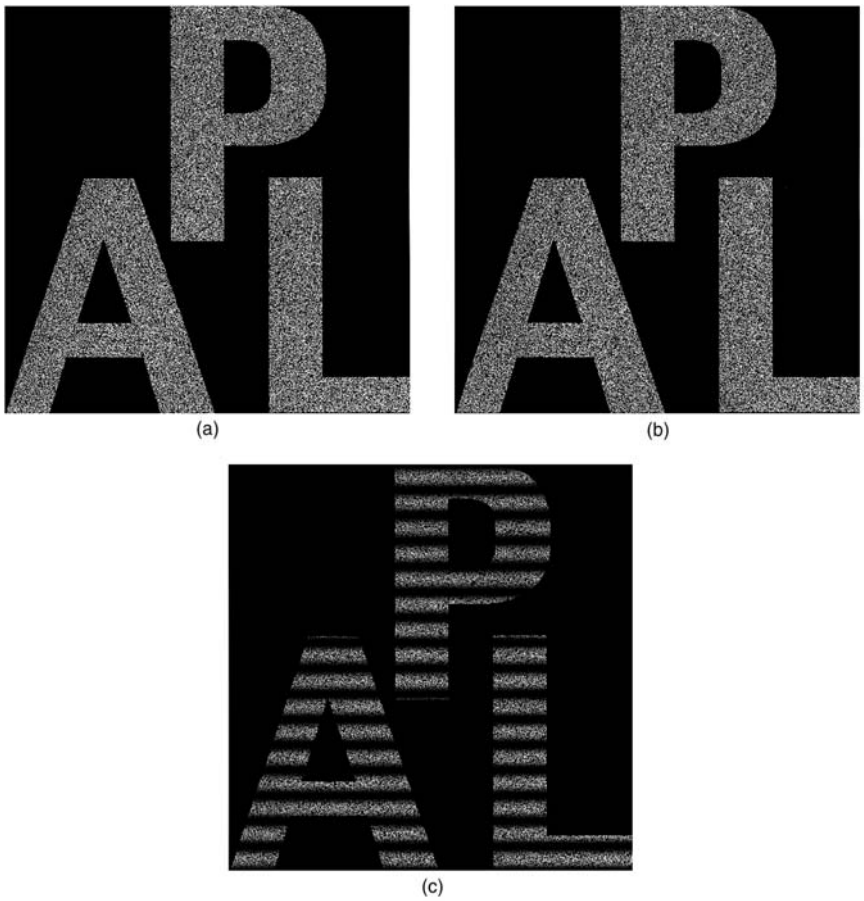


FIGURE 10.29 Simulation of SPC interferometry. (a) Original rough object under coherent illumination; (b) object after deformation; (c) squared difference image showing interference fringes.

Electronic Speckle Pattern Interferometry

A particular implementation of these speckle interference concepts is *Electronic Speckle Pattern Interferometry* (ESPI), sometimes called video holography (Figure 10.30). The figure shows a CCD camera imaged onto a coherently illuminated surface. The reference beam is introduced by means of a beamsplitter. In operation, this configuration is used to capture a video image prior to stressing the object. This reference image is subtracted from all subsequent video frames

Copyright © Marcel Dekker, Inc. All rights reserved.

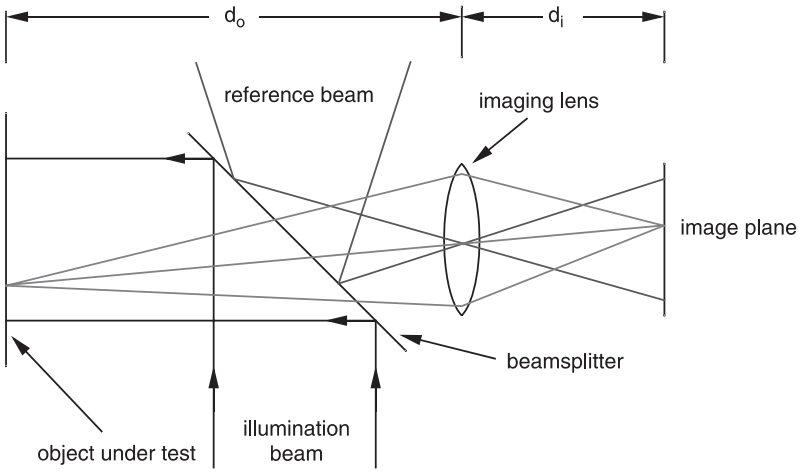


FIGURE 10.30 Typical configuration for electronic speckle pattern correlation interferometry (ESPI).

and the resultant difference is squared. Any out-of-plane strains resulting from the stress are manifested in real time as a series of fringes that are the loci of out-of-plane displacements of one-half wavelength. Systems such as this are commercially available and are often used in NDE of numerous machine components.

Shearing Interferometry

Another interesting implementation of SPC interferometry is *shearography.* This is a technique in which the object is self-referencing. In other words, the object itself provides the reference fields. The name shearography refers to the fact that the image is "sheared", that is, it looks like a double exposure with one image shifted or sheared with respect to the other. A means of producing such a sheared image (there are many others) is illustrated in Figure 10.31. The essential feature of this configuration is that the camera is focused on the test surface, but views this surface through reflections in each of the indicated mirrors. Because one mirror is tilted slightly, one of the image replicates is shifted laterally by a small amount. The mathematical formulation for this type of interferometer parallels that of the previous development. The complex image fields are

$$\begin{aligned} U_r(\bar{r}) &= A(\bar{r})\exp\{i\phi(\bar{r})\} \\ U_r(\bar{r}+\Delta\bar{r}) &= A(\bar{r}+\Delta\bar{r})\exp\{i\phi(\bar{r}+\Delta\bar{r})\} \end{aligned} \tag{86}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

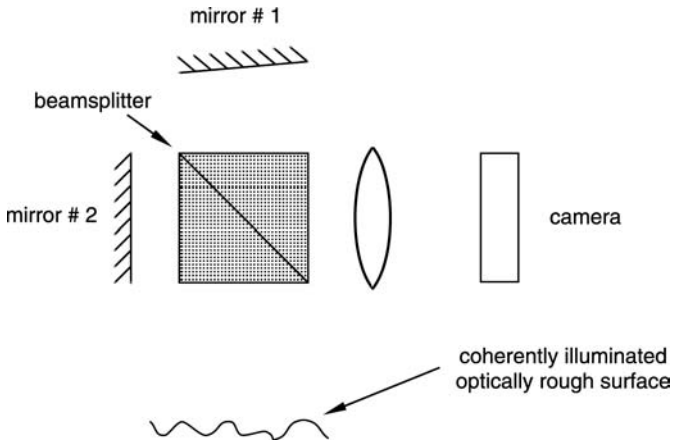


FIGURE 10.31 Illustration of means of performing speckle pattern shearing interferometry (shearography).

where $\Delta\bar{r}$ represents the amount of image shear or lateral offset. The recorded intensity at the detector plane is

$$\begin{aligned} I_r(\bar{r}) &= \left|U_r(\bar{r}) + U_r(\bar{r}+\Delta\bar{r})\right|^2 \\ &= A^2(\bar{r}) + A^2(\bar{r}+\Delta\bar{r}) + 2A(\bar{r})A(\bar{r}+\Delta\bar{r})\cos[\phi(\bar{r}) - \phi(\bar{r}+\Delta\bar{r})] \end{aligned} \quad (87)$$

Now variations in amplitude due to the image shear are small; they correspond to the macroscopic properties of the surface, whereas variations in the phase are due to surface roughness that is at the scale of the wavelength. As a result, the expression for the reference image can be approximated as

$$I_r(\bar{r}) = 2A^2(\bar{r})\left\{1 + \cos[\phi(\bar{r}) - \phi(\bar{r}+\Delta\bar{r})]\right\} \quad (88)$$

Subsequent to deformation of the test surface, the complex fields are

$$\begin{aligned} U(\bar{r}) &= A(\bar{r})\exp\left\{i[\phi(\bar{r}) + 2kz(\bar{r})]\right\} \\ U(\bar{r}+\Delta\bar{r}) &= A(\bar{r}+\Delta\bar{r})\exp\left\{i[\phi(\bar{r}+\Delta\bar{r}) + 2kz(\bar{r}+\Delta\bar{r})]\right\} \end{aligned} \quad (89)$$

The second image is therefore

$$I(\bar{r}) = 2A^2(\bar{r})\left\{1 + \cos[\phi(\bar{r}) + 2kz(\bar{r}) - \phi(\bar{r}+\Delta\bar{r}) - 2kz(\bar{r}+\Delta\bar{r})]\right\} \quad (90)$$

Copyright © Marcel Dekker, Inc. All rights reserved.

As before, we subtract these two images and square the result to obtain

$$S(\bar{r}) = 4I^2(\bar{r}) \sin^2\left[\phi_r(\bar{r}) + \frac{\Delta\phi}{2}\right] \times [1 - \cos(\Delta\phi)] \tag{91}$$

where the interference signal and speckle noise terms are given respectively by

$$\begin{aligned} \Delta\phi &= 2kz(\bar{r}) - 2kz(\bar{r} + \Delta\bar{r}) \\ \phi_r(\bar{r}) &= \phi(\bar{r}) - \phi(\bar{r} - \Delta\bar{r}) \end{aligned} \tag{92}$$

Again, the first sinusoid on the right hand side of Eq. (91) is the speckle term. The second is the interference term. If we assume that the shear is in the x-direction, then this phase difference, for small amounts of shear, is an approximation of the first derivative of the deformation;

$$\begin{aligned} \Delta\phi &\approx 2kz(x, y) - 2k\left[z(x, y) + \Delta x \frac{\partial z(x, y)}{\partial x}\right] \\ &\approx 2k\Delta x \frac{\partial z(x, y)}{\partial x} \end{aligned} \tag{93}$$

If we inspect the series of dark fringes,

$$2k\Delta x \frac{\partial z(x, y)}{\partial x} = 2n\pi \tag{94}$$

we see that these are the loci of constant deformation slope;

$$\frac{\partial z(x, y)}{\partial x} = \frac{n\lambda}{2\Delta x} \tag{95}$$

In other words, these fringes lie along curves where the slope (in the x-direction) of the deformation is constant. Likewise, had we sheared in the y-direction, we would observe fringes along

$$\frac{\partial z(x, y)}{\partial y} = \frac{m\lambda}{2\Delta y} \tag{96}$$

These results point out the beauty of this NDE concept: By varying the amount of shear, we can directly adjust the fringe spacing, i.e., the sensitivity of the measurement technique. For a commonly used shear of 3 mm and a HeNe laser, the fringe spacing would be approximately 10^{-4}.

The books by Cloud (1995), Erf (1978), Jones and Wykes (1988) Robinson and Reid (1993) and Gasvik (1995) are excellent references for ESPI, shearography and other speckle metrology techniques.

Data Processing Issues

There are a number of issues that we have glossed over in our discussion of these speckle techniques. One very important one was illustrated with the "APL"

Copyright © Marcel Dekker, Inc. All rights reserved.

simulation. Recall that the uniform fringe spacing told us that the deformation was a pure tilt, and the total number of fringes told us the amount of tilt. From these results, however, we could not discern the *sense* of the tilt, i.e., whether the top of the object was tilted towards the camera or away. Another issue is the effect of the high contrast, high spatial frequency speckle noise. It turns out that the human "eye" is very adept at discerning patterns embedded in noise. The word eye is in quotes because vision is a very complicated phenomenon that involves not just the eye but the brain as well. If one uses the camera/computer analog of the eye/brain, it is very clear what constitutes the information sensor and what constitutes the processor of this information. In the case of human vision, however, this line of demarcation is much fuzzier. For instance, there is a certain amount of noise suppression that is performed as part of the detection process itself, simply because of the way that the retina is wired to the brain. Subsequently, a sophisticated preprocessing is performed (again, based on the physical structure of the eye/brain connection) even before the information is presented to the conscious mind.

From this discussion one gathers that teaching a computer to process these fringe patterns is a nontrivial task. In fact, shearographic images described by Eq. (91) are generally only useful for qualitative analysis of deformation. Nevertheless there are some "tricks" we can use to extract quantitative phase maps in an effort to describe a deformation in greater detail. In many types of interferometry including shearography, phase stepping techniques are used to directly determine the signal phase, $\Delta\phi$. Generally speaking, a known phase change is introduced by either changing the path length along one arm of the interferometer or by varying the phase of the laser light across the object's surface. The former technique, known as *temporal* phase stepping, requires multiple images to quantify a single state of object deformation. In the latter method, referred to as *spatial* phase stepping or alternatively as the *Fourier technique*, all of the phase data are obtained by capturing one image per deformation state through the introduction of carrier fringes. We will begin with a discussion of temporal phase stepping and follow with a presentation of the spatial phase stepping technique.

A simple modification to the standard shearography experimental configuration allows for temporal phase stepping. As depicted in Figure 10.32, a piezoelectric element can be used to translate one of the mirrors to introduce a phase delay in one arm of the interferometer. (Note that the trick is to calibrate the voltage required to translate the mirror by the appropriate distance.) It turns out that by recording two images prior to and two images after the deformation, the signal phase, $\Delta\phi$, can be determined unambiguously (modulo π). If we rewrite Eq. (90) accounting for the possible phase shift, denoted by α, we get

$$I_i(\bar{r}) = 2I(\bar{r})\left\{1 + \cos\left[\phi_r(\bar{r}) + \Delta\phi + \alpha_i\right]\right\} \tag{97}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

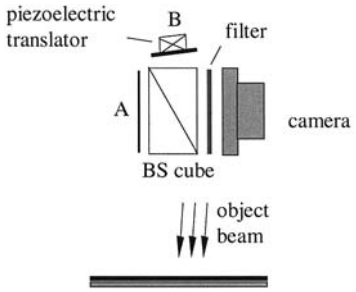


FIGURE 10.32 Shearography configuration for temporal phase stepping.

By choosing the phase steps listed in Table 10.1 for images of the deformed and undeformed object, the wrapped signal phase can be calculated from

$$\Delta\phi = 2\arctan\left(\frac{I_3 - I_2}{I_4 - I_1}\right) \tag{98}$$

We note that this operation completely eliminates the noise associated with the speckle, $\phi_r(\bar{r})$!

The results that are obtained from Eq. (98) may have discontinuities at $\pm\pi/2$. If this is the case, the result is referred to as being "wrapped". By identifying the discontinuities and adding or subtracting multiples of π, the data can be unwrapped. This subject is discussed further in a subsequent section.

The other common technique of solving the problem of directional ambiguities is to introduce *carrier fringes*. In the case of the "APL" example, these carrier fringes could be introduced by intentionally tilting the test object in between the two exposures. For shearographic measurements, the position of the illumination source can be shifted slightly in between the reference and measurement exposures to produce a set of evenly spaced fringes (as in the APL example). Any additional deformation that takes place between these two

TABLE 10.1 Required Phase Shifts for Temporal Phase Stepping Shearography

Image	α	Displacement
I_1	$-\pi/2$	0
I_2	0	0
I_3	0	$z(x, y)$
I_4	$\pi/2$	$z(x, y)$

Copyright © Marcel Dekker, Inc. All rights reserved.

exposures will manifest itself as a perturbation of these carrier fringes. A typical method for implementing this technique calls for the illuminating laser light to have a spherical wave front. In other words, the light expands as if it were emitted from a point source, either real or virtual. It is the shifting of this point source between exposures that creates the carrier fringes. The greater the shift, the higher the spatial frequency of the carrier fringes.

Once again the experimental configuration for spatial phase stepping is essentially the same as that for standard shearography. As before, the mirrors are sheared only along a direction parallel to one of the axes in the recorded image for analysis simplification. The only modification is to the origin of the illumination beam, which can be simply delivered by a single mode optical fiber, as depicted in Figure 10.33. The benefits of using optical fiber are threefold. First, a cleanly cleaved end provides a reasonable approximation of a point source when viewing in the far field. Second, the single mode characteristic of the fiber acts as a spatial filter, which provides relatively uniform illumination (with a Gaussian dependence). And third, it is a simple matter to translate the fiber to change the wave front's radius of curvature.

As usual, the recorded data is subtracted and squared, revealing the carrier fringes with frequency f_0:

$$S(\bar{r}) = 4I^2(\bar{r}) \sin^2\left[\frac{\Delta\phi(\bar{r})}{2} + \pi f_o x - \phi_r(\bar{r})\right]\left\{1 - \cos\left[\Delta\phi(\bar{r}) + 2\pi f_o x\right]\right\} \quad (99)$$

Note that the above equation contains a signal term (in brackets) modulated by a speckle noise term (sine squared term).

Next, Fourier transform techniques are used to demodulate the signal from the carrier fringes and to suppress the speckle noise. This process is accomplished in a fashion that is identical to the detection process that goes on in an AM radio. Specifically, the modulated signal (carrier plus perturbation) is multiplied by a sinusoid of the same frequency as the carrier (the mixing stage) and the result is low-pass filtered.

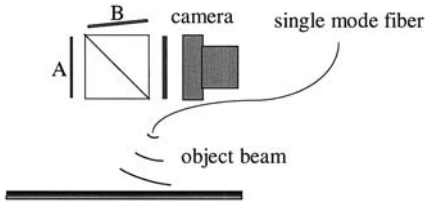


FIGURE 10.33 Shearography configuration for spatial phase stepping.

Copyright © Marcel Dekker, Inc. All rights reserved.

As a means of introducing the Fourier technique, consider the Fourier transform of a cosine dependence;

$$\begin{aligned} g(x) &= A\left[1 + \cos(2\pi f_0 x)\right] \\ G(f) &= \int_{-\infty}^{\infty} g(x) e^{-i2\pi f x}\, dx \\ &= A\delta(f) + \frac{A}{2}\delta(f + f_0) + \frac{A}{2}\delta(f - f_0) \end{aligned} \tag{100}$$

The result is three spectral components: one at zero frequency and a pair at the frequencies of $\pm f_0$. The latter are commonly called the *sidebands*. Now, instead of this strict cosine dependence, consider the case in which the signal contains a slowly varying phase component such as

$$g(x) = A\left\{1 + \cos\left[2\pi f_0 x + \phi(x)\right]\right\} \tag{101}$$

In this case the information contained in the phase, $\phi(x)$, yields a spectral component at zero frequency plus a smearing of the frequency components centered about the carrier components as illustrated in Figure 10.34.

For a (two-dimensional) interference pattern as illustrated in Figure 10.35, a two-dimensional Fourier transform produces the results shown in Figure 10.36. In this illustration, "DC" or zero frequency is at the center of the figure. The effect of the carrier fringes is seen clearly here; they give rise to the "halos" centered about the sidebands on either side of DC. A straightforward demodulation process consists of selecting a single sideband and moving it so that the halo is centered about DC. Finally, an inverse Fourier transform yields the demodulated fringe pattern, say,

$$g(x) = A \exp\left\{i\phi(x)\right\} \tag{102}$$

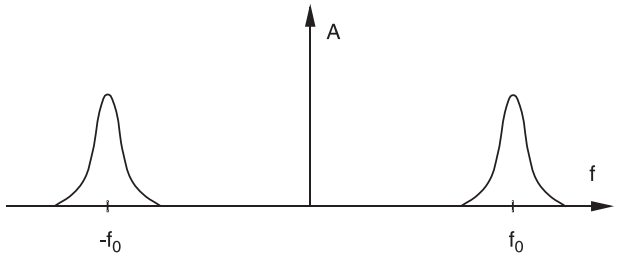


FIGURE 10.34 One-dimensional Fourier transform of function with varying phase (see Eq. (101)).

Copyright © Marcel Dekker, Inc. All rights reserved.

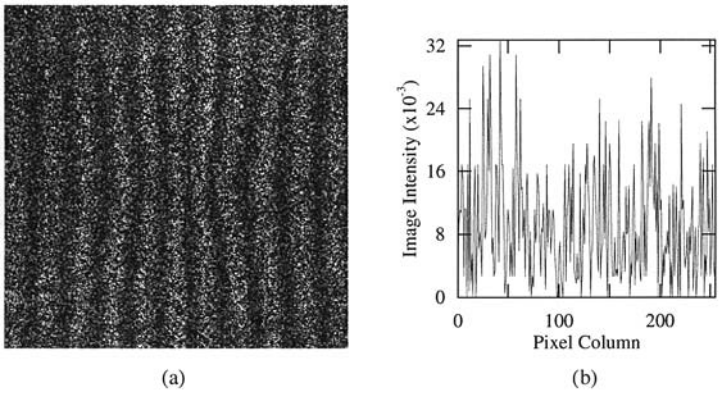


FIGURE 10.35 Shearographic fringe pattern with carrier fringes.

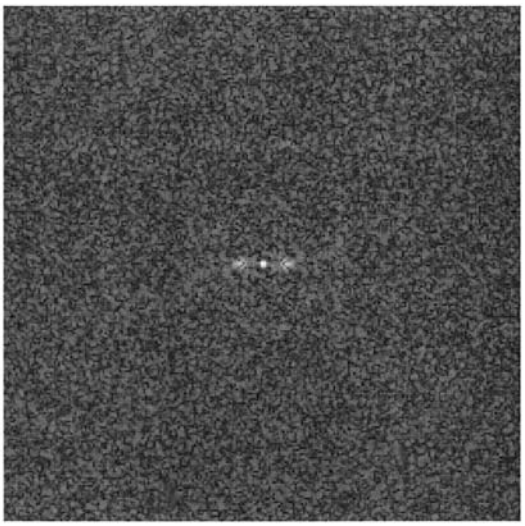


FIGURE 10.36 Two-dimensional Fourier transform of fringe pattern shown in Figure 10.35.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

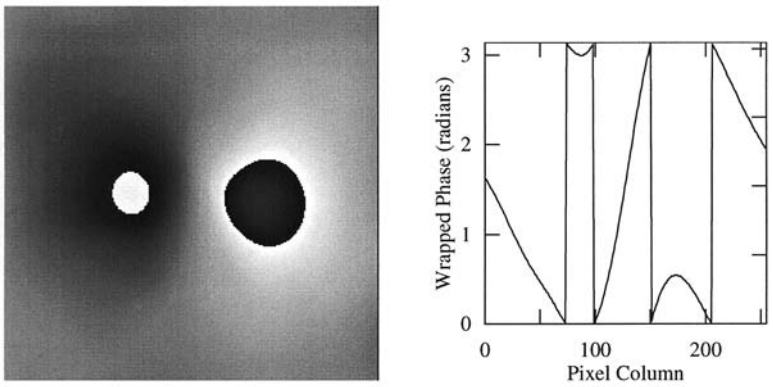


FIGURE 10.37 Demodulated fringe pattern displaying phase wrapping.

