

WELDING

Processes, Quality, and
Applications



*Mechanical
Engineering Theory
and Applications*

Richard J. Klein
Editor

NOVA

**WELDING:
PROCESSES, QUALITY,
AND APPLICATIONS**

No part of this digital document may be reproduced, stored in a retrieval system or transmitted in any form or by any means. The publisher has taken reasonable care in the preparation of this digital document, but makes no expressed or implied warranty of any kind and assumes no responsibility for any errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of information contained herein. This digital document is sold with the clear understanding that the publisher is not engaged in rendering legal, medical or any other professional services.

MECHANICAL ENGINEERING THEORY AND APPLICATIONS

Additional books in this series can be found on Nova's website under the Series tab.

Additional E-books in this series can be found on Nova's website under the E-books tab.

MECHANICAL ENGINEERING THEORY AND APPLICATIONS

**WELDING:
PROCESSES, QUALITY,
AND APPLICATIONS**

**RICHARD J. KLEIN
EDITOR**



Nova Science Publishers, Inc.
New York

Copyright © 2011 by Nova Science Publishers, Inc.

All rights reserved. No part of this book may be reproduced, stored in a retrieval system or transmitted in any form or by any means: electronic, electrostatic, magnetic, tape, mechanical photocopying, recording or otherwise without the written permission of the Publisher.

For permission to use material from this book please contact us:

Telephone 631-231-7269; Fax 631-231-8175

Web Site: <http://www.novapublishers.com>

NOTICE TO THE READER

The Publisher has taken reasonable care in the preparation of this book, but makes no expressed or implied warranty of any kind and assumes no responsibility for any errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of information contained in this book. The Publisher shall not be liable for any special, consequential, or exemplary damages resulting, in whole or in part, from the readers' use of, or reliance upon, this material. Any parts of this book based on government reports are so indicated and copyright is claimed for those parts to the extent applicable to compilations of such works.

Independent verification should be sought for any data, advice or recommendations contained in this book. In addition, no responsibility is assumed by the publisher for any injury and/or damage to persons or property arising from any methods, products, instructions, ideas or otherwise contained in this publication.

This publication is designed to provide accurate and authoritative information with regard to the subject matter covered herein. It is sold with the clear understanding that the Publisher is not engaged in rendering legal or any other professional services. If legal or any other expert assistance is required, the services of a competent person should be sought. FROM A DECLARATION OF PARTICIPANTS JOINTLY ADOPTED BY A COMMITTEE OF THE AMERICAN BAR ASSOCIATION AND A COMMITTEE OF PUBLISHERS.

Additional color graphics may be available in the e-book version of this book.

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Welding : processes, quality, and applications / editor, Richard J. Klein.

p. cm.

Includes index.

ISBN 978-1-61761-544-3 (eBook)

1. Welding. I. Klein, Richard J., 1966-

TS227.W4135 2010

671.5'2--dc22

2010029834

Published by Nova Science Publishers, Inc. + New York

CONTENTS

Preface		vii
Chapter 1	Design of High Brightness Welding Electron Guns and Characterization of Intense Electron Beam Quality <i>G. Mladenov and E. Koleva</i>	1
Chapter 2	Process Parameter Optimization and Quality Improvement at Electron Beam Welding <i>Elena Koleva and Georgi Mladenov</i>	101
Chapter 3	Automation in Determining the Optimal Parameters for TIG Welding of Shells <i>Asif Iqbal, Naeem Ullah Dar and Muhammad Ejaz Qureshi</i>	167
Chapter 4	Friction Stir Welding: Flow Behaviour and Material Interactions of Two Similar and Two Dissimilar Metals and Their Weldment Properties <i>Indra Putra Almanar and Zuhailawati Hussain</i>	227
Chapter 5	Plastic Limit Load Solutions for Highly Undermatched Welded Joints <i>Sergei Alexandrov</i>	263
Chapter 6	Fracture and Fatigue Assessment of Welded Structures <i>S. Cicero and F. Gutiérrez-Solana</i>	333
Chapter 7	Laser Transmission Welding: A Novel Technique in Plastic Joining <i>Bappa Acherjee, Arunanshu S. Kuar, Souren Mitra and Dipten Misra</i>	365
Chapter 8	Effect of in Situ Reaction on the Property of Pulsed Nd:YAG Laser Welding SiCp/A356 <i>Kelvii Wei Guo and Hon Yuen Tam</i>	389
Chapter 9	Residual Stress Evolution in Welded Joints Subject to four-Point Bending Fatigue Load <i>M. De Giorgi, R. Nobile and V. Dattoma</i>	407
Index		421