Note that information we seek is in the phase of the signal, and that the resultant function is complex. We can retrieve this information by calculating

$$\phi(x) = \tan^{-1}\left[\frac{\mathrm{Im}\{g(x)\}}{\mathrm{Re}\{g(x)\}}\right] \tag{103}$$

Now for the bad news: because the domain of the arctangent function is $\pm\pi/2$ we obtain an estimate of the phase only to within a factor of π. The phase is said to be *wrapped*. The wrapped phase map that corresponds to our example is shown in Figure 10.37. We discuss the requisite *unwrapping* step next.

Phase Unwrapping

There are a variety of phase unwrapping algorithms for detecting and compensating the π phase jumps that Eq. (103) produces. The simplest technique involves examining the difference between the phases of adjacent pixels, denoted as $i-1$ and i, and is described schematically by the flow chart in Figure 10.38. With a discrete image, the phase differences between adjacent pixels, $\Delta\phi_i - \Delta\phi_{i-1}$ will only approach π. It is necessary to define a phase jump as a discontinuity that comes within a few percent of π. The sign of the difference determines whether π is added or subtracted to pixel i; π is added if the difference is negative and it is subtracted if the difference is positive. In practice, one must keep track of the reference phase, ϕ_{ref}, an integer multiple of π, that is added to each pixel to account for all previously encountered phase jumps. For noisy images, more complex unwrapping algorithms (Ghilia and Pritt, 1998) can be used in which the phase of the pixel in question is compared with other neighboring pixels. Finally,

Copyright © Marcel Dekker, Inc. All rights reserved.

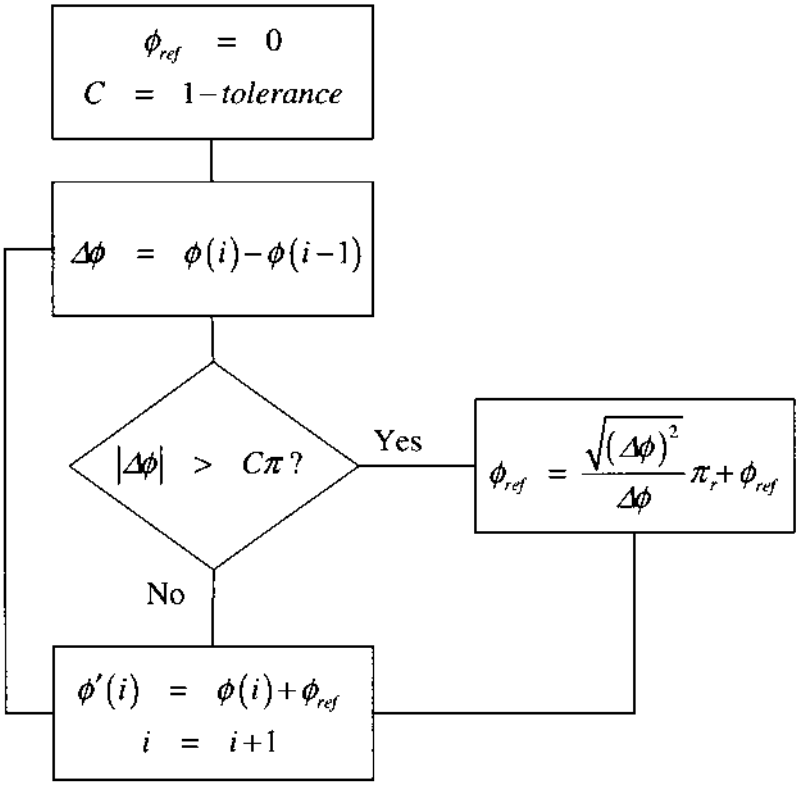


FIGURE 10.38 Flowchart of a simple unwrapping algorithm.

the measured values for the directional derivatives are obtained by scaling the unwrapped phase such that

$$\frac{\partial z}{\partial x} = \frac{\lambda}{4\pi}\frac{\Delta\phi}{\Delta x} \tag{104}$$

The result of applying this unwrapping algorithm to the wrapped phase map (Figure 10.37) is shown in Figure 10.39.

10.3.3 Structured Light

There are many measurement techniques that use "structured" light. For example one may project a bright line onto an object and observe it obliquely. Any height variations in the object manifest themselves as lateral displacements of the bright

Copyright © Marcel Dekker, Inc. All rights reserved.

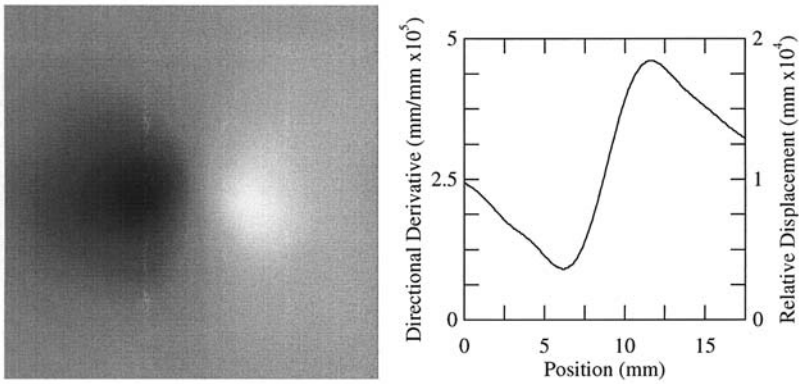


FIGURE 10.39 Unwrapped version of phase map shown in Figure 10.37.

line (see Figure 10.40). With knowledge of the projection and viewing angles the surface height variation can be calculated.

Structured light can take the form of something as simple as a straight line, or a whole family of lines. One type of such structured light uses the interference between a reference mask and a mask whose phase has been modified:

$$\begin{aligned} t_1(x,y) &= a\left[1+\cos\left(\frac{2\pi}{p}x\right)\right] \\ t_2(x,y) &= a\left\{1+\cos\left[\frac{2\pi}{p}x+\Delta\phi(x,y)\right]\right\} \end{aligned} \tag{105}$$

This modification ($\Delta\phi$) is the displacement of the light pattern caused by the surface irregularity. Suppose that we lay one of these gratings atop the other. The total transmittance is simply the product;

$$\begin{aligned} t(x,y) &= t_1(x,y)t_2(x,y) \\ &= a^2\left[1+\cos\left(\frac{2\pi}{p}x\right)\right]\left\{1+\cos\left[\frac{2\pi}{p}x+\Delta\phi(x,y)\right]\right\} \end{aligned} \tag{106}$$

Multiplying out the two grating terms and using a simple trig identity allow us to write this in the form

$$\begin{aligned} t(x,y) = a^2\Bigg\{&1+\cos\left(\frac{2\pi}{p}x\right)+\cos\left[\frac{2\pi}{p}x+\Delta\phi(x,y)\right] \\ &+\frac{1}{2}\cos\left[\frac{4\pi}{p}x+\Delta\phi(x,y)\right]+\frac{1}{2}\cos[\Delta\phi(x,y)]\Bigg\} \end{aligned} \tag{107}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

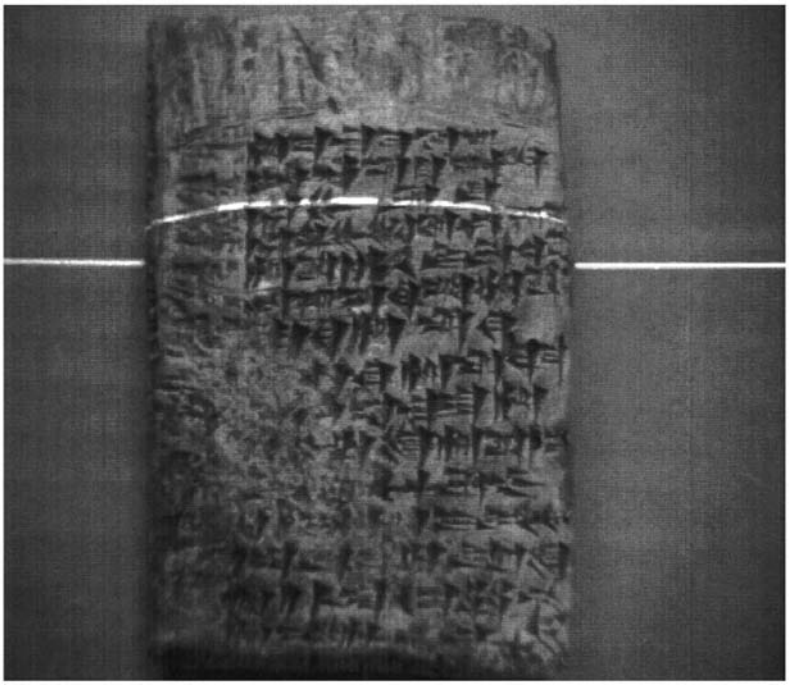


FIGURE 10.40 Illustration of structured light concept.

Of the five individual terms on the right hand side of this expression, the last one is the one that is perceived as the lowest spatial frequency "interference" or moiré term. This is the term containing information about the object surface. Moiré is the French for "watered" or "wavy appearance" that one observes when layers of silk are laid atop one another at an angle. The effect is often seen when looking through a screen at a brick wall, or through two rows of fence. It turns out that the moiré effect is useful for visualizing a wide range of optical effects.

Consider a grating with a square-wave amplitude transmission, illustrated conceptually in Figure 10.41. Such a grating, called a Ronchi (Ron′ke) ruling, can be drawn using the Postscript program listed in Appendix B. An example is shown in Figure 10.42. If a pair of these gratings is laid in contact with each other at an angle, one observes the low frequency moiré pattern as shown in Figure 10.43. An explanation of this moiré pattern can be derived by considering the problem of two plane waves intersecting at $\pm\theta$. From our discussion of interference, we recall that the resulting fringe or "beat" frequency is

$$f_{\text{beat}} = \frac{2\sin\theta}{\lambda} \tag{108}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

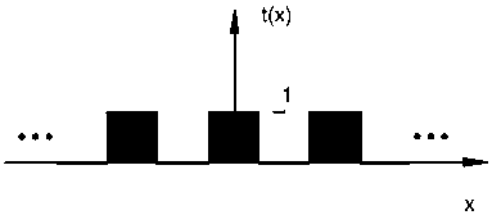


FIGURE 10.41 Transmission profile of a Ronchi grating.

Figure 10.42 is a photographically reduced version of the grating generated using the Postscript program listed in the appendix. The period of the full-size grating is 1 cycle/mm. Figure 10.43 shows two such gratings laid atop one another at a total angular rotation of 20° (again at reduced scale). In this analog, the wavelength is 1 mm and thus from the above equation, the beat frequency should be 0.35 cycle/mm, which can be easily verified from the figure (actually from the figure generated by the program in the appendix). From this example, we see that the moiré analog is useful for visualizing interference phenomena.

To see how this effect might be used to measure surface height variations, we will discuss the shadow moiré approach. Consider Figure 10.44 that illustrates a grating lain atop a surface with some height variation. The point P_1 is on the surface of the object. Illumination is provided from direction θ_1 and the object is viewed from direction θ_2. The shadow of the point indicated by P_0 is cast upon the point P_1, but the ray scattered towards the viewer will be blocked or passed by the grating at point P_2. Another way of thinking about this is that the point P_0 appears to have been shifted to the point P_2.

$$\begin{aligned} d &= d_1 + d_2 = z(x, y)(\tan\theta_1 + \tan\theta_2) \\ \Delta\phi &= \frac{2\pi d}{p} = \frac{2\pi z(x, y)}{p}(\tan\theta_1 + \tan\theta_2) \end{aligned} \tag{109}$$

where p is the "pitch" of the grating, i.e., the period.

This explanation makes use of the concept of the method of images. Consider the reflection of a point source viewed in a mirror, as illustrated in Figure 10.45. The object is perceived as being located at the point *below* the mirror surface at the point indicated.

Analysis of the resulting data can be done by any of the procedures discussed previously.

10.3.4 Photoelastic Techniques

All of the previously discussed techniques were applied to the visualization of surfaces and surface deformations. We now take up a discussion of a technique

Copyright © Marcel Dekker, Inc. All rights reserved.



FIGURE 10.42 Ronchi grating.

that is useful for measuring full-field stress distributions within the volume of a specimen, photoelastic analysis. This is typically done by making a model of the specimen with a material of known photoelastic properties. Some of the more common such materials are Homolite 100, polycarbonate (Lexan), epoxy, or urethane rubber. Each of these materials is naturally optically isotropic, but

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016



FIGURE 10.43 Moiré produced from crossed Ronchi gratings.

exhibits a stress-induced change in the refractive index making it artificially birefringent. Artificially birefringent materials such as these exhibit an induced birefringence that is directly proportional to the applied stress. Therefore, measurement of the degree of induced birefringence and knowledge of the

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

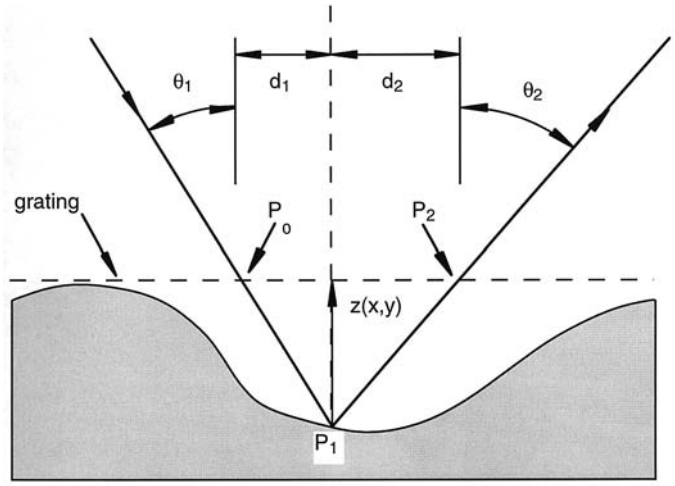


FIGURE 10.44 Production of moiré fringes using a contact grating.

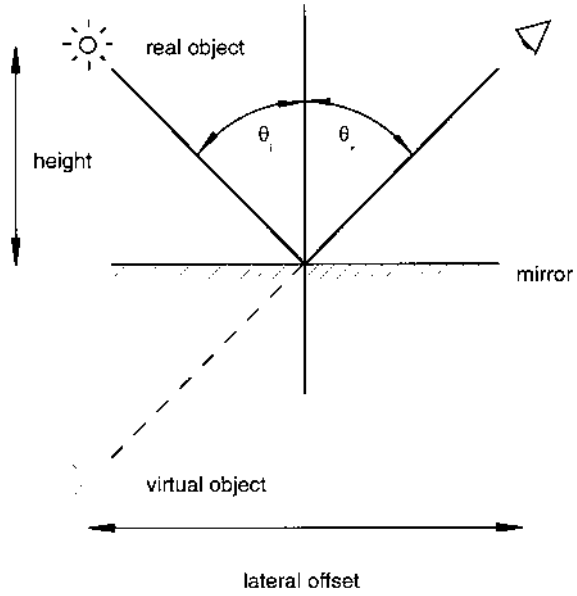


FIGURE 10.45 Method of images.

Copyright © Marcel Dekker, Inc. All rights reserved.

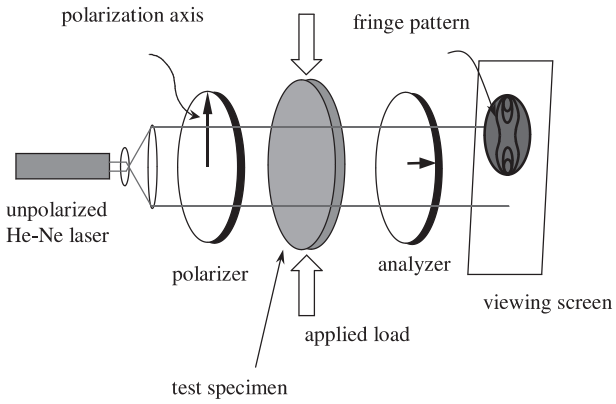


FIGURE 10.46 Illustration of a plane polariscope.

stress-optic constants, which relate the change in refractive index to the applied stress, allows the stresses in the model to be determined.

A polariscope, either a plane or circular configuration, is one means of measuring the induced birefringence in a specimen. Shown in Figure 10.46 is a plane polariscope, which consists of a laser source, a beam expander, a linear polarizer, the specimen being tested, an analyzer (which is another linear polarizer with its optical axis perpendicular to the first linear polarizer), and a viewing screen. A digital camera can be used to record the fringe pattern on the viewing screen for postprocessing of the real-time changes of the specimen's stress distributions. In this configuration, a diametral load is being applied to the specimen with the resulting fringe patterns being monitored.

To understand how the plane polariscope works, we return to our description of the basic wave nature of light. The light exiting the polarizer can be represented by

$$U_p(z, t) = U_0 \cos(kz - \omega t) \tag{110}$$

Because the initial phase of the light is not important for this analysis, Eq. (110) can be reduced to

$$U_p = U_0 \cos \omega t \tag{111}$$

The test object, by nature of its birefringent properties, acts as a waveplate, and resolves the incident light vector into two components, U_1 and U_2. These two

Copyright © Marcel Dekker, Inc. All rights reserved.

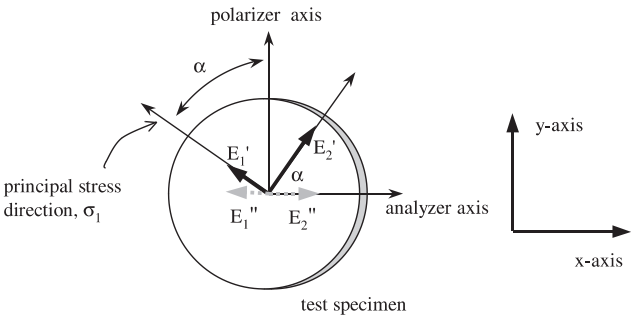


FIGURE 10.47 Wave components for plane polariscope.

components have vibrations parallel to the principal stress directions at that point. Thus, the two field components can be represented by

$$\begin{aligned} U_1 &= U_0 \cos\alpha \cos\omega t \\ U_2 &= U_0 \sin\alpha \cos\omega t \end{aligned} \tag{112}$$

where α defines the angle between the largest principal stress direction, σ_1, and the axis of the polarizer (see Figure 10.47).

Since the induced birefringence causes the two field components to propagate through the test object with different velocities, they can be expressed as

$$\begin{aligned} U_1' &= U_0 \cos\alpha \cos\left\{\left[\frac{2\pi h}{\lambda}\,(n_1 - 1)\right] - \omega t\right\} = U_0 \cos\alpha \cos(\Delta_1 - \omega t) \\ U_2' &= U_0 \sin\alpha \cos\left\{\left[\frac{2\pi h}{\lambda}\,(n_2 - 1)\right] - \omega t\right\} = U_0 \sin\alpha \cos(\Delta_2 - \omega t) \end{aligned} \tag{113}$$

where n_1 and n_2 are the refractive indices seen by the two field components and λ is the wavelength of the laser. The two horizontal components are then resolved when the field vectors enter the analyzer. The resulting component, after transmission through the analyzer, is given by

$$\begin{aligned} U_r &= U_2'' - U_1'' \\ &= U_2'' \cos\alpha - U_1'' \sin\alpha \\ &= U_0 \sin 2\alpha \sin\left(\frac{\Delta_2 - \Delta_1}{2}\right) \sin\left(\frac{\Delta_2 + \Delta_1}{2} - \omega t\right) \end{aligned} \tag{114}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Since the intensity of light is proportional to the square of the amplitude of the light wave, the intensity of the light exiting the analyzer is given by

$$I = |U|^2 \sin^2 2\alpha \sin^2\left(\frac{\Delta_2 - \Delta_1}{2}\right)$$
$$= |U|^2 \sin^2 2\alpha \sin^2\left(\frac{\pi h}{\lambda}\ \Delta n\right) \tag{115}$$

This result shows that extinction occurs whenever $\sin^2 2\alpha = 0$ or $\sin^2(\pi h \Delta n/\lambda) = 0$. The first term is related to the principal stress directions, and results in what is known as an isoclinic fringe pattern. An isoclinic fringe occurs at $2\alpha = m\pi$, where $m = 1, 2, \ldots$. These fringes represent loci of points where the principal stress directions (either σ_1 or σ_2) coincide with the axis of the polarizer.

The second term, $\sin^2(\pi h \Delta n/\lambda)$, is related to changes in the refractive index, Δn. As explained earlier, even materials that are naturally optically isotropic can exhibit a stress-induced birefringence, or change in refractive index. These resultant fringes, known as isochromatic fringes, occur when

$$\pi h/\lambda\ \Delta n = m\pi \tag{116}$$

To obtain information on the principal stress differences in a material, one must first relate the change in the refractive indices to the applied stresses. It can be shown that for isotropic materials the change in refractive is linearly related to the stress difference, $\sigma_1 - \sigma_2$, by the relative stress-optic coefficient, c, as

$$\Delta n = \Delta n_2 - \Delta n_2 = c(\sigma_1 - \sigma_2)$$
$$= c\Delta\sigma \tag{117}$$

Combining Eqs. (116) and (117) allows the principal stress difference, $\Delta\sigma$, at each fringe location to be calculated as

$$\Delta\sigma = \frac{h}{mc\lambda} \tag{118}$$

Shown in Figure 10.48 is a plane polariscope monitoring the static stress field in a disk under diametral compression. The load is being applied using an MTS machine and the resulting fringe pattern is shown in Figure 10.49. In this figure, the first three fringe orders are clearly marked as well as the partial outline of the disk.