PREFACE

Chapter 1 - At the beginning of this chapter the integral description and the micro-characterization of an intense electron beam are discussed. The beam parameters determination is given on base of the distribution functions and other beam characteristics in coordinate and impulse planes.

The analysis of powerful beams, utilized for electron beam welding (EBW) of machine parts, could be perfect, if we measure or calculate both: the radial and the angular beam current distributions. The beam emittance, involving these parameters, is the chosen value for the quality characterization of technology electron beams. In this way monitoring of the beam profile (i.e. distribution of the beam current density in a beam transverse cross-section) and evaluation the beam emittance are needed at standardization of EBW equipment and at providing the reproducibility of the EBW conditions.

Techniques, schemes and limits of such monitoring are described and analyzed. The signal formation features at devices for estimation of the beam profile of intense continuously operated electron beams are given. The role of space-frequency characteristics of the sampling scanning (modulation) system; limitations and peculiarities at assuming normal distribution of the monitored beam current density; the use of Abel back transformation; the application of computer-tomography method for the measuring the beam profile and the methods for simplification the estimation of the beam emittance are discussed.

In this chapter the effects of the negative space charge of beam electrons in the intense electron beam on the current and on the radial dimensions as well as the role of total and local compensation of that charge by the generated ions in zone of interaction beam/material or through the residual gases in the technology chamber are discussed.

The more important data and relations for the design of technology electron guns and for the simulation of the generated intense electron beams are given. Computer simulation of the technology guns, based on phase analysis of the beam, instead of the conventional trajectory analysis is described. In the presented original computer code, the velocity distribution of the emitted from cathode electrons, is taken into account too.

Some examples of computer simulation of technology electron guns for electron beam welding and beam diagnostics of high power low voltage electron beams are given.

Chapter 2 - The complexity of the processes occurring during electron beam welding (EBW) at intensive electron beam interaction with the material in the welding pool and the vaporized treated material hinders the development of physical or heat model for enough accurate prediction of the geometry of the weld cross-section and adequate electron beam welding process parameter selection. Concrete reason for the lack of adequate prognostication

is the casual choice of the heat source intensity distribution, not taking into account the focus position toward the sample surface and the space and angle distribution of the electron beam power density. This approach, despite extending the application of solution of the heat transfer balance equations with the data of considerable number of experiments, results in prognostication of the weld depth and width only in order of magnitude. Such models are not suitable for the contemporary computer expert system, directed toward the aid for welding installation operator at the process parameter choice and are even less acceptable for automation EBW process control.

Various approaches for estimation of adequate models for the relation between the electron beam weld characteristics and the process parameters, the utilization of these models for process parameter choice and optimization are considered.

A statistical approach, based on experimental investigations, can be used for model estimation describing the dependence of the welding quality characteristics (weld depth, width, thermal efficiency) on the EBW process parameters - beam power, welding speed, the value of distance between the electron gun and both the focusing plane of the beam and the sample surface as parameters. Another approach is to estimate neural network-based models. The neural networks were trained using a set of experimental data for the prediction of the geometry characteristics of the welds and the thermal efficiency and the obtained models are validated.

In the EBW applications an important task is to obtain a definite geometry of the seam as well as to find the regimes where the results will repeat with less deviations from the desired values. In order to improve the quality of the process in production conditions an original model-based approach is developed.

Process parameter optimization according the requirements toward the weld characteristics is considered. For the quality improvement in production conditions, optimization includes finding regimes at which the corresponding weld characteristics are less sensitive (robust) to variations in the process parameters.

The described approaches represent the functional elements of the developed expert system.