10.3.5 Equipment

To make any of the techniques outlined in this chapter a reality, a certain equipment base is required. Given here is just a brief listing of likely equipment (and why you would need it) as well as possible vendors.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

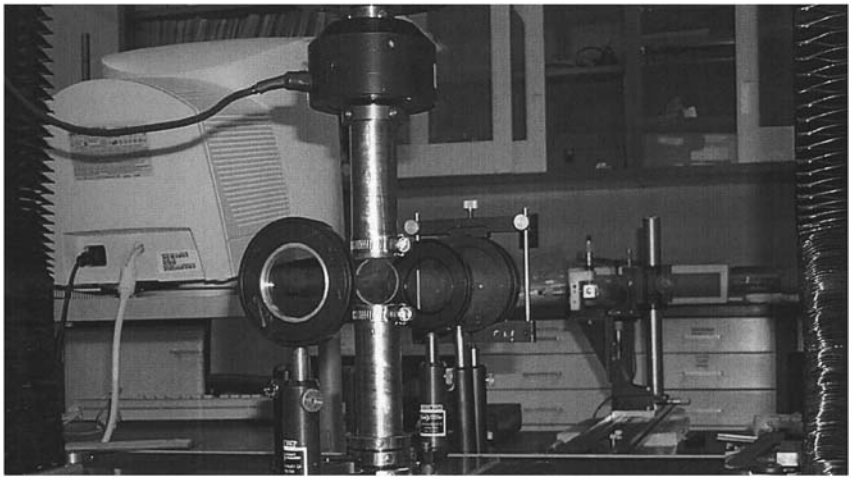


FIGURE 10.48 Plane polariscope being used to monitor stress fields in a thin disc.

Vibration Isolation

The measurements that are being made are on the order of a fraction of a wavelength of light. This requires that some optical measurement experiments be conducted in areas that are free from ambient vibrations. Source of such vibrations range from HVAC units to people walking by. Typically, vibration isolation is accomplished by using damped pneumatic legs to support the table

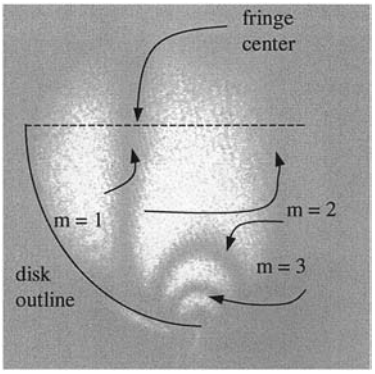


FIGURE 10.49 Measured fringe pattern for thin disc in diametral compression.

Copyright © Marcel Dekker, Inc. All rights reserved.

top, as well as honey-combed internal structure and a solid metal table top. Since a pneumatic table is a rather large investment, simple experiments in holography (such as demonstrating making a hologram) may be conducted by using small inner tubes for vibration isolations. Optical components, of the experiment, can be set in sand to further deaden the effects of vibrations.

Possible Vendors

Newport Corporation
1791 Deere Avenue
Irvine, CA 92606
In U.S. (800) 222-6440
Tel: (949) 863-3144
Fax: (949) 253-1680
E-mail: info@newport.com
http://www.newport.com/Vibration_Control/

Melles Griot
Photonics Components
16542 Millikan Avenue
Irvine, California 92606
Tel: (800) 835-2626/(949) 261-5600
Fax: (949) 261-7589
E-mail: sales@irvine.mellesgriot.com
http://www.mellesgriot.com/

The Laser Source

The selection of a laser source depends upon the application. In some cases, pulsed laser sources are required for the application. A pulsed laser is analogous to using a camera flash bulb, upon exposing an object to a short duration of light, that moment will be "frozen" on film. Pulse lengths can range from femtoseconds (Ti:Sapphire lasers) to 10–100 s of milliseconds (frequency-doubled Nd:YAG.) These sources have short coherence lengths, requiring that the reference and object arm path lengths be closely matched.

Continuum
3150 Central Expressway
Santa Clara, CA 95051
Tel: (800) 956-7757
E-mail: continuum@ceoi.com
http://www.continuumlasers.com/

Big Sky Laser: Quantel
P.O. Box 8100
601 Haggerty Lane
Bozeman, MT 59715
Tel: (800) 224 4759
Fax: (406) 586 2924
Email: info@bigskylaser.com
http://www.bigskylaser.com

Continuous (or CW) laser sources are also available. An example of one is the popular red Helium Neon (HeNe) laser. Such lasers are convenient when object motion and/or vibration isolation are not an issues. These types of laser sources

Copyright © Marcel Dekker, Inc. All rights reserved.

can have very long coherence lengths, on the order of kilometers. Naturally, these sources will require a means to control the exposure time, such as the use of a shutter.

Coherent, Inc. 5100 Patrick Henry Drive Santa Clara, CA 95054 Tel: (408) 764-4983 Fax: (408) 988-6838 E-mail: tech_sales@cohr.com http://www.coherentinc.com/	Spectra-Physics Headquarters 1335 Terra Bella Avenue Post Office Box 7013 Mountain View, CA 94039-7013 Tel: (800)-775-5273 E-mail: sales@splasers.com http://www.spectra-physics.com/

Optics, Mounts, and Bases

A number of "supporting" components will generally be needed:

- Front surface mirrors for beam steering of CW laser sources. For pulsed laser sources, dielectric mirrors are typically used. These are optical flats that have been coated with a wavelength dependent, highly reflective, high damage threshold material.
- Lenses for beam expansion/compression (forming an afocal telescope) as well as imaging.
- Shutters to control the exposure time of an experiment by effectively turning a beam "off" and "on."
- Spatial filters are a valuable tool to conducting any experiments requiring illumination of an object with a laser beam. The output of a laser may not be a spatially "smooth" intensity that one would think it should be, due to the mode structure of the laser. In addition, spatial noise can be introduced into the beam by any optics that the beam passes through. A spatial filter, which is typically composed of a microscope objective and a pinhole, is used to ensure that that laser beam does provide "smooth" illumination. The laser output is focused onto a pinhole that is matched to the numerical aperture of the microscope objective. The result is that the "noise" in the beam falls outside the pinhole and, thus, it is effectively removed (filtered) from the beam.
- Optics mounts which provide a means to support and position the optical components. These mounts typically provide a means to achieve fine adjustment of the angular positioning of the element.
- Beam splitters for breaking a single laser source into separate (object and reference) beams. Various beam splitter configurations exist, providing for either a fixed or adjustable ratio of the output beam power.

Copyright © Marcel Dekker, Inc. All rights reserved.

Posts and post holders are for elevating the optical mounts and making sure that the optics are collinear. Posts are for controlling the elevation of the optical mounts and post holders are for securing these to the vibration-isolated table.

Some vendors, in addition to Newport and Melles Griot, are:

Thorlabs, Inc. 435 Route 206 N. Newton, NJ 07860 Tel: (973) 579-7227 Fax: (973) 383-8406 E-mail: Info@thorlabs.com http://www.thorlabs.com/	CVI LASER CORPORATION 200 Dorado Place SE Albuquerque, NM 87123 Tel: (800) 296-9541 Fax: (505) 298-9908 E-mail: cvi@cvilaser.com http://www.cvilaser.com

Detectors

Detectors provide a means of measuring the irradiance at the film plane. Knowing the irradiance allows one to determine the proper film exposure time. Newport is an example of a vendor providing calibrated detector systems for such applications.

Film (holographic plates)

(For recording holograms of course!)

Integraf 745 N. Waukegan Rd. Lake Forest, IL, 60045 Tel: (847) 234-3756 Fax: (847) 615-0835 E-mail: integraf@aol.com http://members.aol.com/~integraf/	VinTeq, Ltd. 611 November Lane/Autumn Woods Willow Springs, NC 27592-7738 Tel: (877) 639-9424 Fax: (919) 639-7523 E-mail: vinson@vinteq.com http://www.vinteq.com/

CCD Cameras

CCD cameras are an integral part of the ESPI system. They come in a variety of grades ranging from low cost video cameras (7–8 bits of dynamic range) to scientific-grade, cooled systems (14–16 bits of dynamic range).

Copyright © Marcel Dekker, Inc. All rights reserved.

PULNiX America, Inc. 1330 Orleans Drive Sunnyvale, CA 94089 Tel: (800) 445-5444 Fax: (408) 747-0880 E-mail: pulnix@pulnix.com http://www.pulnix.com/	Roper Scientific 3660 Quakerbridge Road Trenton, NJ 08619 Tel: (609) 587-9797 Fax: (609) 587-1970 E-mail: info@roperscientific.com http://www.roperscientific.com

Software

The proper software is of considerable importance in conducting experiments and processing the resultant data. Considerable development has gone into software packages designed to help the experimentalist. These packages enable the user to control an experiment (through serial/parallel/GPIB/TTL signal interfaces with instruments) or to conduct complex analysis of data with very little time devoted to developing specific applications.

National Instruments Corporation 11500 N Mopac Expwy Austin, TX 78759-3504 Tel: (512) 794-0100 Fax: (512) 683-8411 E-mail: info@ni.com	The Math Works, Inc. 3 Apple Hill Drive Natick, MA 01760-2098 Tel: (508) 647-7000 Fax: (508) 647-7001 E-mail: info@mathworks.com http://www.mathworks.com

10.4 SUMMARY

In this chapter we have attempted to give a flavor for the power and versatility of optics in nondestructive evaluation. In particular a brief view of some popular whole-field optical NDE techniques has been given, that is, techniques in which an image of the object under test is formed. Due to space limitations there were a number of interesting techniques left out, such as Schlieren techniques (Malacara, 1978) and *Particle Imaging Velocimetry* (PIV) (Lading, *et. al*, 1994), etc. Furthermore, implicit in this discussion was the fact that we were addressing active techniques in which the object being tested or evaluated was illuminated in some fashion. Thus, we neglected what may be called passive techniques in which thermal emissions from the object under test are measured (the subject of a separate chapter). We also neglected the entire field of interferometric optical component testing which has a very rich history (Malacara, 1978). Finally, we did

Copyright © Marcel Dekker, Inc. All rights reserved.

not treat any of the myriad point measurement techniques. Examples of such nonimaging techniques include laser radars, lidars, objective (nonimaging) speckle schemes, fiberoptic based techniques, etc. Finally, it is important to note that many of the techniques discussed in this chapter require only moderate skills and a minimum of equipment. As an example, moiré measurements can be conducted using an overhead projector and a transparency made by using a simple Postscript program.

PROBLEMS

1. Show that the time harmonic spherical wave,

$$u(R, t) = \frac{U_0}{R}\exp\{ikR - \omega t\}$$

   is an exact solution to the wave equation. Hint: Make use of the symmetry of this problem; express the Laplacian in spherical coordinates.
2. In the text, we derived the expression for the interference fringes caused by plane waves striking an interface from the same side,

$$I(x) = 2I_0[1 + \cos(2kx\sin\theta)]$$

   Derive the corresponding result when the fields are incident from opposite sides. Note that in deriving the above equation it was not necessary to choose the $z = 0$ plane; the same fringe pattern exists wherever the fields intersect.
3. Provide the missing steps between Eqs. (21) and (23).
4. In the expression for the output of the Michelson interferometer, account for the factor of 2 in the argument of the cosine function.
5. Derive the expression for the output of the Michelson interferometer if the reflection–transmission ratio of the beamsplitter is m. For the case treated in the text, we assumed $m = 0.5$.
6. Starting with the equation for the coherence length in terms of the speed of light and the mean optical frequency,

$$l_c = \frac{c}{\bar{\nu}}$$

   derive the corresponding formula

$$l_c = \frac{\lambda^2}{\Delta\lambda}$$

7. What would you expect the output of a Michelson interferometer to be if the path difference exceeded the coherence length of the source?
8. Depth of field refers to the amount by which the object distance can change and still be in focus. Likewise, depth of focus is the distance the focal plane

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

position can move and still be effectively in focus. Starting with the imaging condition,

$$\frac{1}{d_o}+\frac{1}{d_i}=\frac{1}{f}$$

derive the relationship between depth of focus and depth of field,

$$\frac{\Delta d_i}{\Delta d_o}=m^2$$

where m is the ratio of image and object distances.

9. Based on what you know about speckle, how might an astronomer discriminate between light arriving at her telescope from a distant planet versus light arriving from a star? Assume that the telescope is ground based and that the phase relationship amongst parallel photons is jumbled by atmospheric turbulence.
10. Demonstrate that for the values listed in Table 10.1, the phase stepping technique eliminates the term associated with laser speckle.
11. For laser contouring, it is not always possible to control the refractive index surrounding an object. Show how the following procedure leads to an alternate technique for contouring by determining the transmission function/wavefront that corresponds to each step.

    Make a holographic exposure of the object using a source of wavelength, λ_1.
    Keeping the film in place, develop the hologram.
    Reconstruct the hologram using a source of wavelength λ_2, which propagates along a different angle with respect to the optical axis.
    View the interference of the object reconstructed wavefronts in real time.

    The result of this technique will be two distinct fringe patterns. Identify these patterns in the interference expression.

    Identify the "undesired" fringe pattern. How can it be removed?
    Once it has been removed, determine what one "contour" fringe equates to in terms of a change in the object surface.

GLOSSARY

Aperture stop: The limiting physical aperture within an imaging system whose image constitutes the entrance and exit pupils of the system.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Birefringence: A property of certain dielectrics; different refractive indices along different crystallographic axes (can be a function of stresses in the material).

Boltzman's constant: A fundamental constant of nature that relates temperature and kinetic energy.

CCD (charge-coupled device): A one or two-dimensional detector array that conceptually may be likened to a series of buckets that collect photoelectrons.

Chief ray: The ray that crosses the optical axis at the aperture stop.

Coherence, longitudinal: The degree to which light can interfere constructively and destructively with a time-delayed version of itself; proportional to the spectral purity of the light. Also known as temporal coherence.

Coherence, spatial: The degree to which light can interfere constructively and destructively with a laterally shifted version of itself; proportional to how much the source appears as a geometrical point. Also known as transverse coherence.

Coherence, temporal: See *coherence, longitudinal*.

Coherence, transverse: See *coherence, spatial*.

Color: Subjective term associated with wavelength.

Degeneracy parameter: The number of indistinguishable photons, i.e., the number within a coherence volume.

Depth of field: The distance an object can move, towards and away from the camera, and still remain in focus.

Depth of focus: The distance the focal plane may move while still maintaining the object in focus; the "image" of the depth of field.

Dielectric: Materials that possess no carriers of free charge.

Diffraction: An interference phenomenon; the apparent bending of light rays in response to a discontinuity in the propagation medium.

Diffraction grating: Generally a periodic structure which deflects light of a given wavelength in a specific direction.

Direction cosines: Projections of the unit vector (along the direction of propagation) onto the Cartesian axes.

Doppler shift: Shift in frequency caused by relative motion between the source and receiver (and possibly the medium through which the disturbance propagates).

Entrance pupil: The "conjugate" or image of the exit pupil of an imaging system.

ESPI: Electronic speckle pattern interferometry.

Euler expansion: Expression of the trigonometric sine and cosine functions in terms of complex exponentials.

Exit pupil: The "terminal" of an optical system, i.e., the apparent source of the fields that converge to the image plane.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

FFT: "Fast Fourier Transform," an efficient algorithm for the implementation of the finite discrete Fourier transform.

f-number: For an imaging system, the ratio of the focal length to the exit pupil diameter.

Fraunhofer zone: The far-field of a diffractive aperture or source; beyond the Rayleigh range.

Frequency, spatial: Number of cycles of a periodic function per unit distance, commonly expressed in cycles per mm or line-pairs per mm.

Fresnel reflection: The amplitudes of the fields on either side of a dielectric interface are related through the Fresnel reflection and transmission coefficients.

Fresnel zone: The near field of a source or diffracting aperture; less than the Rayleigh range.

FWHM (full-width at half-maximum): A measure of the width of a structure taken at the point where the value is one half its maximum.

Holodiagram: A graphical device useful for visualizing interference phenomena.

Holography: The storage and retrieval of amplitude and phase information in a square-law medium.

Hue: Subjective term associated with wavelength.

Huygens' principle: The view that a wavefront can be thought of as a series of secondary point sources; at a later time, i.e., at a point removed in space, the wavefront is then the envelope of this family of fictitious sources.

Intensity: In the field of physical optics this is generally the square magnitude of the electric field variable.

Interferogram: Temporal (as from a Michelson interferometer) or spatial interference pattern.

Interferometer, Michelson: Of the class referred to as amplitude splitting; a device for splitting a wavefront in two and then recombining so that two portions interfere constructively and destructively.

Interferometer, shearing: A device for splitting and recombining wavefronts with a lateral offset.

Interferometer, Young's: Of the class referred to as wavefront splitting; a device for causing two points on a wavefront to interefere constructively and destructively.

Interferometry: Technique involving the superposition of two or more electromagnetic fields.

Kelvins: A degree on the absolute temperature or Kelvin scale.

Laser: An acronym for "light amplification through the stimulated emission of radiation."

Method of images: A technique used to calculate electric fields in the vicinity of a conductor that involves the conceptual reflection of the source about the boundary and subsequent removal of the boundary.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Moiré: An interference pattern observed when two periodic patterns are superimposed.
Morphology: Referring to form or structure.
NDE: See nondestructive evaluation.
Newton's fringes: Interference pattern generally observed due to reflection from dielectric interfaces.
Nondestructive evaluation: A class of techniques for assessing certain properties of test objects noninvasively.
Orthoscopic image: The image with the correct orientation and proportions; the virtual image is said to be orthoscopic.
Paraxial approximation: The assumption that the observation distance from the propagation axis is small with respect to the source distance.
PIV: particle image velocimetry: a technique typically used to measure flow by seeding the fluid with small particles; a double exposure with known interexposure time allows estimate of flow velocity by measuring distance between particle images.
Phase unwrapping: Refers to any technique used to remove phase jumps of integer multiples of π from interferometric data.
Photoelastic technique: An NDE technique that relies on strain-induced birefringence to assess stress distributions in objects.
Photon: A particle of light.
Planck's constant: A fundamental constant of nature that relates a photon's wavelength and energy.
Point-spread function: The degree to which an imaging system "spreads" or blurs the image of a geometrical point object; a measure of a system's ability to resolve two closely-spaced objects.
Polariscope: A device for determining the degree of birefringence of an article.
Polarization, P: The electric field is parallel to the plane defined by the direction of propagation and the interface normal.
Polarization, S: The electric field is perpendicular ("Senkhrect") to the plane defined by the direction of propagation and the interface normal.
Polarizer: An optical device that passes only a single polarization.
Polaroid: Registered trade name of a polarizer that consists of a plastic that is stretched to align the long chain molecules and then dyed so that the electric field that is aligned with the long axis of the molecules is absorbed.
Pseudoscopic image: The real image which is "conjugate" to the virtual; its appearance is "inside-out."
PSF: See *point-spread function*.
Rayleigh criterion (for resolution): Referring to objects separated in angle by λ/D, where D is the linear dimension of the aperture.
Rayleigh range: Rough demarcation between near and far field, D^2/λ, where D is the linear dimension of the radiating aperture.
Refraction: Bending of the light ray at a dielectric interface.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Refractive index: Quotient of the velocity of light in a vacuum and the velocity in the medium.
Resolution: The ability to distinguish two closely-spaced objects.
Retarder, phase: See *waveplate*.
Ronchi ruling: A periodic grating with square wave transmission function.
Saturation: Subjective term associated with spectral purity.
Schlieren technique: Interferometric technique often used for volumetric visualization of small refractive index variations.
Sensitivity vector: Vector difference between the illumination and observation directions.
Shearography: An interferometric technique that relies on interfering two coherent images that are laterally shifted with respect to each other.
Shot noise: Noise due to the discrete nature of light.
Snell's law: Relates the propagation directions on either side of a dielectric interface.
Senkhrect: From the German for perpendicular; see *polarization S.*
Speckle: An interference phenomenon caused by scatter from adjacent points on an object that is rough compared to the illumination wavelength.
Square-law: Refers to a physical detector that is sensitive to the square of the electric field.
Structured light: A surface measurement strategy in which a light pattern is projected onto the object to be measured.
TE: transverse electric: The electric field oriented perpendicular to the plane of incidence defined by the direction of the incident ray and the interface normal; see *polarization S.*
Telecentric: imaging system: A system in which the chief ray strikes the entrance and exit pupils parallel to the optical axis.
TM: transverse magnetic: The electric field lies within the plane of incidence defined by the direction of the incident ray and the interface normal; see *polarization P.*
Virtual image: A apparent image; literally "so in effect, but not in reality."
Wavefront: A surface of constant phase.
Waveplate: A birefringent optical device that introduces a prescribed phase delay between two polarization components.

SYMBOLS

Arabic letters

c	Speed of light in vacuum, stress-optic coefficient
d_o	Object distance
d_i	Image distance

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

f	Focal length of a lens
$f^{\#}$	"f" number of an imaging system ($f^{\#} \equiv f/D$)
h	Planck's constant
I	Optical intensity
$\overline{k}$, or, $\overline{\kappa}$	Wave vector $\overline{k} = (\kappa_x, \kappa_y, \kappa_z)$
k	Wavenumber $\left(k = \lvert\overline{k}\rvert = 2\pi/\lambda\right)$, or Boltzman's constant
n	Refractive index
U	Generic E-field variable
v	Velocity
(x, y, z)	Cartesian coordinates

Greek letters

(α, β, γ)	Direction cosines
ϕ	Phase
θ	Angle, or phase
λ	Wavelength
ν	Temporal frequency
ω	Radian frequency ($\omega = 2\pi\nu$)

Mathematical operations

Re{. . .}	Real part of
∇^2	Laplacian

Abbreviations

CCD	Charge-coupled device (camera)
ESPI	Electronic speckle pattern interferometry
PSF	Point-spread function
NDE	Nondestructive evaluation
SPC	Speckle pattern correlation
SPP	Speckle pattern photography

APPENDIX A: COHERENCE CONCEPTS

Now that we have discussed interference, we wish to take up the topic of coherence. The word "coherence" is nothing more than jargon from the field of optics than means "correlation." If fields are correlated they are said to be coherent. This property of coherence is one that is usually associated with the output of a laser. Yet there are aspects of this concept of coherence that are rather fuzzy in the minds of many workers in the field. We shall attempt to shed some light (sorry) on this confusion.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

If electromagnetic fields are "coherent," they are able to interfere constructively and destructively with each other. Specifically, the property refers to correlation between the fields at two different points in space or time. Consider the light from some source. If we select two points in a plane that is transverse to the direction of propagation, then we speak of spatial coherence (sometimes called transverse coherence for obvious reasons). If we select two points in space that are separated along the direction of propagation, then we speak of temporal coherence (spatial separation along the direction of propagation is equivalent to separation in time). This is sometimes called longitudinal coherence.

To introduce the concept of temporal coherence, we first need to introduce the concept of a "quasimonochromatic" field. This is expressed in complex notation as follows:

$$u(R, t) = U(R, t) \exp\{-i2\pi\bar{\nu}t\} \tag{A1}$$

where $\bar{\nu}$ is the center frequency of the field and $U(R, t)$ is a phasor that is slowly varying in time. Specifically, this phasor has a bandwidth, $\Delta\nu$, such that

$$\frac{\Delta\nu}{\bar{\nu}} \ll 1 \tag{A2}$$

This slowly varying phasor is usually referred to as the *envelope* function. The temporal coherence of this field at two points in time is expressed in terms of the self-coherence function,

$$\Gamma(\tau) = \langle u(R, t)u^*(R, t + \tau)\rangle \tag{A3}$$

The indicated time average is over a time span that is long compared to the period $1/\bar{\nu}$. Using our definition of a quasimonochromatic field, the expression for the self coherence function becomes

$$\Gamma(\tau) = \langle U(R, t)U^*(R, t + \tau)\rangle \exp\{i2\pi\bar{\nu}\tau\} \tag{A4}$$

Now we make the observation that if the time difference, τ, is short compared to the temporal variations in the envelope function, the terms in angular brackets will be well correlated. On the other hand, if the time separation is large, we would expect the envelope functions to be uncorrelated. Since variations in the envelope function are no faster than $\sim 1/\bar{\nu}$, one might conclude that this bandwidth would serve a reasonable estimate of the *coherence time* of these fields;

$$\tau_c = \frac{1}{\Delta\nu} \tag{A5}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

The coherence length is simply the distance the field propagates within the coherence time:

$$l_c = \frac{c}{\Delta\nu} \tag{A6}$$

Another way of expressing this is in terms of the optical bandwidth of the disturbance,

$$l_c = \frac{\lambda^2}{\Delta\lambda} \tag{A7}$$

It may not be obvious at this point, but this equation tells us that temporal coherence is not an intrinsic property of the field. Rather, it suggests that if a source does not display an appreciable temporal coherence, it is possible to introduce temporal coherence by selecting only a narrow band of wavelengths by using the appropriate filter. For example, with a filter bandwidth of 2 nm and the central wavelength of 0.55 μm, the coherence length is approximately 0.15 mm.

Spatial coherence is proportional to the degree to which a source appears like a point source. Mathematically, it is expressed in terms of the mutual intensity:

$$J(R_1, R_2) = \langle u(R_1, t)u^*(R_2, t)\rangle \tag{A8}$$

To ground this concept on the physics of the problem, consider the sketch in Figure A.1. Here we have shown two independently radiating sources, S_1 and S_2 and two field points, P_1 and P_2. Also illustrated are the wavetrains emitted by the sources that arrive at the field points. We write for the fields at the observation points:

$$\begin{aligned} u(P_1) &= A_1 + B_1 \\ u(P_2) &= A_2 + B_2 \end{aligned} \tag{A9}$$

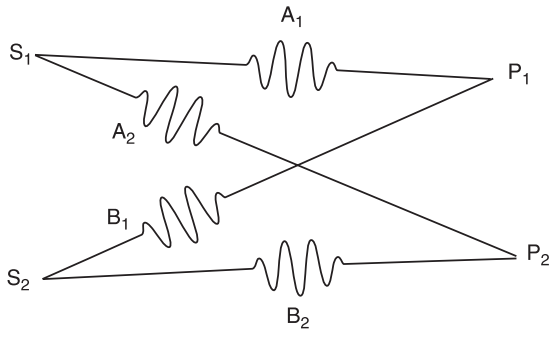


FIGURE A.1 Illustration of two uncorrelated sources.

Copyright © Marcel Dekker, Inc. All rights reserved.

If the path differences are small so that

$$\overline{S_1P_1} \approx \overline{S_1P_2} \tag{A10}$$

then

$$A_1 \approx A_2 \tag{A11}$$

Similarly, for

$$\overline{S_2P_1} \approx \overline{S_2P_2} \tag{A12}$$

then

$$B_1 \approx B_2 \tag{A13}$$

Obviously, this argument leads to,

$$u(P_1) \approx u(P_2) \tag{A14}$$

i.e., the fields at the observation plane are correlated. One might logically deduce that the region of the observation plane over which the fields are correlated simply depends upon the various path differences as the (pairs of) points explore the source and the observation planes.

A simple way of quantifying the effects of source size on the degree of spatial coherence uses the Young's double slit experiment. Consider Figure A.2

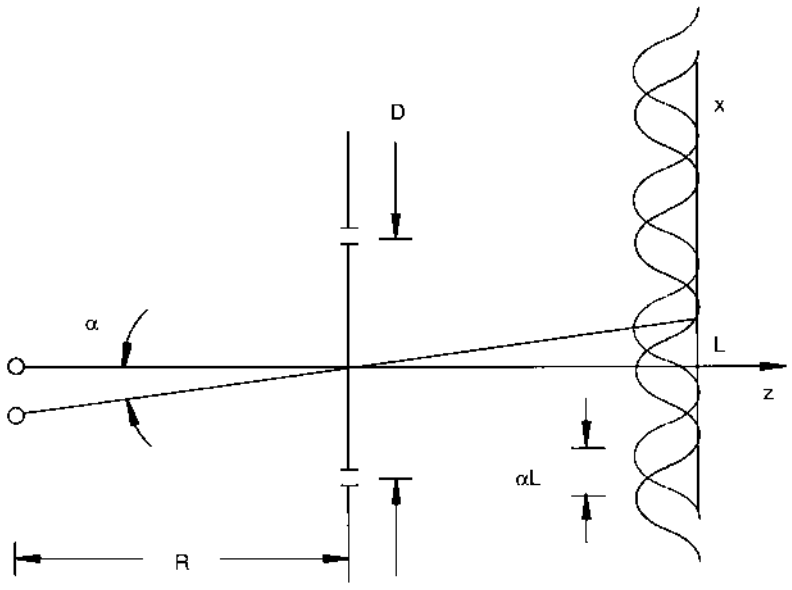


FIGURE A.2 Young's experiment for two uncorrelated point sources.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

that illustrates an opaque screen with two small apertures. Illumination is provided by two uncorrelated sources. As such, the intensity in the observation plane associated with each source adds, rather than the fields. Nevertheless, because these sources are small, each will produce fringes in the observation plane. If these sources are displaced in angle α, so too will be the fringe patterns. The total will have good visibility until the peak of one fringe pattern corresponds to the minimum of the other. This happens when the offset, αL, is equal to one half the fringe period;

$$\begin{aligned} \alpha L &= \frac{\lambda L}{D} \\ \alpha &= \frac{\lambda}{D} \end{aligned} \tag{A15}$$

In other words, if α is less than this limit, the fringe pattern will have good visibility. But recall, that the fringe pattern will have good visibility only if the fields in the plane of the apertures are coherent. Turning the problem around, we find that if the separation of the apertures, D, is such that

$$D < \frac{\lambda}{\alpha} \tag{A16}$$

then the fields will be coherent and we will observe good fringe visibility. Thus we arrive an expression for the *coherence diameter* of a source;

$$D_c \sim \frac{\lambda}{\alpha} \tag{A17}$$

where α is the angular subtense of the source. As an example consider the sun. It has an angular subtense of approximately 30 min of arc (one half degree). The corresponding spatial coherence diameter in the midvisible range ($\lambda = 0.55\ \mu m$) is approximately 0.06 mm. On the other hand, for the star Betelguese, which has an approximate angular subtense of 0.047 sec of arc (0.228 μrad), the coherence diameter at this wavelength is about 2.4 m.

Again, from this discussion we conclude that spatial coherence is not an intrinsic property of the source, but is inversely related to its angular subtense, i.e., its *apparent* size. Clearly, it is possible to introduce a degree of spatial coherence in a field by limiting the apparent size of its source.

Sources

In order for interferometry to have any practical application, there must exist a source capable of generating coherent optical energy. Probably the most used source of coherent light for interferometry and holography is the laser. But we have just seen that spatial and temporal coherence are not intrinsic properties of the light source. So what is it that makes the laser such a special light source? To

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

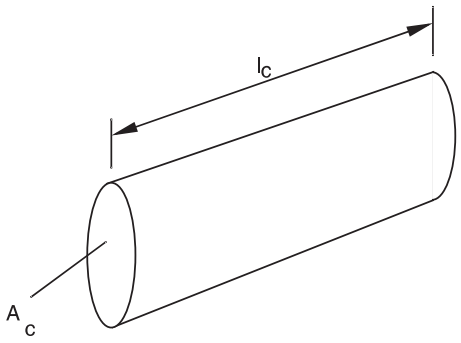


FIGURE A.3 Illustration of coherence volume.

answer this question we introduce the concept of the coherence volume. This is a right circular cylinder (see Figure A.3) of length equal to the coherence length and cross-section equal to the coherence area, A_c;

$$A_c = \pi\left(\frac{D_c}{2}\right)^2 \tag{A18}$$

For a blackbody radiator, it can be shown that the number of photons within such a coherence volume is given by

$$N = \frac{1}{\exp\{h\nu/kT\} - 1} \tag{A19}$$

where h is Planck's constant, k is Boltzman's constant and T is the absolute temperature in Kelvins. This expression is known as the Bose-Einstein distribution, and N is commonly known as the *degeneracy* parameter. For our sun, which is an approximate blackbody radiator at 5900 K, and at a wavelength of 0.6328 μm, the degeneracy parameter is approximately 0.022. On the other hand, for a well-stabilized (single longitudinal mode) HeNe laser of 1 mW, $\Delta\nu = 500$ Hz, and $\lambda = 0.6328$ μm, the degeneracy parameter is 6.4×10^{12}.

This points out the benefit of the laser, its ability to achieve an enormous degeneracy parameter. Now what does this mean? It means that it can produce a huge number of correlated photons. If they are correlated, they are indistinguishable, hence the nomenclature "degenerate."

The Speckle Phenomenon

The final topic of this appendix is a directly observable effect of coherence, laser speckle. Consider the problem of illuminating a piece of ground glass with a laser beam. (Figure A.4). The field impinging from the left hand side is assumed

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

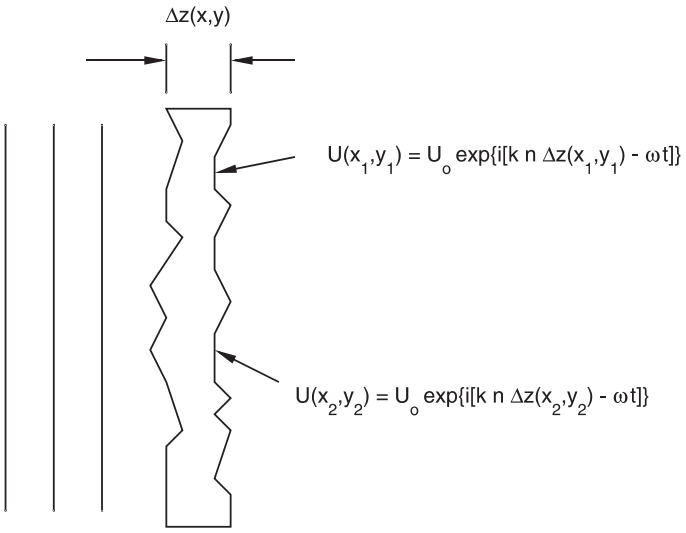


FIGURE A.4 *Gedanken* (German for "thinking") problem in which a ground glass is transilluminated with a coherent beam.

monochromatic and spatially coherent. We characterize the transmission of the ground glass as giving rise to a phase (dictated by its local physical thickness and refractive index) that varies in a complex manner as a function of the transverse coordinates. Although the field at one point may have a phase difference with respect to that at another point on the right-hand side of the ground glass,

$$U(x_1, y_1, t) = U(x_2, y_2, t) \exp\{ikn[\Delta z(x_2, y_2) - \Delta z(x_1, y_1)]\} \tag{A20}$$

we see that the temporal behavior of the field at these two points is identical. It follows that the fields on the right hand side of the ground glass are still perfectly spatially coherent. The only effect of the ground glass is to put a "crinkle" on the wavefront.

Now we wish to explore the effect of coherently illuminating an optically rough surface (see Figure 10.26). The field at an observation point removed from this coherently illuminated surface can be considered to be due to contributions from fictitious secondary point sources on the surface, much as in Huygens' principle. As in the previous discussion, the scattered field is still spatially (and temporally) coherent. The contributions for each of these secondary source points are randomly phased because of the rough surface. The effect is to produce a highly structured intensity variation in the observation plane that is referred to as laser speckle. These speckle patterns, resulting from the illumination of a surface

Copyright © Marcel Dekker, Inc. All rights reserved.

with highly coherent light, were first observed after the invention of the laser in the early 1960s, although similar phenomena were studied as early as the nineteenth century (Dainty, 1975).

These randomly distributed bright and dark spots, known as speckles, are the result of interference between light reflected from many independent scattering areas on a rough surface. In a simple (nonimaging) free space geometry, the intensity of an objective speckle pattern in the observation plane is a result of the superposition of the complex amplitudes of the wavelets scattered from each point on the surface. For geometries where the surface is imaged onto a detector, a subjective speckle pattern is obtained which also includes the effects of diffraction from a finite aperture. An objective speckle pattern would be observed, for instance, if one were to replace the observation plane of Figure 10.27 with a CCD camera without a lens. If one were to place a lens on the camera, one would observe a subjective speckle pattern. The nomenclature alludes to the fact that details of the observed speckle pattern may be subject to the "interpretation" (hence "subjective") of the imaging system. In the literature, one sometimes sees these objective and subjective classifications referred to as Fresnel and Fraunhofer.

A statistical treatment of the objective speckle phenomenon treats speckle as a random walk (Dainty, 1975). Under this analysis, it is found that for intensity measurements of polarized speckle the standard deviation equals the mean intensity and, therefore, the contrast, defined as the ratio of the mean intensity to the standard deviation, is always equal to one.

For the case of objective speckle, the smallest speckle in the observation plane is

$$\sigma_{\text{lat}} = \frac{\lambda L}{D} \tag{A21}$$

where the subscript denotes the lateral dimension, L is the distance between illuminated surface and observation plane, λ is the wavelength, and D is the size of the illuminated spot. This formula can be understood from the standpoint of the Young's double slit experiment. Consider the sketch in Figure A.5 where we have illustrated a rough surface illuminated with a spot of linear dimension, D. Consider two points at the extreme edges of the illuminated spot and determine the field at the observation plane. From our previous discussion of Young's double slit experiment, we know that we would observe the sinusoidal intensity pattern

$$I(x, L) = I(0, L)\left[1 + \cos\left(\frac{kxD}{L}\right)\right] \tag{A22}$$

and the period of this fringe pattern is

$$T = \frac{\lambda L}{D} \tag{A23}$$

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

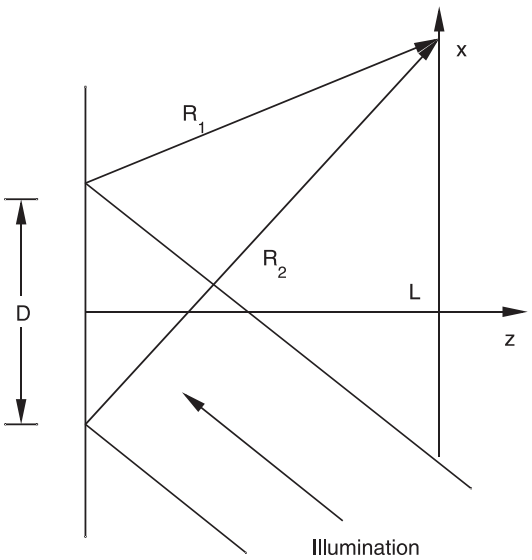


FIGURE A.5 Construction for estimating lateral extent of smallest speckles.

All other possible point pairs of scatterers are closer together. As a result, the corresponding fringes will have larger spatial periods. This expression thus represents the smallest speckle that one would observe in the plane $z = L$.

A very similar argument using Figure A.6 leads to an expression for the axial extent of the smallest speckles. Here, we inspect the interference pattern caused by a scatterer that is closest to the observation point (L) with one that is the furthest (R_1). The conclusion is that the smallest axial extent of a single speckle is of the order

$$\sigma_{\text{axial}} = 8\lambda\left(\frac{L}{D}\right)^2 \tag{A24}$$

This is a rough estimate of how much we can move the observation plane before the speckle pattern changes appreciably. If we look at the ratio of the axial and lateral minimum speckle sizes,

$$\frac{\sigma_{\text{axial}}}{\sigma_{\text{lat}}} = 8\frac{L}{D} > 1 \tag{A25}$$

we are led to the conclusion that the speckles are cigar-shaped with the long axis perpendicular to the illuminated spot.

For subjective speckle, similar arguments lead to the equations for the minimum speckle size except that the observation distance is the imaging

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

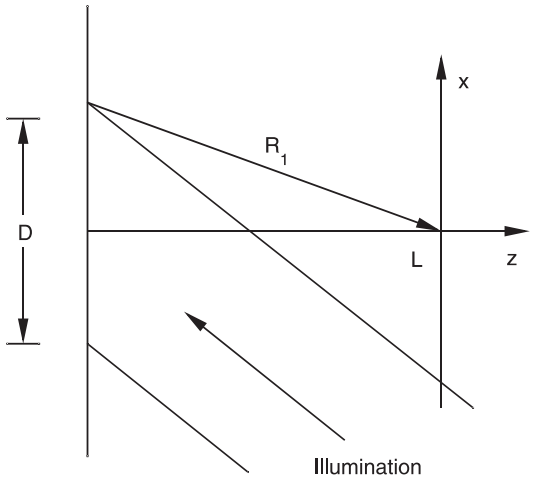


FIGURE A.6 Construction for estimating longitudinal extent of smallest speckles.

distance, d_i, and the D is interpreted as the diameter of the pupil of the imaging system;

$$\begin{aligned} \sigma_{\text{lat}} &= \frac{\lambda d_i}{D} \\ \sigma_{\text{axial}} &= 8\lambda\left(\frac{d_i}{D}\right)^2 \end{aligned} \tag{A26}$$

With the familiar thin lens imaging condition,

$$\begin{aligned} \frac{1}{d_i} + \frac{1}{d_o} &= \frac{1}{f} \\ 1 + \frac{d_i}{d_o} &= \frac{d_i}{f} \\ d_i &= (1+m)f \end{aligned} \tag{A27}$$

where f is the focal length of the lens and m is the image magnification, we can write

$$\begin{aligned} \sigma_{\text{lat}} &= (1+m)\lambda f^{\#} \\ \sigma_{\text{axial}} &= 8\lambda(1+m)^2(f^{\#})^2 \end{aligned} \tag{A28}$$

Note that the expression for the lateral extent of the smallest speckle is the same as for the PSF. Further, the expression for the axial extent of the smallest speckle

Copyright © Marcel Dekker, Inc. All rights reserved.

is a rough estimate of the *depth of focus* (distance the observation plane can be moved before the speckle pattern changes) of the speckle pattern.

These speckle effects can be observed directly by inspecting a laser beam illuminating a painted wall surface (which is typically rough compared to the wavelength). When one looks at the spot, a subjective speckle pattern is cast onto the retina. The pupil of the eye corresponds approximately to the exit pupil of the "imaging system." If the head is moved slowly side to side, the speckle pattern will "twinkle." The analysis above suggests that if the $f^{\#}$ of the eye were increased (smaller pupil) the speckles would appear larger. This can be accomplished by punching a small hole in a piece of thin cardboard, placing it as close as possible to the eye, and again peering at the speckle pattern. The speckles should appear appreciably larger.

APPENDIX B: POSTSCRIPT PROGRAMMING

PostScript is a language used by many printers. For such a printer, a given computer application will convert all the printing and plotting commands to this language prior to sending the output to the printer. One can also program directly in PostScript as the following examples demonstrate. These programs can be created with any text editor and then sent to the printer using the manufacturer's print utility. An alternative is to use a "Ghostscript" graphical interface such as "GSView," available from http://www.ghostgum.com.au/.

```
%! Linear grating, 1 1p/mm
/period 2.83465 def % period of grating
%
1.41732 setlinewidth % width of ruled lines
%
/height 576 def % height of ruled region
%
/width 72 def % width of ruled region
%
306 396 translate % reset origin at page center
90 rotate % rotate page about page center
%
-288 -36 translate % reset origin
0 period height {0 moveto 0 width rlineto stroke}
  for
show page
```

FIGURE B.1 PostScript commands to produce grating shown in the text. There are 72 PostScript "points" to the inch.

Copyright © Marcel Dekker, Inc. All rights reserved.

```
%! Linear grating, 1 1p/mm
/period 2.83465 def % period of grating
%
1.41732 setlinewidth % width of ruled lines
%
/height 576 def % height of ruled region
%
/width 72 def % width of ruled region
%
% draw first grating
306 396 translate % reset origin at page center
100 rotate % rotate page about page center
-288 -36 translate % reset origin
0 period height {0 moveto 0 width tlineto stroke}
  for
%
%draw second grating
288 36 translate % reset origin at page center
-20 rotate % rotate page about page center
-288 -36 translate % reset origin
0 page height {0 moveto 0 width rlineto stroke}
  for show page
```

FIGURE B.2 PostScript commands to produce moiré pattern shown in the text.