Chapter 3 - Residual stresses and distortion are the two most common mechanical imperfections caused by any arc welding process and Tungsten Inert Gas (TIG) Welding is no exception to this. A high degree of process complexity makes it impossible to model the TIG welding process using analytical means. Moreover, the involvement of several influential process parameters makes the modeling task intricate for the statistical tools as well. The situation, thus, calls for nonconventional means to model weld strength, residual stresses and distortions (and to find trade-off among them) based on comprehensive experimental data.

Comprehensive Designs of Experiments were developed for the generation of relevant data related to linear and circumferential joining of thin walled cylindrical shells. The base metal utilized was a High-Strength Low Alloy Steel. The main process parameters investigated in the study were welding current, welding voltage, welding speed, shell/sheet thickness, option for trailing (Argon), and weld type (linear and circumferential).

For simultaneous maximization/minimization and trade-off among aforementioned performance measures, a knowledge base – utilizing fuzzy reasoning – was developed. The knowledge-base consisted of two rule-bases: one for determining the optimal values of the process parameters according to the desired combination of maximization and/or minimization of different performance measures; while the other for predicting the values of

the performance measures based on the optimized/selected values of the various process parameters. The optimal formation of the two rule-bases was done using Simulated Annealing Algorithm.

In the next stage, a machine learning (ML) technique was utilized for creation of an expert system, named as EXWeldHSLASteel, that could: self-retrieve and self-store the experimental data; automatically develop fuzzy sets for the numeric variables involved; automatically generate rules for optimization and prediction rule-bases; resolve the conflict among contradictory rules; and automatically update the interface of expert system according to the newly introduced TIG welding process variables.

The presented expert system is used for deciding the values of important welding process parameters as per objective before the start of the actual welding process on shop floor. The expert system developed in the domain of welding for optimizing the welding process of thin walled HSLA steel structures possesses all capabilities to adapt effectively to the unpredictable and continuously changing industrial environment of mechanical fabrication and manufacturing.

Chapter 4 - In friction stir welding of two similar and dissimilar metals, the work materials are butted together with a tool stirrer probe positioned on the welding line. The work materials in the welding area are softened due to heat generation through friction between the probe and the surface of the work materials. Upon the softening of the work materials, the friction will be diminished due to the loss of frictional force applied between the tool stirrer probe and the softening surface of work materials. The probe then penetrates the work material upon the application of the axial load and the tool shoulder confines the working volume. In this configuration, the advancing and retreating zones are created relevant to the direction of the probe rotational direction. At the same time the leading and trailing zones are also created relevant to the direction of motion of the tool. These zones determine the flow behavior of the softened work materials, which determine the properties of the weldment. Since the chemical, mechanical, and thermal properties of materials are different, the flow behavior of dissimilar materials becomes complex. In addition, material interaction in the softened work materials influences material flow and mechanical intermixing in the weldment. This review discusses the fundamental understanding in flow behavior of metal during the friction stir welding process and its metallurgical consequences. The focus is on materials interaction, microstructural formation and weldment properties for the similar and dissimilar metals. Working principles of the process are explained beforehand.

Chapter 5 - Limit load is an essential input parameter in many engineering applications. In the case of welded structures with cracks, a number of parameters on which the limit load depends, such as those with the units of length, makes it difficult to present the results of numerical solutions in a form convenient for direct engineering applications, such as flaw assessment procedures. Therefore, the development of sufficiently accurate analytical and semi-analytical approaches is of interest for applications. The present paper deals with limit load solutions for highly undermatched welded joints (the yield stress of the base material is much higher than the yield stress of the weld material). Such a combination of material properties is typical for some aluminum alloys used in structural applications.

Chapter 6 - The presence of damage in engineering structures and components may have different origins and mechanisms, basically depending on the type of component, loading and environmental conditions and material performance. Four major modes or processes have

generally been identified as the most frequent causes of failure in engineering structures and components: fracture, fatigue, creep and corrosion (including environmental assisted cracking), together with the interactions between all of these. As a consequence, different Fitness-for-Service (FFS) methodologies have been developed with the aim of covering the mentioned failure modes, giving rise to a whole engineering discipline known as structural integrity.