For other PostScript programs see the articles by Knotts (1996) or Clark *et al*. (1991).

REFERENCES

N Abramson. *Light in Flight or The Holodiagram: The Columbi Egg of Optics*. Bellingham: SPIE Engineering Press, 1996.

CF Bohren, DR Huffman, *Absorption and Scattering of Light by Small Particles*. New York: John Wiley & Sons, 1983.

M Born, E Wolf. *Principles of Optics, 6th edition*. New York: Pergamon Press, 1989.

VA Borovikov, B Ye Kinber. *Geometrical Theory of Diffraction*. London: Institute of Electrical Engineers, 1994.

RN Bracewell. *Two-Dimensional Imaging*. Englewood Cliffs: Prentice-Hall, 1995.

C Brosseau. *Fundamentals of Polarized Light: A Statistical Optics Approach*. New York: John Wiley & Sons, 1998.

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

WT Cathey. *Optical Information Processing and Holography.* New York: John Wiley & Sons. 1974.
GW Clark, YN Demkov. 'Making zone plates with a laser printer.' Am J Phys, **59**: 158–162, February 1991.
G Cloud. *Optical Methods of Engineering Analysis*. Cambridge: Cambridge University Press. 1995.
E Collett. *Polarized Light: Fundamentals and Applications*. New York: Marcel Dekker Inc., 1992.
RJ Collier, CB Burckhardt, LH Lin. *Optical Holography.* San Diego: Academic Press, Inc., 1971.
JC Dainty, ed. *Laser Speckle and Related Phenomena*, Faller-Verlag, 1975.
JC Dainty, ed. *Laser Speckle and Related Phenomena Second Enlarged Edition*. Berlin: Springer-Verlag, 1984.
YN Denisyuk, Soviet Phys.-Dokl., **7**: 543, 1962.
RK Erf. *Holographic Nondestructive Testing*. New York: Academic Press, 1974.
RK Erf, ed., *Speckle Metrology*, New York: Academic Press, 1978.
GR Fowles. *Introduction to Modern Optics, Second Edition*. New York: Dover Publications, Inc., 1975.
M Françon. *Laser Speckle and Applications in Optics*. New York: Academic Press, 1979.
D Gabor. Microscopy by Reconstructed Wavefronts, Proc. Roy. Soc, 197: 454, 1949.
JD Gaskill. *Linear Systems, Fourier Transforms, and Optics*. New York: John Wiley & Sons, 1978.
KJ Gåsvik. *Optical Metrology Second Edition*. Chichester: John Wiley & Sons, 1995.
DC Ghilia, MD Pritt. *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*. New York: John Wiley & Sons, 1998.
JW Goodman. *Statistical Optics*. New York: John Wiley & Sons, 1985.
JW Goodman. *Introduction to Fourier Optics. Second edition*. San Francisco: McGraw-Hill, 1996.
R Guenther, *Modern Optics*, New York: John Wiley & Sons, 1990.
R Greenler. *Rainbows, Halos, and Glories*, Cambridge: Cambridge University Press, 1980.
P Hariharan. *Optical Holography: Principles, Techniques, and Applications*. Cambridge: Cambridge University Press, 1984.
OS Heavens and RW Ditchburn, *Insight into Optics.* Chichester: John Wiley & Sons, 1991.
E Hecht. *Optics, 3rd edition*. New York: Addison-Wesley, 1997.
K Iizuka. *Engineering Optics, Second Edition*. Berlin: Springer-Verlag, 1983.
GL James. *Geometrical Theory of Diffraction for Electromagnetic Waves, 2nd edition*. London: Institute of Electrical Engineers, 1980.
FA Jenkins, H E White. *Fundamentals of Optics*. New York: McGraw-Hill, 1957.
R Jones, C Wykes. *Holographic and Speckle Interferometry. A Discussion of the Theory, Practice and Applications of the Techniques*. Cambridge: Cambridge University Press, 1983.
O Kafri and I Glatt. *The Physics of Moiré Metrology.* New York: John Wiley & Sons, 1990.
JE Kasper, SA Feller. *The Complete Book of Holograms: How They Work and How to Make Them*. New York: John Wiley & Sons, Inc., 1987.
M Knotts, Fun with moiré patterns, *Optics and Photonics News*, pp 54–55, August 1996.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

T Kreis. *Holographic Interferometry Principles and Methods*. Berlin: Akademie Verlag, 1996.
L Lading, G Wigley, P Buchhave, eds. *Optical Diagnostics for Flow Processes*, New York: Plenum Press, 1994.
EN Leith, J Upatnieks, Reconstructed Wavefronts and Communication Theory, J. Opt. Soc. Am., **54**: 1295, 1964.
D Malacara. *Optical Shop Testing*, New York: John Wiley and Sons, 1978.
D Malacara, M Servin, Z Malacara. *Interferogram Analysis for Optical Testing*. New York: Marcel Dekker, Inc., 1998.
M Mineart. *The Nature of Light and Color in the Open Air*. Dover Publications, Inc., 1954.
JA Ogilvy. *Theory of Wave Scattering from Random Rough Surfaces*. Bristol: Institute of Physics Publishing, 1991.
EL O'Neill. *Introduction to Statistical Optics*. Reading MA: Addison-Wesley, 1963.
PK Rastogi, ed. *Optical Measurement Techniques and Applications*. Boston: Artech House, Inc., 1997.
GO Reynolds, JB DeVelis, GB Parrent, Jr., and BJ Thompson. *The New Physical Optics Notebook: Tutorials in Fourier Optics*. Bellingham WA, SPIE Engineering Press, 1989.
B Saleh and M Teich. *Fundamentals of Photonics*. New York: John Wiley-Interscience, 1991.
AE Siegman. *Lasers*, Mill Valley: University Science Books, 1986.
CS Williams, OA Becklund. *Introduction to the Optical Transfer Function*. New York: Wiley-Interscience, 1989.
JW Wagner. Optical detection of ultrasound. In: RN Thurston, AD Pierce, eds. Ultrasonic Measurement Methods. Volume XIX in series Physical Acoustics. Boston: Academic Press, Inc., 1990.
WJ Smith. Modern Optical Engineering. New York, NY: McGraw-Hill, 1966.

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Index

Ablative, laser generation, 134
Abrasive cleaning, PT technique, 36–37
Absolute coil probe, 308–310
  RFT data, 354
Absolute EC probes, 307–308, 310
Absolute loss factor, 655
Absolute permittivity, 655
Absolute probe, definition, 361
Absorbed dose, definition, 579
Absorbed dose rate, definition, 580
Absorption, 470
  and attenuation, 107, 109, 111
  definition, 580
  physical and mathematical representation, 112–114
Absorption edge (radiation), 476
  definition, 580
Absorption holograms, 740–743
Accelerating potential, definition, 580
Accelerometers, 396
Acoustic emission (AE), 369–443
  activity, 406–407
  advantages, 372
  analysis techniques, 406–410
  applications, 422–440
  audible, wood, 369
  definition, 369–370, 442
  fundamentals, 377–406
  historical perspective, 372–375
  measurement, 395–406
  measurement systems, 401–406
  potential, 375–377
  seismic analogy, 371
  sources, 377–383
  technical overview, 369–370
  testing complications, 371–372
  waveform analysis, 419–422
  wave propagation, 383–395

Copyright © Marcel Dekker, Inc. All rights reserved.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Acoustic energy, dissipation, 394
Acoustic Fresnel equations (UT)
  definition, 185
  normal incidence wave, 96
Acoustic NDE, 1
Acoustic pressure, 89–92
  continuity, 178–180
Acoustic waves, description, 173
Active element, piezoelectric transducer, 127–128
Active thermographic measurement, temperature distribution, 598
Active thermography, 597–644
  background, 597–604
  components, 598
  data analysis, 627
  experimental setup, 603–604
  heating methods, 628–629
    NDE problem, 632
  heating source, 628–630
  historical perspective, 599–602
  measurements, 631–636
  NDE, 619
  nondestructive evaluation, 597
  problems, 641–642
  spatial heating patterns, 629–630
  specific applications, 636–640
  technique overview, 592–602
  techniques, 602–604, 619–622
  variables, 642
Activity curve shape analysis, 407–408
Activity monitoring, AE, 401–402
Adhesive force, PT fluid flow principle, 25
Advanced Light Source (ALS) synchrotron, 468
Aerospace applications, EC, 355–357
Aerospace engines, components, 355–357
Aerospace industries, PT application, 53–54
Aerospace Materials Specification (AMS), 51
Aerospace structures
  fatigue crack detection, 435–440
  test description, 436–440
Agencies, radiation safety responsibilities, 578
Agreement States, U.S. Nuclear Regulatory Commission, 573–574
Air-coupled mechanism, piezoelectric transducer characterization, 122–124
Aircraft, maintenance cycle, 356
Air-gap thickness, disbond thickness, 702
Airlines, retirement-for-cause NDE, 11
Airports, inspection, 535
Alpha particle, definition, 580
Alternating current, 217–218
  definition, 361
Alternative transverse EMAT, 132
Aluminum
  liquid penetrant dwell time, 41
  surface breaking cracks, flaw response trajectories, 326
Aluminum plate, AE signal, 399
Ambient temperature, piezoelectric transducer characterization, 126
American Society for Nondestructive Testing (ASNT), 52, 574
American Society for Testing and Materials (ASTM)
  standard test methods, 51
American Society of Mechanical Engineers (ASME), 52
Amorphous thin-film semiconductor technology, 509
Amplification, 114
Amplitude
  ratios, normal incidence wave (UT), 96
  response, transducer characteristic (UT), 139–140
  splitting, 734
  ultrasonic inspection principle, 148
  of wave, 72
AMS (*see* Society of Automotive Engineers (SAE))
AMS-2645, 51
AMS-2646, 51
AMS-3155 to 3158, 51
Analog radiation detector (film), 503–504
  disadvantages, 503

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

[Analog radiation detector (film)]
efficiency, 504
exposure, 503
hazardous waste, 503
latent image, 503
overdeveloped, 503
structure, 503
Analog film, vs. digital radiographs, 528–529
Analog radiographic image, definition, 580
Analog to digital converter, 401
definition, 580
Angled beam (UT)
piezoelectric transducer characterization, 124
transducer illustrated, 125
ultrasonic inspection principle, 156
Ångstrom, definition, 580
Angular frequency, 655
Angular spatial frequency, 73
definition, 186
Angular temporal frequency, 73
definition, 186
Anisotropic, definition, 185
Anisotropic materials, 388–390
slowness surfaces (UT), 180
Anisotropy, rubber compound, 686
Annihilation photons, 473
Anode, definition, 580
Anode current, definition, 580
Anomaly, definition, 361
ANST standard tomographic MTF, 518–519
Antennas
horn, 666
microwave, 666
Aperture size, transducer characteristic, 139
Aperture stop, definition, 788
Applied strain, UT bulk waves, 80
ARACOR, 535
Arctan function, 701
Area-array (2D) radiation detector-based system, 538–539
Area-array detectors (radiography)
source utilization, 521
spatial distortions, 538–539
systems, 536–547
Area-array parallel beam DR/CT systems, 541–542
Areal density, 475
Array, piezoelectric transducer characterization, 124
Artifact, definition, 580
Artifact standards, EC calibration, 350
A-scan, definition, 185
ASTM E-165, 51
ASTM E-1220, 51
ASTM E-1418, 51
Atoms
excited state, 464
ground state, 464
Attenuation, 107–114, 391–394, 418
definition, 185, 442
by dispersion, 114
geometric, 113–114
mathematical characteristics, 112–114
physical characteristics, 109–112
ultrasonic waves, 64
Attenuation coefficients, 473–479
definition, 580
Attenuation cross section (radiation), definition, 580
Attenuation mechanisms, frequency dependency, 108
Attenuation spiral, RFT, 346–348
Audible AE, wood, 369
Automated eddy current inspection system, automotive steering push rods, 324
Automotive industry, PT application, 56
Automotive steering push rods, automated eddy current inspection system, 324
Automotive water pump
CT image, 543
digital radiographic image, 543
Axles, ultrasonic testing, 167–170

Backscatter, 528
Back scattered ration, definition, 580

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Bandwidth
  definition, 185
  transducer characteristic, 140
Barium titanate (BaTi)
  piezoelectric material, 122
Bar magnet
  vs. solenoid
  magnetic field, 199
Barn, definition, 580
Base density, 504
Bases, 784
  vendors, 785
Beam characterization, transducer, 141–147
Beam divergence, 145–147
Beam lenses, 784
Beamlines, 468
Beam splitters, 784
Beam spreading
  and attenuation, 107
  definition, 185
  physical and mathematical representation, 112–114
Becquerel, 495
  definition, 580
Beer's law, homogeneous material, 476
Bel, definition, 113
Bell, A.G., 113
  NDE measurements, 600
Beryllium sheet, 558–559
Betatron, definition, 580
Biot-Savart law, 270
Birefringence
  definition, 185, 789
  light wave, 93
Black light
  definition, 60
  PT illumination principle, 30
Blades, aerospace engines, 356
Blooming, definition, 580
B-mode, definition, 185
Bobbin coil, definition, 361
Bobbin probe, 302, 308–310
  definition, 361
  EC probe windings, 308
  tube, normalized impedance plane, 306
Boilers, RFT damage, 352
Boltzmann's constant, 789
  Plank's constant, 798
Boundary, reflection and refraction, 92–107
Bounded waves, guided waves (UT), 117–120
β particles, 479
Bragg's Law, 550–551
Braking radiation, 467, 488
Brass, liquid penetrant dwell time, 42–43
Breast cancer, X-rays, 452
Bremsstrahlung, 467, 488, 490
  definition, 581
Bridge circuit, definition, 361
Bridge pins
  CT, 543–544, 567
  digital radiography, 543–544, 567
  failure, investigation, 562–563
  photographs, 543
Brightness, 468, 469
  definition, 468
Brightness-contrast dyes, PT illumination principle, 29
Broadband
  photon radiation, 468
  piezoelectric transducer characterization, 125
Broad beam attenuation, definition, 581
Broad beam geometry, 480
Bronze, liquid penetrant dwell time, 42–43
Brushing, PT technique, 39
B-scan, definition, 185
Bubblers, 128
Bulk longitudinal wave, 80
Bulk modes, isotropic materials, 383–384
Bulk waves, 80–89
Burmah-Castrol strips, 244
Burst emission, definition, 442

Calibration standard, definition, 361
Capacitive reactance, definition, 361
Capacitor, definition, 361

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Capillarity
  definition, 60
  PT fluid flow principle, 26–28
Capillary, PT fluid flow principle, 26–28
Capillary action, PT fluid flow principle, 26–28
Carbide tipped tools, liquid penetrant dwell time, 43
Carbon fiber contaminants
  detection of, 638
  microwave heating, 638
Carbon filters, PT technique, 46
Cargo containers, inspection, 535
Carrier fringes, 767
Cassette, definition, 581
Cathode ray tube, 450, 452
  Röntgen, Wilhelm, 452–453
Central conductors (MP), 226–228
  definition, 255
Centrifuge tube, particle concentration (MP), 238
Ceramics
  materials, PT application, 56–57
  piezoelectric material, 122
  porosity levels, 690
  quality assessment, 374
Characteristic curve, 525–526
  definition, 581
Characteristic length, definition, 361
Characteristic (limiting) frequency, 361
Characteristic parameter, definition, 361
Charge-coupled devices (CCD), 508–509, 627, 722
  cameras, 785–786
    vendors, 786
  costs, 514
  definition, 789
  radiation detectors, 500
Chief ray, definition, 789
Chrysippus, 66
Cine-radiography, definition, 581
Circular aperture transducer, 145–147
Circular defect, ultrasonic reflection, 65
Circular magnetization, 211–213
  definition, 255
  vs. longitudinal, 213–215
Circumferential coil, definition, 361
Cleaning, PT technique, 32–37
Clients, radiation safety responsibilities, 578
Coatings, EC sensing, 334
Coercive force, definition, 255
Coherence diameter, 797
Coherence electromagnetic fields, 794
Coherent scattering, 462, 470
Cohesive forces, PT fluid flow principle, 22–24
Coil fill factor, definition, 362
Coils (EC), 222–223
  Bobbin, definition, 361
  circumferential, definition, 361
  excitation, 262
  external reference, 311–312
  feed-through, definition, 362
  inductive coupling, 273
  internal reference, 311–312
  parameters, EC probes, 314–315
  pickup, 262
  probe, 308–310
    RFT data, 354
  reference
    definition, 365
    EC probes, 311–312
  self-reference, 311–312
  single-loop, Faraday's law, 314–315
  test, definition, 365
  tube, normalized impedance plane, 305
Coils (MP)
  demagnetization, 237
  hair pin, 256
  solenoid, 212
  split, 258
Collimator, 498, 513–515
  definition, 581
Color, definition, 789
Color-contrast
  penetrant, PT technique, 38–39
  PT illumination principle, 29
Color plots, 159
Commercial liquid penetrant testing system, schematic and photograph, 33

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Composite materials
  dielectric, inspection of, 695–703
  measured amplitudes, 418
  permeability, 661
  permittivity, 661
  porosity levels, 690
  TMC study, 429–435
Compression wave, propagation
    illustrated, 77
Compressive force, PT fluid flow
    principle, 22–24
Compressor disk, aerospace engines, 356
Compton scattering, 462, 470
  definition, 581
Computed radiology, definition, 581
Computer aided tomography (CAT) scan
  heart, 451–452
  invention, 457
Computer tomography (CT), 449, 458,
    528–529, 558–570
  angular sampling, 519
  configuration, 529
  data acquisition geometry, 522
  definition, 463–464
  detector systems, 538
  digital (projection) radiography,
    558–571
  high explosive materials, process
    control, 566–571
  radiology, 481–484
  reverse engineering, 559–562
  scanners, preprocessing steps, 525
Concrete-based materials, 690
Conductive coating, 334
  impedance loci, 335
Conductive materials, EC probe, 263
Conductivity, 358
  definition, 362
Conductors
  definition, 362
  microwave NDE, 652
  spacing between, EC, 334–336
Cone-beam CT geometry systems, 541
Cone-beam geometry, 521
Cones (human eye), 28, 31
  cone response curve, 28
Constructive interference
  definition, 187
  diffraction beam characterization, 141
  ultrasound, 170–171
Contact angle, PT fluid flow principle,
    24–26
  listed cases, 25–26
Contact mechanism, piezoelectric
    transducer characterization,
    122–124
Continued emission, definition, 443
Continuity of acoustic pressure, 178–180
Continuity of particle velocity, 178–180
Continuity-of-phase law (*see* Snell's law)
Continuous laser sources, 783
Continuous method (MP), definition, 255
Continuous recording, 404–405
Continuous source, AE, 378
Continuous wave, definition, 699
Continuous wave generator (UT),
    definition, 186
Continuous wave pulsers (UT), 157
Contrast sensitivity (radiography),
    definition, 581
Contrast stretch, definition, 581
Conventional film, holographic film, 741
Core soft metal, definition, 255
Corrosion resistant steels, 204
  aircraft, 355–357
Coupling mechanism, piezoelectric
    transducer characterization,
    122–124
Cracks
  aircraft, 355–357
  depths, flaw response trajectories,
    328–329
  impedance plane measurement (EC),
    326–334
    probe diameter, 333–334
    subsurface-breaking, 327
    surface-breaking, 326–330
Creeping waves, 116
Critical angle (UT), definition, 186
  Snell's law, 101–102
Crystalline material, X-ray diffraction,
    551

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

C-scan, 159–160
  definition, 186
Cumulative dose, definition, 581
Curie, 495
  definition, 581
Curie brothers, 67
Curie temperature, definition, 255
Current-induced magnetic field, 269–271
Current integrating detectors, 507, 530
Current integrating systems, 529
Current levels, 221–222
Cycle
  period, 73
  wave propagation, 72

Damage activity, AE, 375
Damage location, AE, 376
Damage mechanism identification, AE, 376
Dark adaptation, human eye, 30
Dark current image, 540
Data analysis algorithms, step heating cases, 633
Data (contrast) resolution, 506
Data processing, optical methods, 765–771
Daughter product (radiation), 495
Dead zone, pulse-echo, 155
Decibel, definition, 186
Defect, definition, 60, 362
Deformation, bulk material, 83–89
Degeneracy parameter, 798
  definition, 789
Delamination, 695
  definition, 186
Demagnetization
  coils, 237
  hysteresis loop, 236
  sample, 235
Densitometer, definition, 581
Density
  definition, 581
  gradient, definition, 581
  images, 569–570
Department of Defense (DOD), 52
Depth informing techniques (radiography), 552–554
Depth of field, definition, 789
Depth of focus, 803
  definition, 789
Depth of penetration, definition, 362
Descaling, PT technique, 36
Destructive
  definition, 187
  interference
    diffraction beam characterization, 141
    ultrasound, 170–171
Detectors (radiography), 785
  blur, 521
  categorization, 529
  physical configuration, 524
  quantum efficiency, 521
  source efficiency, 521
Detector spectral sensitivity, 624–625
Detergent cleaning, PT technique, 32–35
Developers
  coating, 46
  definition, 60
  types, PT technique, 45–48
Diamagnetic magnetization, 202
  definition, 256
Dielectric
  composites, inspection of, 695–703
  constant, 652–656
  definition, 789
  insulators, 652–653
  interface, electric magnetic fields, 697
  material characterization, 679–692, 692–695
  microwave NDE, properties, 656
  plates, 701
Differential EC probes, 307–308, 310
Differential probe (EC), definition, 362
Diffraction
  beam characterization, 141
  definition, 582, 789
  efficiency, 743–744
    hologram, 744
  grating, definition, 789
  transducer characteristic, 139

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Diffusion equation, 608
Diffusive heat flow, 608
Digital detectors (radiation), 506–507
  vs. film, 513
Digital image
  definition, 582
  enhancement, definition, 582
Digital radiography, 520, 528–529, 558–571
  vs. analog film, 528–529
  with computed tomography, 535
  configuration, 529
  frame rates, 528
Digitized film radiographs
  flat butt welds, defects, 557
  teeth, 450
Dilatation wave, 80
  definition, 186
  propagation, 89
Dim-light, PT illumination principle, 29
Dipole, definition, 256
Dipping, PT technique, 39
  degreasing, 35–36
Dirac's delta function, 615
Direct current, 217
  definition, 362
Direction cosines, 728
  definition, 789
Directivity patterns
  beam divergence, 145
  far field, 146
Direct piezoelectric effect, 120
Disbond thickness, 701–702
  air-gap thickness, 702
  detecting, 703
  evaluating 703
  remedies, 702
Discontinuity
  definition, 60, 362
  signatures, RFT, 348–349
Discrete source, AE, 378
Disintegration, 495
  rate, 495
Dispersion, 394
  and attenuation, 107
  attenuation by, 114
[Dispersion]
  curves
    definition, 186
    Lamb waves, 119
    plate waves, 174
  definition, 186, 443
Displacement, bulk material, 83–89
Distance, radiation exposure reduction, 574–575
Divergence, and attenuation, 107
Domain, definition, 256
Doppler effect, definition, 186
Doppler frequency shift, oscillator, 671
Doppler radar, 671
Doppler shift, 736
  definition, 789
Dose, radiation, 484
Dose equivalent, radiation, 484, 485
Dosimetry, 484–486
Double-coil probes (EC), 315
Double exposure holographic interferometry, 752
  illustration, 753
Drain off period, lipophilic emulsifiers, 44
Drivers
  definition, 186
  excitation pulsers, 156–158
Drug detection, fan-beam digital radiography system, 534
Dry method (MP), definition, 256
Dry powders, developer, PT technique, 46
Dual-element transducer
  illustrated, 126
  piezoelectric transducer characterization, 124
Dual energy, 478
Dual refractive index technique, 756
Dunegan corollary, 408
Dwell time
  definition, 60
  liquid penetrant, 41–43
  PT fluid flow principle, 28
  PT procedure, 19
  PT technique, 39–40
Dye, visible, PT technique, 38

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Dye penetrants, 37
  methods, 20
Dyes, penetrant, PT illumination principle, 29
Dynamic range, 524
  definition, 582
  vs. precision, eddy current, 316–317

Earing, NDE, 3–4
Echo-reflection, 68
Echo transmission, ultrasonic NDE, 106–107
Eddy current (EC), 261–365
  advantages/disadvantages, 267–268
  applications, 350–357
  coil impedance, 279–280
  decay, 287
  definition, 256, 362
  density, skin depth, 285–287
  discovery, 264
  electromagnetic principles, 268–305
  example, 274–279
    qualitative results, 275–279
  geometric and electromagnetic factors, 274
  historical development, 264–266
  inductive impedance, 280
  inspection principles, 325–350
  instrument display, computer based, 321
  magnetic induction, 269–274
  measurement equipment, 316–325
  measurement optimization, 336–343
  method, definition, 362
  NDE methods, 604
  phase lag, 287
  phasor notation and impedance, 280–285
  potential, 266–267
  probes, 263, 275, 283, 307–316
    coil parameters, 314–315
    common configurations, 308–310
    cracks, 333–334
    differential, 307–308, 310
[Echo transmission, ultrasonic NDE]
    guiding and shielding magnetic fields, 312–314
    mutual inductance, 272–274
    reference coils, 311–312
    sinusoidal excitation, 288
  resistive impedance, 279–280
  resonant circuit, 322
  sensing, coatings, 334
  technical overview, 262–263
  transducers, 306–315
    probes, 307–308
Edge effect (EC), definition, 362
Effective depth of penetration (EC), definition, 362
Effective dose, 485
Effective dynamic range, 524
Effective permeability (EC), definition, 362
Effective resonator capacitance, 663
Effective-Z (radiography), 569–570
Elastic forces, PT fluid flow principle, 22–24
Elastic medium, 143
Elastic potential energy, PT fluid flow principle, 23
Elastic scattering, 462
Elastic wave equation, 83
Elastomeric O-ring, 3D-solid rendering, 547–548
Electrical components, energy, 285
Electrical power industry, PT application, 56–58
Electric flux density, 654
Electric magnetic fields dielectric interface, 697
Electric polarization vector, 654
Electric susceptibility, definition, 654
Electromagnetic acoustic transducers (EMATs), 129–132, 396
  definition, 186
  longitudinal wave generation, ultrasonic inspection principle, 149
  schematic, 129
Electromagnetic coupling, definition, 362

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Electromagnetic energy, wave equation, 725
Electromagnetic field, 794
  amplitude, optical detectors, 729
  coherence, 794
  properties, near-field region, 706
Electromagnetic induction
  definition, 362
  discovery, 264
Electromagnetic interference (EMI), 379, 419, 650
Electromagnetic theory, 648
  radiation, 650
Electromagnetic waves
  basic concepts, 656–664
  microwave NDE, 651
  propagation of, 656
  X-rays, 448
Electromotive force (EMF), definition, 362
Electronic bridge circuit, 317
Electronic speckle pattern interferometry
  definition, 789
  video holography, 762
Electronic transducers, AE, 374
Electrons, 464–465
  capture, 465
  radioactivity, 464
Electron volt, 459–460
  definition, 582
Elevator shafts, ultrasonic testing, 167–170
Emissivity, 623
Emitted signals, AE, 371–372
Empirical filing factor, 663
Employers, radiation safety responsibilities, 577–578
Emulsifying agent (PT), definition, 60
Emulsifying agents, PT procedure, 20
Encircling coil (EC), tube, normalized impedance plane, 305
Encircling probe, 308–310
  definition, 362
  EC probe windings, 308
Energy discriminating detector-based (radiographic), 507, 512–513, 548–550
Energy discriminating DR/CT detectors, 524
Energy integrating detectors, 507
Energy ratios (UT), attenuation, 113
Energy resolution, 524
Engineering shear strain, 79–80
  bulk material, 85
Entrance pupil, definition, 789
Epoxy coating, imaging entrapped water, 636–638
Epoxy matrix, 694
Equivalent conductivity, 655
Equivalent I.Q.I. sensitivity, definition, 582
Equivalent penetrameter sensitivity, definition, 582
Equivalent power, NEP, 626
Equivolumetric component (UT), 89
Euler expansion, definition, 789
Event duration (AE), sources, 418–419
Examination, PT, 48–49
Excitation pulsers, ultrasonic inspection principles, 156–158
Excited neutron state, 466
Excited state, 464
Exit pupil, definition, 789
Explosives
  detection, 535
  fan-beam digital radiography system, 534
Exposed film, archival storage, 528
Exposure (radiography)
  charts, 526–527
  definition, 582
  radiation, 484
  range, definition, 582
External factors, NDE reliability, 13
External reference coil, 311–312
Eye
  dark adaptation, 30
  frequency response, 29
  noise, 766

F-16
  bulkhead subcomponent, 436
  engine noise tests, 439

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Fabrication industries, PT application, 52–53
Fabry-Perot method, 186
Failure Mode and Effect Analysis (FMEA), 12
False color, 159
Fan-beam digital radiography system, drug detection, 534
Fan-beam geometry, 521
Fan-beam LAD radiography system
  LAD system geometry, 533
  with shielding system, 535
Fan disk, aerospace engines, 356
Faraday, Michael, 264
Faraday's law of induction, 288
  single-loop coil, 314–315
Far field (UT)
  definition, 186
  transducer characteristic, 143–145
Far field region (optical), definition, 704
Fast Fourier transform, definition, 790
Fast-responding light receptors, PT illumination principle, 31
Fatigue crack
  detection, aerospace structures, 435–440
  growth, waveforms, 437
  photograph, 438
Feature-based systems, AE, 405
Feature measurement, AE, 402–403
Feature set analysis, AE, 410–419
Feed-through coil, definition, 362
Felicity effect, 408
  definition, 443
Felicity ratio, 408
  definition, 443
Female X-ray technicians, 455–456
Ferrite
  cores, 313
  definition, 362–363
Ferroelectric domains, 122
Ferromagnetic, definition, 363
Ferromagnetic conductive material, 296–301
  EC, 278–279
  edge effects, 298–299
[Ferromagnetic conductive material]
  liftoff, 296–298
  material thickness, 300–301
  reversible (effective) magnetic permeability, 296–298
Ferromagnetic core, transformer, 278
Ferromagnetic domains, 122
Ferromagnetic magnetization, 202
  definition, 256
Ferromagnetic material
  hysteresis loop, 206
    diagram, 205
  magnetic domains, 203
  magnetic hysteresis, 203–207
Fiberoptic based techniques, 787
Field
  depth of, definition, 789
  flattening, 533
  point, definition, 727
Fill-factor, 288, 506
  definition, 363
Film, 525–528, 785 (*see also* Analog detector (film))
  characteristics, 504, 505
  contrast, definition, 582
  costs, 514
  vs. digital detectors, 513
  dynamic range, 519
  efficiency, 504
  processing, 527
  radiation detectors, 500
  radiography, density curve, 505
  speed, 506
    definition, 582
Filter
  circuit, 317
  definition, 582
Filtered particle penetrant, PT application, 57
Final cleaning, PT technique, 50
Finite-size material body, 660–661
Finite thickness, 610, 613–615
  NDE applications, 613
Firestone, 68
First critical angle, definition, 186
Flat butt welds, defects, digitized film radiographs, 557

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Flat panel detectors, 509–512
  arrays, 528
Flat transducer, piezoelectric transducer characterization, 124–125
Flaw
  definition, 60, 363
  ultrasonic waves, 65
Fluid flow, PT principle, 21–28
Fluid viscosity, PT fluid flow principle, 27
Fluorescence, definition, 60, 582
Fluorescent penetrant inspection, photograph, 55
Fluorescent penetrants, 37
  PT technique, 38
Fluorescent screens, radiation detectors, 501–502
Flux density (ϕ), 467
Flux-Flo (MP), definition, 256
F-number (optical), definition, 790
Focal spot, definition, 582
Focus, depth of, 803
  definition, 789
Focused wave generation, ultrasonic inspection principle, 152–153
Fog (radiography), definition, 582–583
Fog density (radiography), 504
  definition, 583
Forster, Federico, 266
Foucault, J.B., 265
Foucault currents, 265
  definition, 363
Fourier's law
  temperature distribution, 607–609
  thermal conductivity, 606
Fourier transformation, 615
  illustration, 769
  techniques, 768
Frame rates, 507
  digital radiographs, 528
Franklin, Benjamin, 451
  X-rays, 449
Fraunhofer zone (optical), definition, 790
Frequency
  angular, 73, 186, 655
  characteristic (limiting) (EC), 361
  definition, 186, 363
[Frequency]
  dependency, attenuation mechanisms, 108
  fundamental, 186
  linear spatial, 73, 186
  linear temporal, 73, 186
  modulated wave, ultrasound, 172
  sampling, AE, 405
  skew
    definition, 188
    transducer characteristic, 140
  spatial, 73, 74, 790
  transducer characteristic, 139–140
  ultrasonic inspection principle, 148–149
Fresnel equation
  acoustic
    definition, 185
    normal incidence wave, 96
  definition, 186
  normal incident wave, 99–100
  oblique incidence, examples, 102–106
  and Snell's law, 101
Fresnel reflection, definition, 790
Fresnel zone (optical), definition, 790
Fresnel zone plates, focused wave generation, ultrasonic inspection principle, 152–153
Fringe pattern
  illustration, 782
  light, 781
Front surface mirrors, 784
Fuel injector, CT image, 544
Full-width at half-maximum, definition, 790
Fundamental frequency, definition, 186

Galilei, Galileo, 66
Gamma-radiography, definition, 583
Gamma-rays
  attenuation, 475–479
  definition, 583
  mechanisms, 469
  spectroscopy, 529
Gas/liquid (PT), surface tension, 25
Gauge measurement, 448
  X-rays, 448

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Gauging, 463, 530–531
Gauss meter, 242
  definition, 256
Gedanken, 799
Generators, 157
Geometrical attenuation, 107, 113–114
  definition, 186
Geometric spreading, 392
  amplitude, 393
Geometric unsharpness, definition, 583
Giga (G), 64
Glass, liquid penetrant dwell time, 43
Graininess, definition, 583
Graphite-epoxy composite pressure vessels, impact damage, 425–429
  results, 428–429
  test, 425–427
Graphite-epoxy tubing, eddy current measurements, 324
Grass (UT)
  and attenuation, 109, 110
  definition, 186
Gray (Gy), 572
  definition, 583
Green-blue spectrum, PT illumination principle, 28
Greens functions, 422
Ground state, 464
  atoms, 464
Group velocity (UT), 171–173
  definition, 189
Guaging, 463
Guided waves (UT), 114–120
  bounded waves, 117–120
  longitudinal creeping waves, 116
  Love waves, 116
  Rayleigh waves, 115
  SAW, 114–116
  Scholte waves, 115–116
  Stoneley waves, 116
Guiding magnetic fields, EC probes, 312–314

Hair pin coil (MP), definition, 256
Half-value layer (HVL)
  definition, 583
Half-wave current (MP), 218–219
Hall effect meter, 241
  definition, 256
Halogen, definition, 60
Halos, 769
Hand
  CT segmentation, 563
  digital radiograph, 556, 561
Handheld EC units, 322–323
Hard magnets, definition, 256
Harmonic wave, 76
H-D curve, 525–526
Head- and tailstock beds, 224–226, 233
Head shot, definition, 256
Head wave (*see* Longitudinal creeping waves)
Health physics, 484–486
Heart, CAT scan, 451
Heat conduction
  equations, periodic solutions, 609–617
  one dimension, 607–609
  schematic, 606
Heat diffusion, 604–619, 630
  other extensions, 619
  techniques, 619–636
Heat exchangers (EC), 351–353
  flaws, 353
  tube bundle, NDE, 9
  tubes, multifrequency method, 341–343
Heating induction, NDE, 638–640
Heating methods, active thermography, NDE problem, 632
Heating pulse, 616
Heating source
  active thermography, 628–630
  phase shift, 612
Helium neon laser, 783
Helmholtz equation, 725
  shear and longitudinal potential, 175–176
Henry, Joseph, 264
Hertz, definition, 186
High-conductivity coating, 334
High explosive materials, process control, CT, 566–571
High fidelity sensors (AE), 399

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

High fidelity transducer, definition, 443
High speed digitizers, 374
High temperature alloys, liquid penetrant dwell time, 43
High temperature delay lines, piezoelectric transducer characterization, 126
Hit Lockout Time (AE), 404
Holodiagram, definition, 790
Holograms
  diffraction efficiency, 744
  equations, 754
  sensitivity vector, 752
  wavefronts, 740
Holographic contouring, 755–757
  illustration, 755
Holographic film, conventional film, 741
Holographic interference pattern, 741
Holographic interferometry, 744–756
Holography, 740–744
  definition, 790
  invention, 724
  wavefronts, 745
Homogeneous, definition, 186
Homogeneous isotropic medium longitudinal phasic velocity, 91
Honeycomb deposits, damage detection, 708
Hooke's law, linear elastic media, 91
Hounsfield, Godfrey, 457
Hsu-Nielsen source (AE), 371
  definition, 443
Hue, definition, 790
Human eye
  dark adaptation, 30
  frequency response, 29
  noise, 766
Human factors, NDE reliability, 13
Huygens-Fresnel wavelet principle (UT), 143–144
  beam divergence, 145
Huygens' principle, definition, 790
Huygens' wavelet model, 143
Hydrophilic
  definition, 60
  emulsifiers, 45
[Hydrophilic]
  postemulsified
    penetrant types and methods, 37
    PT flow charts, 34
  postemulsifying washes, penetrant removal process, 46
Hysteresis, definition, 256, 363
Hysteresis loop
  definition, 256
  demagnetization, 236
  ferromagnetic material, 206
    diagram, 205

Ideal testing environment, NDE reliability, 13
Illumination
  PT principle, 28–32
    type and intensity, 28
  PT technique, 49–50
Image, pupils, 739
Image blurring (radiography), 523
Image quality indicators (IQI), 527
  definition, 584
Image quality standards, 520
Image shear, variations in amplitude, 764
Imaging, 159
Imaging entrapped water, epoxy coating, 636–638
Imaging sensors, 677
Imaging systems, 736–740
  nomenclature, 759
  physical aperture stop, 737
Immersion mechanism, piezoelectric transducer characterization, 122–124
Impedance, 89–92
  definition, 186
Impedance analyzer, 323–325
  graphite-epoxy tubing, 324
Impedance method, definition, 363
Impedance plane analysis, tubes and rods, 301–306
  characteristic (frequency) parameter, 302–303
  fill factor, 302
  tube wall thickness, 303–306

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Impedance plane diagrams (EC), 291–301
  characteristic (length) parameter, 295
  definition, 363
  ferromagnetic conductive material, 296–301
  normalized impedance, 291–293
  test sample conductivity, 293–294
Impedance plane measurement (EC), 325–343
  cracks, 326–334
    surface-breaking, 326–330
Impulse response, definition, 737
Incident electric field, 696
Incident intensity, DR image, 540
Incident longitudinal wave (UT), Fresnel coefficients plot
  aluminum/water boundary, 105
  perspex/liquid/steel boundary, 107
  steel/air boundary, 103
  water/aluminum boundary, 104
Incident photon, orbital electron, 472
Incident shear wave (UT), planar boundary, critical angles, 102
Incident transverse wave (UT), Fresnel coefficients plot
  aluminum/water boundary, 106
  steel/air boundary, 103
Indirect piezoelectric effect, 120
Induced current, definition, 256
Induced strain, bulk waves, 80
Inductance
  definition, 363
  mutual
    definition, 364
    eddy current probes, 272–274
  self, 272
Induction rings, 228
Inductive coupling, 273
  coils, 273
Inductive impedance, eddy current, 280
Inductive reactance, definition, 363
Inductor, definition, 363
Industrial digital radiography, data acquisition geometry, 522
Industrial film radiographic imaging, 450
Industrial X-ray machines, 488
Infinite half-space, 610–612
Infrared cameras, 626–627
  temperature distribution, 621
Infrared detectors, 625–626
  responsivity, 625
Infrared radiometry, 622–628
  radiation laws, 622–625
Inherent defects, definition, 60
Inhomogeneity
  and attenuation, 109
  definition, 186
Inspection
  percentage, 8–12
  reason for, 9
Inspector influence, NDE reliability, 13
Intensification factor (radiography), 501
Intensifying screen (radiography), definition, 584
Intensity, 497
  definition, 790
Intensity ratios, normal incident wave example, 98–99
Intensity transformations, 159
Interface
  dielectric, electric magnetic fields, 697
  reflection and refraction, 92–107
Interference, 729–736
  definition, 187
  diffraction beam characterization, 141
  ultrasound, 170–171
Interference pattern, standing wave pattern, 668
Interference term, speckle term, 765
Interferogram, definition, 790
Interferometers, 396
  distance measurement, 735
Interferometric, NDE, 724
Interferometric measurement concept, 759
Interferometry, 744–756
  definition, 790
  laser, 797
  simulation of, 762
  wavefront, 747
Internal friction, absorption, 392

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Internal probe (coil), definition, 363
Internal Rotary Inspection System (IRIS) (UT), 350–351
International Annealed Copper Standard (IACS), 265
  definition, 363
International System of Units, 572
Interpretation, PT, 48–49
  Interpretation phase, NDE, 722
Intrinsic impedance, definition, 657
Inverse Fourier transformation, 615
Inverse square law, 468
InVision CTX system, 536–537
InVision Technologies, 535–536
Irrotational component, 89
Isotropic
  definition, 187
  naturally optically, 776
Isotropic homogeneous medium, normal incidence wave, 96
Isotropic materials
  bulk modes, 383–384
  lamb (plate) waves, 385–388
  Rayleigh (surface) waves, 384–385
  slowness surfaces, 180

Joule, 459–460
J-probes, 356

Kaiser effect (AE), 374, 375, 408, 409
  definition, 443
Kay-jay (MP), definition, 256
KCAT detector, schematic diagram, 545
KCAT DR/CT system, 543–544
  Kel-F, 569–570
Kelvins, definition, 790
Ketos test ring, 244–245
  definition, 256
KeV (kilo electron volt), definition, 584
Kevlar
  composite, 694
  porosity, 708
Kilo, 64
Knife-edge method
  definition, 187
[Knife-edge method]
  laser detection, 134–136
Known defect standard (KDS), 51
Kroneker delta function, bulk material, 87

LAD system geometry, 533
Lamb mode dispersion, 386
  vs. plate theory dispersion, 388
Lamb waves, 66, 117–120
  definition, 187
  dispersion curves, 119
  isotropic materials, 385–388
Lamé constants (UT), 83
  bulk material, 87
  definition, 187
Laminography, 552
  definition, 584
Laser (*see also* Optical)
  definition, 790
  detection
    methods, 134–137
    ultrasound, 133–138
  generation
    ultrasonic inspection principle, 149
    ultrasound, 133–138
  holography, 797
  interferometry, 797
    definition, 187
  optics, 723
  radars, 787
  source, 783
    continuous, 783
  speckle, origin, 758
  vendors, 784
Latent image (radiography), definition, 584
Laws of thermodynamics, PT fluid flow principle, 23
Lead, mass attenuation coefficient, 477–478
Lead metaniobate (PMN), piezoelectric material, 122
Lead zirconate titanate (PZT), piezoelectric material, 122