At the same time, welds can be considered as singular structural details, as they may have, among others features, noticeably different mechanical properties from the base material (both tensile properties and toughness), geometrical singularities causing stress concentrations, and residual stresses with specific profiles depending on the type of weld and welding process. Traditional approaches to the assessment of welds have consisted in making successive conservative assumptions that lead to over-conservative results. This has led to the development, from a more precise knowledge of weld behavior and performance, of specific Fitness-for-Service (FFS) assessment procedures for welds which offer great improvements with respect to traditional approaches and lead to more accurate (and still safe) results or predictions.

The main aim of this chapter is to present these advanced Fitness-for-Service (FFS) tools for the assessment of welds and welded structures in relation to two of the above-mentioned main failure modes: fracture and fatigue.

Chapter 7 - Plastics are found in a wide variety of products from the very simple to the extremely complex, from domestic products to food and medical product packages, electrical devices, electronics and automobiles because of their good strength to weight ratio, ease of fabrication of complex shapes, low cost and ease of recycling. Laser transmission welding is a novel method of joining a variety of thermoplastics. It offers specific process advantages over conventional plastic welding techniques, such as short welding cycle times while providing optically and qualitatively high-grade joints. Laser transmission welding of plastic is also advantageous in that it is non-contact, non-contaminating, precise, and flexible process, and it is easy to control and automate.

This chapter discusses all major scientific and technological aspects concerning laser transmission welding of thermoplastics that highlights the process fundamentals and how processing affects the performance of the welded thermoplastic components. With the frame of this discussion the different strategies of laser transmission welding of plastic parts are also addressed. Finally, applications of laser transmission welding are presented, which demonstrates the industrial implementation potential of this novel plastic welding technology.

Chapter 8 - The effect of in situ reaction on the properties of pulsed Nd:YAG laser welded joints of particle reinforcement aluminum matrix composite SiCp/A356 with Ti filler was studied, and its corresponding temperature field was simulated. Results shows that in situ reaction during the laser welding restrains the pernicious Al_4C_3 forming in the welded joints effectively. At the same time, the in situ formed TiC phase distributes uniformly in the weld, and the tensile strength of welded joints is improved distinctly. Furthermore simulation results illustrate that in addition to the lower heat-input into the substrate because of Ti melting, in situ reaction as an endothermic reaction decreases the heat-input further, and its temperature field distributes more smoothly with in situ reaction than that of laser welding directly. Also, the succedent fatigue test shows the antifatigue property of welded joints with in situ reaction is superior to that of traditional laser welding. It demonstrates that particle

reinforcement aluminum matrix composite SiCp/A356 was successfully welded by pulsed Nd:YAG laser with in situ reaction.

Chapter 9 - Residual stresses, introduced into a component by manufacturing processes, significantly affect the fatigue behaviour of the component. External load application produces an alteration in the initial residual stress distribution, so it is reasonable to suppose that residual stress field into a component subject to a cyclic load presents an evolution during the total life. In this work, the authors analysed the evolution that the residual stress field, pre-existing in a butt-welded joint, suffers following the application of cyclic load. The comparison between two residual stress measurements, carried out on the same joint before and after the cyclic load application, allowed to obtain interesting information about the residual stress evolution. It was found that in particular condition, unlike the general opinion, a cyclic load application produces an increasing in the residual stress level rather than a relaxation. This phenomenon is to take well in account in order to avoid unexpected failure in components subjected to a fatigue load.

Chapter 1

DESIGN OF HIGH BRIGHTNESS WELDING ELECTRON GUNS AND CHARACTERIZATION OF INTENSE ELECTRON BEAM QUALITY

G. Mladenov and E. Koleva

Institute of Electronics, Bulgarian Academy of Sciences, Sofia, Bulgaria

ABSTRACT

At the beginning of this chapter the integral description and the micro-characterization of an intense electron beam are discussed. The beam parameters determination is given on base of the distribution functions and other beam characteristics in coordinate and impulse planes.