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Leakage field (MP), 206
  definition, 256
  diagram, 208
Leak testing, PT application, 57
Leaky surface waves, 116
Lenz's law, definition, 257
Lidars, 787
Liftoff (EC), definition, 363
Light
  basic properties, 725–729
  behavior, 721
  equations, 773
  equipment, 781–785
    vendors, 783
  fringe pattern, 781
  illustration, 774
  intensity, 781
  interaction, 721
  measurement techniques, 772–775
  properties, 721
  specific properties, 723
  structured, 772–775
  type of structured use, 773
Lignocellulose abrasives, PT technique, 37
Linacs, 494
Linear accelerator, definition, 584
Linear array, 507–508
  radiation detectors, 500
Linear-array detector (LAD)
  CT systems, 532–533, 562
    second-generation mode, 532–533
    third-generation mode, 533
  parallel-beam, 533
  source utilization, 521
  systems, 531–536
Linear array digital radiography system, with shielding system, 535
Linear attenuation coefficient, 473
Linear elastic media, Hooke's law, 91
Linear energy transfer (LET), 485
Linear laminography, movement arrangements, 554
Linear rigid body displacement, 83
Linear sensor arrays, location zone, 411
Linear source location, sensor spacing distances
  pipe, 413
Linear spatial frequency, 73
  definition, 186
Linear temporal frequency, 73
  definition, 186
Linear wave equation, one dimension, 76
Line pair test pattern, definition, 584
Lipophilic (PT)
  definition, 60
  PT flow charts, 34
Lipophilic emulsifiers, 44–45
  removing excess penetrant steps, 44
Lipophilic postemulsified, penetrant types and methods, 37
Lipophilic postemulsifying washes, penetrant removal process, 46
Lippmann, 67
Liquid, density, velocity, and specific acoustic impedance, 82
Liquid penetrant, 17–58
  dwell time, 41–43
  testing system, schematic and photograph, 33
Liquid petroleum gas (LPG), NDE, 8
Liquid/solid, surface tension, 25
Liquid/solid-solid/liquid interface, Fresnel equation, examples, 104–106
Lissajous figure, definition, 363
Little moon, PT fluid flow principle, 26–28
Local rigid rotation, 83
Location marker (radiography), definition, 584
Lock-in thermography, 615
Longitudinal, definition, 257
Longitudinal coherence, definition, 789
Longitudinal creeping waves, guided waves, 116
Longitudinal magnetization, 211
  vs. circular, 213–215
Longitudinal modes
  liquid/solid interface, 104–105
  one dimension, 76–78

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Longitudinal phasic velocity, homogeneous isotropic medium, 91
Longitudinal plane wave incident, planar boundary, critical angles, 102
Longitudinal potential
  Helmholtz equation, 175–176
  particle displacement, 173
Longitudinal pressure wave, 90
Longitudinal transmitted waves, water/aluminum boundary, echo plots, 108
Longitudinal velocity, 89
Longitudinal wave
  definition, 187
  generation
    EMAT, 131
    ultrasonic inspection principle, 149–150
  mode, 71
  piezoelectric transducer characterization, 122
  propagation, 89
Lorentz force, EMAT, 129–130
Loss factor, 655
Lost mechanisms, frequency dependency, 108
Love waves, guided waves, 116
Low-conductivity coating, 334
Low-energy gamma radiation, definition, 584
Luggage, inspection, 535–537

Macroscopic dipole movement, 653
Magnesium, liquid penetrant dwell time, 41–42
Magnesium casting, photograph, 55
Magnetic coupling, definition, 363
Magnetic dipoles, 200
Magnetic effect, definition, 257
Magnetic field, 197–200
  circular ring, 206
  concentration, 270
  current-carrying loop, 199
  density, definition, 257
  generation equipment, 222–230
[Magnetic field]
  induced current, 271–272
  materials, 200–202
  solenoid vs. bar magnet, 199
  straight wire current, 270
  strength, definition, 257, 363
Magnetic flux density, 271
  definition, 257, 364
Magnetic hysteresis, ferromagnetic materials, 203–207, 296, 298
Magnetic induction, 268
  definition, 257
  eddy current, 269–274
  EMAT, 129
  method, 228–229
Magnetic particle, 193–253
  application methods, 232–233
  concentration
    centrifuge tube, 238
    changes, 239
  contaminations, 239–241
  fluorescence loss, 239
  suspension, 240
    controlling, 237–241
  theory, 195–197
  types, 231–232
  viewing methods, 234
Magnetic particle inspection (MPI), 193
  advantages/disadvantages, 195, 196
  applications, 247–252
  avoiding nonrelevant indications, 245–247
  equipment, 207–215
  historical review, 194–195
  inspection aids, 237–241
  parts condition, 236
  pipeline weld, 14
  railroad wheels, 250–252
  techniques, 194, 207–215
  welds, 249–250
Magnetic penetrameter (MP), 241–242
Magnetic permeability, definition, 364
Magnetic saturation, definition, 296, 364
Magnetic susceptibility, definition, 257
Magnetic writing, definition, 257
Magnetism, theory, 195–197

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016



Magnetization
controlling, 241–245
current, types, 215–222
definition, 257
direction, 209–236
types, 202
methods, 210
Magnitude, AE, 380
Manipulators, 515
Marching cubes algorithm, 561
Marine industry, PT application, 56
Mass attenuation coefficient, 473
Maximum photographic contrast, 524
Maxwell, James Clark, 287–291
Maxwell's equations, skin depth, 287–291
Meander coil EMAT, 133
Measurement equipment, eddy current, 316–325
dynamic range vs. precision, 316–317
phase sensitivity, 317–320
resonant circuits, 321–322
response display, 320–321
Mechanical cleaning, PT technique, 36–37
Mediator (UT), definition, 187
Mega (M), 64
Meniscus, PT fluid flow principle, 26–28
Mersenne, Marin, 66
Metal pressure vessels, AE testing, 423–425
description, 423–424
results, 424–425
Metals
density, velocity, and specific acoustic impedance, 81
liquid penetrant dwell time, 43
Method of images, definition, 790
Method of potentials, ultrasound, 173–178
Michelson interferometry, 734–736
definition, 187, 790
method, laser detection, 136
Young's double slit configuration, 734–736
Micro-electro-mechanical systems, 627
Microfocus DR/CT systems, 542–543
Microfocus X-ray tube, definition, 584
Microseismology, 370
Microwave, 645–719
advantages, 649–650
antennas, 666
circulators, 668
detectability, 708
detectors, 669–670
dielectric characterization, 690
directional couplers, 667–668
disadvantages, 649–650
equipment, 664–672
heating, carbon fiber contaminants, 638
heating methods, 636
inspection, 646, 649
with probes, 703–711
isolator, 669
matched loads, 667
material parameters, 651–656
skin depth, 651–652
mixers, 670–672
moisture permeation detection, 707
NDE, 645, 647
applications, 678–713
conductors, 652
definition, 712
dielectric properties, 656
electromagnetic wave, 651
hardware systems, 650
historical perspective, 647
measurements, 646
modern era, 648
near-field techniques, 704
potential applications, 648–649
properties of, 656
sensors, 713
techniques, 695
wavelength, 648
near-field imaging
examples, 707
experimental results, 710
near-field inspection, 711
network analyzers, 665
noncontacting, 673–674
nondestructive evaluation, 713
oscillators, 664–665

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

[Microwave]
voltage controlled, 664
Yttrium Iron Garnet, 664
potential techniques, 647–649
problems, 713–715
radar sensors, 673–674
radiometer sensors, 675
reflection, 673–674
experimental setup, 699
resonator sensors, 661
sensors, 672–678
resonators, 674
versatility, 678
symbols, 715–716
technical overview, 646
techniques, 672–678
thermography, 677
transmission method, 692
transmission sensors, 672
waveguides, 665–666
wavelengths, 647
MIL-6866B, 52
MIL-I-25135E, 52
Mil-Std-6866, 51
MIL-STD-410A, 52
Mixers (microwave)
diode detector, 670
microwave, 670–672
nonlinear solid state device, 670
Modal analysis, AE, 420–422
Mode conversion, 92
acoustic waves, 394–395
definition, 187
Modulation transfer function (MTF), 518
Moiré, definition, 791
Moiré fringes, production of, 778
Moiré pattern, 774–775
Moisture permeation detection, microwave, 707
Monte Carlo method (radiography), 484
Morgan William, 451–452
Morphology, definition, 791
Mounts, 784
vendors, 785
M-scan (UT), definition, 187
Mulhauser, 67
Multidirectional magnetization (MP), 215–216
definition, 257
Multifrequency-multiparameter EC method, 337–338
Multiple attenuation measurements, 482
Multiple element, piezoelectric transducer characterization, 124
Multiple loading, 408–409
Mutual inductance
definition, 364
eddy current probes, 272–274

Narrowband, piezoelectric transducer characterization, 125
Narrow beam (radiography), 473
NAVSEA 250-1500-1, 52
Near field (UT)
definition, 187
transducer characteristic, 143–145
Near field imaging
experimental results, microwave, 710
rust, 711
Near field length, definition, 187
Near field region, electromagnetic field properties, 706
Net density, definition, 584
Network/impedance analyzer, 323–325
graphite-epoxy tubing, 324
Neutron attenuation, 479
Neutron radiography (NRT), definition, 584
Newton, Isaac, 66
Newton's fringes, definition, 791
Newton's law of motion, PT fluid flow principle, 23
Noise
definition, 364, 585
discrimination, AE, 418–419
equivalent temperature, NEP, 626
eye, 766
radiation, 676
sources, 378–379
Nonaqueous suspendable developers (*see* Solvent suspendable)

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Nonconductive coating, 334
Nondestructive characterization (NDC), implies, 3
Nondestructive evaluation (NDE)
  active thermography, 597
  application, 3–5
  basic levels of choosing, 7–8
  definition, 1–2, 791
  diagnostic maintenance tool, 5
  implies, 3
  inspection methods, 614
  methods, ultrasonic, 63–180
  microwave, 713
  optical techniques, 721
  principle, 2
  program statistics, 8–9
  purpose, 8
  reliability, 12–14
  testing procedure and schedule, 12
Nondestructive examination, definition, 3
Nondestructive inspection (NDI), definition, 3
Nondestructive sensing (NDS), definition, 3
Nondestructive testing (NDT), definition, 2
Non-energy discriminating (current integrating) detector systems (radiography), 530–547
Nonferromagnetic conductive materials, EC, 275–278
Nonimaging techniques, 787
Nonmetals, density, velocity, and specific acoustic impedance, 81–82
Nonoptical heating sources, wavelength, 631
Non-safety-critical parts, NDE, 9
Normal beam, piezoelectric transducer characterization, 124
Normal incidence (UT), boundaries, 94–107
  amplitude ratios, 96
  critical angles, 101–102
  echo transmission, 106–107
  Fresnel equations, 102–106
  intensity ratios, 98–99
[Normal incidence (UT), boundaries]
  oblique incidence, 99–100
  reflection and transmission, 94
  reflection coefficient, 95
  Snell's law, 100–101
Normal incident, ultrasonic inspection principle, 156
Normalized impedance (EC), impedance plane diagrams, 291–293
Normal strain, 83, 86
Normal stress, Poisson effects, 86
North/south poles (magnetic), definition, 257
Nuclear activity, definition, 585
Nuclear decay, 465
Nucleus, radioactivity, 465–466
Nucleus decay, 495
Null balance (EC), definition, 364
Null-location curve, 684
Numerical control machining, 560
Nyquist criterion, 517–518

Objective speckle pattern, 759
Objective speckle schemes, 787
Object leg, laser detection, 136
Oblique incidence (UT)
  Fresnel equation, examples, 102–106
  normal incident wave, 99–100
Observation plane
  intensity, 732
  Young's double slit experiment, 732–734
Occupational exposure, radiographers, 573
Offset null balance, EC measurement optimization, 336–337
Ohm's law, 288
  definition, 364
Oil-and-whiting technique, PT procedure, 19
Oil-based, penetrant types and methods, 37
One-dimensional arrays, location zone, 411
Operating point (EC), definition, 364

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Optical, definition, 723 (*see also*, Laser)
Optical absorption properties, 629
Optical beam deflection, 601–602
  NDE measurements, 601
Optical density, definition, 585
Optical detectors, electromagnetic field amplitude, 729
Optical generation and detection, ultrasound, 133–138
Optical interference, 137–138
Optical methods, 721–806
  data processing, 765–771
  historical methods, 723–724
  postscript programming, 804
  problems, 787
  symbols, 792–793
Optical remote sensing, NDE,723
Optical techniques, 744–757
  nondestructive evaluation, 721
Optical transducer (UT), definition, 187
Optic mounts, 784
Optics, 784
  laser, 723
  vendors, 785
Orbital electron, incident photon, 472
Orientation, AE, 381–383
Orsted, Hans Christian, 264
Orthoscopic image, definition, 791
Oscillator
  definition, 364
  Doppler frequency shift, 671
  isolator, 671
Oscilloscope, definition, 187
Oscilloscope display
  signals, 159
  ultrasonic pulser-receiver system, 158
Out-of-plane deformation, 760–761

Pair production, 462, 472
  definition, 585
Parallel-beam area-array system (radiography), 541
Parallel-beam geometry, 521
Parallel-beam LAD, 533
Parallel resonant circuit, 662
Paramagnetic magnetization, 202
  definition, 257
Paraxial approximation, definition, 791
Part failure, NDE, 9
Particle, 479
  definition, 580
Particle displacement, 173–178
  longitudinal potential, 173
  plate waves, 176–178
  SAW, 175–176
  Shear potential, 173–175
Particle DR/CT system, 547–548
Particle image velocimetry, definition, 791
Particle imaging velocity, 786
Particle size (UT), wave scattering, 111
Particle stream, vs. wave response (diffraction), 142
Particle velocity (UT), 72
  continuity, 178–180
  definition, 189
PBX9502 pellet, 568
PCAT DR/CT system, 538–539, 541–542
Pencil beam (radiography), definition, 585
Pencil-Beam Computed Axial Tomography (PBCAT), 568–569
Pencil lead break
  AE sensor, 372
    forcing function, 373
  definition, 443
Penetrameters, 527
Penetrants
  color-contrast
    PT technique, 38–39
  comparator block, 52–53
  definition, 60
  dyes, PT illumination principle, 29
  fluorescent, 37–38
  inspection, photograph, 55
  liquid, 17–58
  removal process, lipophilic and hydrophilic postemulsifying washes, 46
  removing excess, 40–45

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

[Penetrants]
types, 19, 37–39
Penetrant testing (PT), 17–58
advantages/disadvantages, 21, 22
applications, 52–58
aerospace industries, 53–54
automotive and marine manufacturing, 56
electrical power industries, 56–58
fabrication industries, 52–53
petrochemical plants, 54
examination, 48–49
fundamentals, 21–31
history, 19–20
illustrated, 6
interpretation, 48–49
method, 31
photograph, 33
potential, 20
schematics, 32, 33
techniques, 18–19, 31–52
cleaning, 32–37
dwell time, 39–40
examination and interpretation, 48–49
final cleaning, 50
illumination, 49–50
removing excess penetrant, 40–45
specifications and standards, 50–52
temperature, 39
types of developers, 45–48
types of penetrants, 37–39
using requirements, 17
Penetration (EC), depth of, definition, 362
Penetration time, PT technique, 39–40 (*see also* Dwell time)
Penumbra, 493
definition, 585
Period
definition, 187
wave propagation, 72
Periodic heating, experiments, 615
Permeability
composite material, 661
definition, 257, 364
effective, definition, 362
Permittivity
absolute, 655
composite material, 661
Perspex, 106
Petrochemical plants, PT application, 54
Phase
ultrasonic inspection principle, 148
ultrasound, 171
Phased array
definition, 187
piezoelectric transducer characterization, 124
Phase lag, definition, 364
Phase measurement, EC, 319
Phase retarder, definition, 792
Phase rotation, EC, 319
Phase sensitivity, EC, 317–320
Phase shift, heating source, 612
Phase unwrapping, 771–772
definition, 791
Phase velocity, 72, 74
definition, 189
ultrasound, 171–173
Phasor
definition, 364
notation, eddy current, 280–285
sinusoids, 281
Photoacoustic effect, 600–601
Photoacoustic spectroscopy, 601
Photodisintegration, 462, 473
definition, 585
Photoelastic techniques, 775–781
definition, 791
Photoelectric absorption, 471
Photoelectric effect, 462
definition, 585
Photomultiplier tubes, 507
radiation detectors, 500
Photon, 459–462
attenuation, 474
definition, 791
detectors, thermal detectors, 626
energy lines, 467
radiation, 466–479
interaction, 469–470
material interaction, 469

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

[Photon]
NDE, 466–479
physics, 466–473
sources, 466–473
sources
NDE, 468
Photophone, 600
Photopic, PT illumination principle, 28
Photostimulable luminescence, definition, 585
Photostimulable (storage) phosphors, 509
radiation detectors, 500
Photothermal characterization, different classes, 599
Photothermal radiometry, 602
Physical aperture stop, imaging systems, 737
Pickup coils (EC), 262
Pie field indicator (MP), 243
Pie gauge, 241–242
Piezoelectric, definition, 187
Piezoelectric AE transducers, 397–398
Piezoelectric sensors, 396
Piezoelectric transducers, 120–128
characterization, 122–128
construction, 127–128
longitudinal wave generation, ultrasonic inspection principle, 149
Pipe, linear source location, sensor spacing distances, 413
Pipeline weld, magnetic particle inspection, 14
Pitch-catch (UT)
definition, 188
illustration, 68
ultrasonic inspection, 65–66
principle, 154–155
Pixel, 506
definition, 585
size, definition, 585
P-jay (MP), definition, 257
Planar interface
reflection, 658–659
transmission, 658–659
Planar sensor arrays, location zone, 411
Planar source location
graphics, 415–416
intersecting hyperbolas, 416
sensor array geometry, 415
Planar wavefront, 748
Plane waves
characteristic impedance, 656–658
definition, 188
types or modes, 71
Plank's constant
Boltzmann's constant, 798
definition, 791
Plastics, liquid penetrant dwell time, 43
Plate theory dispersion, vs. Lamb mode dispersion, 388
Plate waves (*see also* Lamb waves)
dispersion curves, 174
generation, 133
particle displacement, 176–178
Point-detector based systems (radiography), source utilization, 521
Point detector systems (radiography), 530–531
Point force, illustration, 382
Point location
AE, 411–418
equations, 413
Point-spread function, 737–740
definition, 791
Poisson effects
normal stress-strain, 86
shear stress-strain, 86
Poisson's ratios
bulk waves, 80
Rayleigh waves, 115
Polar diagrams, definition, 188
Polariscope, 779
definition, 791
illustration, 779
stress field, 782
wave components, 780
Polarity (MP), definition, 257
Polarization
definition, 653, 791
macroscopic point of view, 654
mechanisms of, 653

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Polarizer
  definition, 791
  equation, 779–781
Polaroid, definition, 791
Poled, piezoelectric transducers, 122
Polymers
  composites, porosity evaluation, 690
  porosity levels, 690
Porosity
  differences, 691
  permittivity, 691
Positive images, 519
Postemulsified penetrant wash, 44–45
Potential energy, PT fluid flow principle, 23
Power ratios, attenuation, 113
Precision, vs. dynamic range, eddy current, 316–317
Precleaning, PT technique, 32–37
Pressure ratios (UT), normal incident wave example, 98–99
Pressure waves, 71
  definition, 188
Primary field (EC), definition, 364
Primary radiation, definition, 585
Prods, 223
  definition, 257
Projection angles, 519
Projection radiography, 449
Protection standards, radiation, 486
Protons, 479
Pseudoscopic image, definition, 791
Pulsed excitation, 615–617
Pulsed heating, 632
  method, characteristic feature, 616
  surface temperatures, 618
Pulsed laser sources, 783
Pulse-echo (UT)
  definition, 188
  illustration, 68
  ultrasonic inspection, 65–66
    principle, 153–155
Pulse generator (UT), definition, 188
Pulse-height spectroscopy (radiography), 512
Pulsers (UT), 156–158
  options and limits, 157
Pupils
  definition, 736
  image, 739
Pure absorption (UT) and attenuation, 111
Purkinje shift, PT illumination principle, 28
PVDF, piezoelectric material, 122
P-waves (UT), 71
  definition, 188

Quadrature (EC)
  defined, 320
  sensor response voltage, 319
Quality control
  large expensive systems, 10–11
  safety-critical parts, 10–11
Quality factor, 485
Quantitative Quality Indicators (QQIs), 242–244
  definition, 257
Quartz, piezoelectric material, 122
Quick-break circuitry (MP), 229–230
  definition, 257

Rad, definition, 585
Radar guns, nondestructive inspection, 645
Radiating near-field region, 704
Radiation, 745
  electromagnetic theory, 650
  fundamental sources, 464
  how it travels, 461
  protection standards, 486
  sources, 464–466
Radiation accidents, reduction, 576–577
Radiation detectors, 498–515
Radiation dose, 485
  vs. health effect, 572
Radiation effects, 572–573
Radiation exposure, reduction, 574–576
Radiation fundamentals, 459

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Radiation gauge
  schematic, 449
  set up, 530
Radiation laws, infrared radiometry, 622–625
Radiation methods, 458
  safety concerns, 458
Radiation NDE, applications, 461
Radiation safety, 571–578
  procedures, 574–575
  rationale, 571–574
  responsibilities, 577–578
Radiation Safety Officer, 576
Radiation sources, 487–498
  electrically powered, 487–494
Radiation testing, equipment, 486–516
Radiation transport modeling, 484
Radioactive decay, 466
Radioactive half-life, 495
Radioactivity, 464–465
  electrons, 464
  nucleus, 465–466
Radiograph, definition, 585
Radiographers
  occupational exposure, 573
  safety responsibilities, 577
Radiographic contrast, 504, 523
  definition, 585
Radiographic dynamic range, 519
Radiographic energy, NDE, 459
Radiographic equivalence factor, definition, 585
Radiographic exposure, definition, 582
Radiographic techniques, 447
  CT, 528–529
  digital radiography, 528–529
  film, 525–528
  measurement limitations, 520–525
  practical considerations, 519–520
  special, 550–554
Radiography, 449, 463, 480
  accuracy, 456
  definition, 463
  NDE methods, 604
  regulations, 573–574
  safety, 456
[Radiography]
  two dimensional, 447
Radioisotopes, 466–467
  sources, 494–498
    characteristics, 496
    disadvantages, 495
Radiological systems, 515–516
Radiology, 447–595
  CT, 481–484
  fundamentals, 459–468
  radiation detectors, 498–515
  radiation safety, 571–578
  radiation testing equipment, 486–516
  radiographic techniques, 516–554
  radiography, 480
  selected applications, 554–571
  X-rays, 448–451
Radiometers, 676
Radiometer sensors, 675–677
  microwave, 675
Radon transform (radiography), 483
  new inverting methods, 483
Rail cars, inspection (radiography), 535
Railroad wheels
  MPI, 250–252
  roll-by inspection (UT), 161–167
Raster scan, 548
Rayleigh, 66
Rayleigh criterion, definition, 791
Rayleigh range, definition, 791
Rayleigh scattering, 462, 470
  definition, 188
  particle size, 111
Rayleigh speed, SAW generation, 151–152
Rayleigh (surface) saves, isotropic materials, 384–385
Rayleigh velocity, calculation, 384–385
Rayleigh waves, 115
Real time DR/CT system, 544, 547
Real time DR imaging system, photograph, 546
Real time radioscopy, definition, 585
Red dye, PT technique, 38
Reference coils
  definition, 365

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

[Reference coils]
EC probes, 311–312
Reference leg, laser detection, 136
Reflectance (UT), definition, 188
Reflected acoustic pressures, different boundary materials, examples, 96–99
Reflected electric field, 696
Reflected longitudinal wave, 100–101
Reflected shear wave, 100–101
Reflection, 92, 418
acoustic waves, 394–395
coefficient, 659
definition, 188
normal incidence wave, 95
coefficient phase difference, 700
normal incidence wave, 94
sensors, application, 674
Refracted longitudinal wave, 100–101
Refracted shear wave, 100–101
Refracting transducers, piezoelectric transducer characterization, 124
Refraction, 92
acoustic waves, 394–395
definition, 791
wave, 93
Refractive index, definition, 792
Regulations, radiography, 573–574
Relative angular displacement, bulk material, 84
Relative displacement, bulk material, 83
Relative penetration sensitivity, 217
Relative permeability, definition, 365
Relative permittivity, 655
Reluctance, definition, 364
Remote field eddy current (RFEC) (*see* Remote field testing (RFT))
Remote field testing (RFT), 343–349, 350–354
advantages, 350–351, 358
basics, 343–345
boiler damage, 352
challenges, 351
disadvantages, 358
discontinuity signatures, 348–349
EC calibration, 349–350
[Remote field testing (RFT)]
field methods, 353–354
heat exchanger
damage, 351–352
tube, 351
probes, 352
response visualization and interpretation, 345–349
schematic, 344
vector polar plot, 346–348
Residual magnetism, definition, 257
Residual method, definition, 257
Resin binders, 690
Resistance
definition, 365
physical movement, 89–92
Resistive impedance, 272
eddy current, 279–280
Resistivity, 358
definition, 365
Resolution, definition, 792
Resonance, 661–664
definition, 365
Resonant circuits, eddy current, 321–322
Resonant transducer, definition, 443
Resonator
capacitance, 663
microwave sensors, 674
quality factor, frequency deviation, 662
sensors, 674–675
waveguides, 663
Responsivity, IR detectors, 625
Retentivity (MP), definition, 257
Retirement-for-cause NDE, 11
Reverse engineering, CT, 559–562
Reverse geometry, 552
X-ray imaging, 554
Right hand rule, 198
definition, 258
Rigid body rotation, 83
Risk-based determination, NDE testing procedure and schedule, 12
Risk-based matrix, ranking components, inspection needs, 12
Risk-informed inspection, 11–12

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

RMS (root-mean-squared) voltmeter, 401–402, 406
Rod, PT illumination principle, 31
Rod response, PT illumination principle, 28
Roll-by inspection (UT), railroad wheels, 161–167
Ronchi grating, transmission profile, 775
Ronchi ruling, 774
  definition, 792
Röntgen, 572
  definition, 585
Röntgen equivalent in mammals, 572
Röntgen equivalent man, definition, 586
Röntgen rays, 453
Röntgen, Wilhelm, 452
  cathode ray tube, 452–453
Rotating X-ray target, 491
Rubber compound
  anisotropy, 686
  carbon black, 688
  constituents, 684–685
  curative detection, 687–688
  cured, 685–686
  dielectric characterization, 684
  volume content, 685
Rust, 709–711
  near-field imaging, 711

Safety critical, 6
Safety-critical parts, quality control, 10–11
Sampling frequency, AE, 405
Saturation
  definition, 365, 792
  transducer characteristic, 139
Saturation point, definition, 258
Scalar dot product, 751
Scanners, 626
Scanning artifacts, 540
Scanning photoacoustic microscopy, 601
Scanning tank (UT), 123
Scattered waves (UT), normal incident wave, 99–100
Scattering, 462, 470
  and attenuation, 107, 109
  matrix, 660
  parameters, 659–660
  particle size, 111
  physical and mathematical representation, 112–114
  ultrasonic waves, 65
Schlieren techniques, 786
  definition, 792
Scholte waves, 115–116
Scintillators
  definition, 586
  radiation detectors, 501–502
Scotopic, PT illumination principle, 28
Scratch detection, leakage field, diagram, 208
Screens (radiography)
  definition, 586
  radiation detectors, 500–501
    fluorescent, 501–502
Sea containers, inspection, 535
Seat belt buckles, MPI inspection, 232
Secondary field (EC), definition, 365
Secondary radiation, definition, 586
Second critical angle (UT), definition, 188
Seismic monitoring, 370
Seismic waves, 370
  sources, 370
Self-inductance, 272
Self-reference coil (EC), 311–312
Semiconductors, 512–513
Senkhrect, definition, 792
Sensitivity (radiography), 506
Sensitivity vector, 749
  definition, 792
  hologram, 752
  illustration, 751
Sensor response voltage, quadrature, 319
Septa collimators, 515
Shadow moiré approach, 775
Shaped transducer (UT)
  illustrated, 127
  piezoelectric transducer characterization, 124–125

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Shear, NDE concept, 765
Shear constant, bulk material, 87
Shear horizontal (UT), normal incident wave, 99–100
Shearing interferometry, 763–771
  definition, 790
Shearographic images, 766
Shearography, 763
  definition, 792
  spatial phase stepping, 768
  temporal phase stepping, 766
Shear potential (UT)
  Helmholtz equation, 175–176
  particle displacement, 173–175
Shear strain, Poisson effects, 86
Shear stress, 78
  Poisson effects, 86
Shear vertical (UT), normal incident wave, 99–100
Shear wave (UT), 71
  definition, 188
  generation, ultrasonic inspection principle, 150–151
  piezoelectric transducer characterization, 122
  propagation illustrated, 77
Shelby countup ration (SCR) (AE), 429
  burst strength, 430
Shielding magnetic fields, EC probes, 312–314
Shoe fitting, X-rays, 455
Shot noise, definition, 792
Shutters, 784
Sidebands, 769
Side lobes (UT)
  beam divergence, 145
  definition, 188
Sievert, 485, 572
Signal conditioning, AE, 400–401
Signal receiving, ultrasonic inspection principles, 158–161
Signal-to-noise ratio, definition, 365
Simmons, 68
Simple shear, 78
  definition, 188
Simulated AE signals, large vs. small specimen, 396
Single-detector based systems (radiography), source utilization, 521
Single-detector DR/CT system, 548–549
Single element, piezoelectric transducer characterization, 124
Single-loop coil (EC), Faraday's law, 314–315
Single-phase-full-wave direct current (MP), 219–220
Single photon counting system, 529
Sinogram, 483
Sinusoids, phasors, 281
Skew frequency (UT)
  definition, 188
  transducer characteristic, 140
Skin depth (EC)
  definition, 258, 365
  eddy current density, 285–287
  Maxwell's equations, 287–291
Skin effect, definition, 365
Sloppy continuous method (MP)
  definition, 258
Slowness curves, 178–180
  definition, 188
  ultrasound, 178–180
Slowness surfaces, isotropic and anisotropic materials, 180
Slow-responding light receptors, PT illumination principle, 31
Snell's law, 93, 395
  definition, 792
  derivation, 178–180
  description, 100–101
  direction determination, 102
  longitudinal wave generation, ultrasonic inspection principle, 149
  normal incident wave, 99–100
SNT-TC-1A, 52
Society of Automotive Engineers (SAE), 52
Soft magnet, definition, 258
Sokolov, 67
Solenoid, definition, 258

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Solenoid coil (MP), 212
Solenoid vs. bar magnet, magnetic field, 199
Solid/free boundary interface, Fresnel equation, examples, 102–106
Solid/gas (PT), surface tension, 25
Solid/thin liquid layer/solid interface, Fresnel equation, example, 106
Solvent-aided penetrant removal, 45
Solvent cleaning, PT technique, 35
Solvent-removable, PT flow charts, 34
Solvent suspendable, developer, PT technique, 46
Solvent wipe, penetrant types and methods, 37
Sound, definition, 64
Sound field diffraction pattern, 142
Sound field variation, 143
Sound waves, physics of, features, 75
Source, 377–383
  characterization
    AE, 418–419
    waveform analysis, 420
  definition, 586
  detector configurations, types, 521
  event duration, 418–419
  inversion, waveform analysis, 420
  location
    accuracy, 414–418
    waveform analysis, 410–420
  unsharpness, 491–492
  utilization, 521
South poles, definition, 257
Spatial coherence, definition, 789
Spatial distortions, area-array detectors, 538–539
Spatial filters, 784
Spatial frequency, 73, 74
  definition, 790
Spatial heating patterns, active thermography, 629–630
Spatial phase stepping, shearography, 768
Spatial resolution, 506–507, 517
  thermal transit times, 637
Specific acoustic impedance, 89–92
Specific acoustic pressure, 89–92
Specific activity, 495, 497
Specific heat, 604–607
Speckle
  construction, 801
  definition, 758, 792
  effects, 803
  pattern shearing, illustration, 764
  phase terms, 760
  phenomenon, 757–761, 798–803
    statistical treatment, 800
  size, 801
  technique, 757–772
    potential, 757
  term, interference term, 765
Spectral radiant emittance, 622
Spectroscopy DR/CT based systems, 548–550
Spectroscopy DR/CT detectors, 524
Spherical wave, 729
Spherical wavelets, 143
Spike pulsers (UT), 157
Split coil (MP), definition, 258
Spraying, PT technique, 39
Square wave amplitude transmissions, 774
Standing wave, 657
  minimum, 681
  pattern, 658
  ratio, 658, 680
Steel
  exposure chart, 526
  liquid penetrant dwell time, 42
Steel/air boundary interface (UT), Fresnel equation, examples, 103
Steel coil springs (MP), 247–249
  fatigue cracks, 248
  seams, 248
  stress cracks, 247
  surface breaking cracks
    flaw response trajectories, 326
Steel container (radiography), stepped weld joint, 557
Steel/water interface (UT), reflected and transmitted acoustic pressures, 96–99

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Stefan-Boltzmann's law, 623
Step heating, 617–619, 632
Step heating cases, data analysis algorithms, 633
Step-hole-type IQIs (radiography), 527
Stepped weld joint (radiography), steel container, 557
Stereo imaging, 552
Stereo radiography, 553
Stoneley waves, guided waves, 116
Storage phosphors, definition, 586 (*see also* Photostimulable (storage) phosphors)
Storage ring, 468
Strain energy, 380
Strain gages, 396
Strength predictions, AE, 376–377
Stress-corrosion cracks, definition, 60
Stress field, polariscope, 782
Structured light, definition, 792
Strutt, John William, 66
Subsurface defect (MP), leakage field, diagram, 208
Supersonic reflectoscope, 68
Surface acoustic wave (SAW), 114–116
  generation, ultrasonic inspection principle, 151–152
  particle displacement, 175–176
Surface breaking cracks (EC), flaw response trajectories, 326
Surface crack (MP), leakage field, diagram, 208
Surface deformations, 747–749
Surface discontinuities, 18
Surface energy, PT fluid flow principle, 23
Surface probe (EC), definition, 365
Surface tension
  definition, 60
  PT fluid flow principle, 22–24
Surface wave, piezoelectric transducer characterization, 122
Surface wetting, PT fluid flow principle, 24–26
Swing fields (MP), definition, 258
Synchrotrons, 468, 494
Tangential electric fields, dielectric interface, 697
Target (radiography), definition, 586
TATB weight fraction, 569–570
Teeth, digitized film radiograph, 450
Telecentric, definition, 792
Temperature, PT technique, 39
Temperature distribution
  bare rebar, 639
  Fourier's law, 607–609
Temporal coherence, 795
  definition, 789
Temporal phase stepping, shearography, 766
Temporal response, AE, 380
Test coil (EC), definition, 365
Thermal conductivity, 605–607
  Fourier's law, 606
Thermal detectors, photon detectors, 626
Thermal diffusivity, 608
Thermal effusivity, 611
Thermally conductive metals, 635
Thermally-insulating materials, 635
Thermal methods, NDE problem, 631
Thermal mismatch factor, 614, 635
  negative, 614
Thermal neutrons, 479
  definition, 586
Thermal transit times, 629
  spatial resolution, 637
Thermoelastic, laser generation, 134
Thermoelectric noise, 540
Thermography (*see* Active thermography)
Thickness, developer coating, 46
Thick-walled tubing (EC), 302–305
Thin film transistors (TFT), radiation detectors, 500
Thin lens imaging rule, 737
Thin-walled tubing (EC), 302–305
Threader bar, definition, 258
Three-dimensional wave (UT), 71
  equation, 88
Three-phase-full-wave direct current (MP), 220
Threshold, transducer characteristic, 139

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Through transmission (UT), definition, 188
Time, radiation exposure reduction, 574–575
Time averaging operation, 731
Time of flight
  definition, 188
  ultrasonic inspection principle, 147–148
Tin cry, 374
Titanium, liquid penetrant dwell time, 43
TMC study
  composites, 429–435
    results, 432–435
    test, 431–432
Tomography, definition, 586
Tomosynthesis, 553–554
Tone burst generator (UT), definition, 188
Tone burst pulsers (UT), 157
Toroidal field (MP), definition, 258
Torsional velocity (UT), 89
Torsional wave, propagation, 89
Total image unsharpness (radiography), definition, 586
Total internal reflection (UT)
  definition, 188
  normal incident wave, 101–102
Tourmaline, piezoelectric material, 122
Traditional film (projection) radiography, 555–558
  weld inspection, 555–556
Transducer, 395–409
  bandwidth, definition, 140
  configurations, ultrasonic inspection principles, 153–156
  coupling medium, 128
  definition, 188, 443
  flat, piezoelectric transducer characterization, 124–125
  frequency response curves, 398
  ultrasonic inspection, 65–66
  ultrasonic wave generation and detection, 120–147
    characteristics, 138–147
    electromagnetic acoustic, 129–133
    laser, 133–138
    piezoelectric, 120–128
Transformer, ferromagnetic core, 278
Transmission, 92
  coefficient, 659
  definition, 188
  densitometer
    definition, 586
  normal incidence wave, 94
  radiographic image, 540
  sensors
    advantages, 672
    microwave, 672
Transmittance (UT), definition, 188
Transmitted acoustic pressures, different boundary materials, examples, 96–99
Transmitted film density (radiography) definition, 586
Transmitted longitudinal wave, liquid/solid interface, 104–105
Transuranic, 465
  nuclides, 465
Transverse coherence, definition, 789
Transverse electric, definition, 792
Transverse electric-field vector, 657
Transverse electromagnetic waves, 656, 704
Transverse magnetic, definition, 792
Transverse matrix cracking, 418
  location results, 435
  source, 433
Transverse mode, one dimensional, 78–80
Transverse transmitted waves, water/aluminum boundary, echo plots, 108
Transverse velocity, 89
Transverse wave (UT), 71 (*see also* Shear wave)
  generation, EMAT, 132
  propagation, 89
Traveling harmonic wave, wave function, 75–76
Triggers, AE, 405
Trucks, inspection, 535
Tube
  Bobbin probe (EC)

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

[Tube]
normalized impedanced plane, 306
current (radiography)
definition, 586
encircling coil (EC)
normalized impedance plane, 305
impedance plane analysis (EC), 301–306
trailer
AE requalification test, 423
Turbine disk
EC inspection, 357
Two-dimensional arrays
location zone, 411
planar location zones, 412
Two-dimensional waves, 71
Two projection radiographs
results, 481

Ultrasonic, NDE methods, 604
Ultrasonically hard, 91
definition, 189
Ultrasonically soft, 91
definition, 189
Ultrasonic cleaning, PT technique, 36
Ultrasonic free boundary, 102
definition, 189
Ultrasonic inspection
principles, 147–161
excitation pulsers, 156–158
measurements, 147–149
signal receiving, 158–161
transducer configurations, 153–156
wave generation, 149–153
technique, 65–66
Ultrasonic longitudinal wave, propagation illustrated, 77
Ultrasonic NDE (*see* Ultrasound)
Ultrasonic pulser-receiver system, oscilloscope display, 158
Ultrasonic reflection, circular defect, 65
Ultrasonic testing (UT), 63–64
axles, 167–170
elevator shafts, 167–170
Ultrasonic transverse wave, propagation illustrated, 79
Ultrasonic waves
amplitude and attenuation, 112
description, 64–65
propagation, 64–65
Ultrasound, 63–180
advantages and disadvantages, 69, 70
applications, 68–69, 161–170
group velocity, 171–173
history, 66–68
inspection principles, 147–161
interference, 170–171
method of potentials, 173–178
phase velocity, 171–173
slowness curves, 178–180
Snell's law derivation, 178–180
theory, 69–120
transducers, generation and detection, 120–147
Ultraviolet (UV) lamp, definition, 60
Uniform field eddy current (UFEC) probe, 314
Uniform fringe spacing, 766
Uniformity (PT), developer coating, 46
Unprocessed film, storage conditions, 527–528
Unwrapping algorithms, 771
application, 772
flowchart, 772
U.S. Atomic Energy Commission (USAEC), 573–574
U.S. Nuclear Regulatory Commission
Agreement States, 573–574

Vacuum permittivity, 654
Vapor degreasing, PT technique, 35–36
Vector, definition, 365
Vector electric-dipole movement, 654
Vector polar plot, RFE 9, 346–347
Velocity (UT)
definition, 189
ultrasonic waves, 64
Vertical resolution, 405–406
Vibration, 378–379
isolation, 782–783

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Video holography, electronic speckle pattern interferometry, 762
Virtual image, definition, 792
Viscosity (PT), definition, 60
Visible dye, PT technique, 38
Visible penetrants, 37
Visible spectrum
  colors
    frequency and wave lengths, 29
  illumination
    PT principle, 28–32
Visual NDE, 1
Voltage
  definition, 365
  EC probe, 279
Voltmeter
  definition, 365
  RMS, 401–402, 406

Water/aluminum boundary (UT)
  attenuation, echo plots, 108
  Fresnel equation, examples, 104
Water-based, penetrant types and methods, 37
Waterfall plot, 159
Water-soluble, developer, PT technique, 46
Water/steel interface (UT), reflected and transmitted acoustic pressures, 96–99
Water suspendable, developer, PT technique, 46
Water-washable
  penetrants
    removing excess, 40
    types and methods, 37
  PT flow charts, 34
Wave
  components, polariscope, 780
  description, 70, 75
  digitizing AE systems, disadvantages, 404–405
  energy or momentum, 70
  equation, 74–89
    solution, 726
  generation, ultrasonic inspection principle, 149–153
  motion, 74–89
    governing equation, 75
  one dimension, 75–80
    longitudinal mode, 75–78
    transverse mode, 78–80
  propagation, 693
    AE, 383–395
      velocity and dispersion, 383–394
    complete cycle, 72
    illustration, 726
    ultrasonic theory, 70–74
  recording instrumentation, 374–375
  refraction, 93
  split, 93
  three dimensions, 80–89
    displacement/deformation, 83–89
  types of, 100–101
  vector, definition, 729 (*see also* Angular spatial frequency)
Waveform-based systems, AE, 405
Waveform digitization, AE, 403–406
Wavefront, 71
  definition, 189, 792
  example of use, 746
  holograms, 740
  holography, 745
  splitting, 734
Waveguides (microwave)
  bridge, 690
  completely-filled approach, 688
  equipment, 679
  measurement accuracy, 682
  microwave, 665–666
  resonator, 663
  short-circuited technique, 679
  wavelength parameters, 681
Wavelengths
  definition, 189
  microwave, 647
  microwave NDE, 648
  wave propagation, 72
Wavenumber, propagation constants, 657 (*see also* Angular spatial frequency)

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Waveplate, definition, 792
Wear plate, piezoelectric transducer, 127–128
Wedge method
  definition, 189
  SAW, ultrasonic inspection principle, 151–152
Welds
  defects, 558
  inspection, traditional film (projection) radiography, 555–556
  MPI, 249–250
Wet horizontal magnetic particle units, 224–226
Wet method (MP), definition, 258
Wet (wetting ability), definition, 60–61
Wheatstone bridge, 317–318
  definition, 365
White radiation, definition, 581
Wien's law, 623
Wire-type IQIs (radiography), 527
Wood, audible AE, 369
Working environment, NDE reliability, 13
Wrapped phase, 771

X-cut crystals (UT), definition, 189
X-ray fluorescence (XRF), 551–552
X-rays, 448–451
  advantages, 457–464
  adverse effects, 455–456
  amusement machines, 452, 453
  attenuation, 475–479
  breast cancer, 452
  diffraction, 550–551
  disadvantages, 457–464
  electromagnetic waves, 448
  film
    latitude, 505
    range, 505
    span, 505
  gauge measurement, 448
  hazards, 456
  history, 451–457
  imaging, new methods, 457
  introduction, 448
  machines, 487–494
    filament, 488
    tube current, 491
    tubes, 488
      industrial, 493
      medical, 493
  mechanisms, 469
  medical use, 453
  NDE, 447
  popularity, 455
  shoe fitting, 455
  technicians, 455–456
  tubes, 467, 488, 491, 493
    electrically powered, 467
    radiation spectrum, 468
  women, 455

Y-cut crystals (UT), definition, 189
Yellow-red spectrum, PT illumination principle, 28
Yoke method (MP), 223–224, 225
  definition, 258
Young's double slit experiments, 732–734
  illustration, 732–734
  observation plane, 732–734
Young's interferometer, definition, 790
Young's modulus, linear elastic media, 91

Zonal location, AE, 410–411

Copyright © Marcel Dekker, Inc. All rights reserved.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016