The analysis of powerful beams, utilized for electron beam welding (EBW) of machine parts, could be perfect, if we measure or calculate both: the radial and the angular beam current distributions. The beam emittance, involving these parameters, is the chosen value for the quality characterization of technology electron beams. In this way monitoring of the beam profile (i.e. distribution of the beam current density in a beam transverse cross-section) and evaluation the beam emittance are needed at standardization of EBW equipment and at providing the reproducibility of the EBW conditions.

Techniques, schemes and limits of such monitoring are described and analyzed. The signal formation features at devices for estimation of the beam profile of intense continuously operated electron beams are given. The role of space-frequency characteristics of the sampling scanning (modulation) system; limitations and peculiarities at assuming normal distribution of the monitored beam current density; the use of Abel back transformation; the application of computer-tomography method for the measuring the beam profile and the methods for simplification the estimation of the beam emittance are discussed.

In this chapter the effects of the negative space charge of beam electrons in the intense electron beam on the current and on the radial dimensions as well as the role of total and local compensation of that charge by the generated ions in zone of interaction beam/material or through the residual gases in the technology chamber are discussed.

The more important data and relations for the design of technology electron guns and for the simulation of the generated intense electron beams are given. Computer simulation of the technology guns, based on phase analysis of the beam, instead of the conventional trajectory analysis is described. In the presented original computer code, the velocity distribution of the emitted from cathode electrons, is taken into account too.

Some examples of computer simulation of technology electron guns for electron beam welding and beam diagnostics of high power low voltage electron beams are given.

INTRODUCTION

The conventional method for setting the beam power distribution in a plant for electron beam welding (EBW) relies on the operator visually to focus the beam on a secondary target situated near the welded parts. This requires significant operator experience and judgment, but in each case different settings could be obtained due to the subjective visual interpretation of the observed picture of the interaction of intense beam with the sample surface.

For the applications of the advantages of electron beam welding it is necessary to know in details the properties of the electron beam. There are only standards for measurements of electron beam current and accelerating voltage as beam characteristics, applicable at the acceptance inspection of electron beam welding machine [1] or at process investigations. These parameters could not characterize the quality of produced electron beam in terms of their ability to be transported over long distances, to be focused into a small space with a minimum of divergence. The directional energy flow is the main feature of the non-conventional welding heat sources- the electron beam and the laser beam. At the case of use of laser beams the photon intensity profile and M^2 measures [2] are the quality parameters of the beam that evaluation are important step to standardization of powerful laser beams.

The reproducibility of the product performance characteristics, the optimization and quality improvement of the results of EBW, as well as the transfer of concrete technology from one EBW installation to another, need quantitative diagnostics of the intense electron beams quality. At responsible joining of details periodic measurements of the beam parameters could safe the obtaining welds with equal parameters. During the design stage of EBW guns such characterization is useful as a measure used for their optimization and comparison.

High brightness electron beams are a subject of interest among researchers and designers promoting technology applications of concentrated energy beam sources, namely in the field of EBW. Computer simulation of generated beam is of considerable importance for creation of a perfect from electron-optical point of view welding electron gun. The quality of electron beam welds is directly connected with the generated intense beam characteristics and in that way with the optimization of electron gun parts.

1. CHARACTERIZATION OF INTENSE ELECTRON BEAMS

General Description of the Behavior of Electron Beams

A beam is ensemble of moved in nearly one direction electrons. The beam electrons are accelerated to a kinetic energy in an electrical field. Often, together with these quick (high energy) electrons in the beam space there are a quantity of low energy electrons and ions. The beam particles velocity distribution is non-isotropic, and these particles are non-uniformly distributed in the space. In such a way the beam is a non equilibrium system from thermodynamic point of view. The kinetic energy of the beam particles is much higher than the energy of interactions forces between the beam electrons.

The interaction forces between the beam electrons are usually of electrostatic character. Electromagnetic interactions have place only in case of relativistic velocities of beam electrons or in the case of full compensation of the electrostatic forces between the beam particles by low energy ions, situated also in the beam space [3, 4].

The behavior of the beam electrons is determined strongly by their space density. In the case of a low density of the beam current and correspondingly at low interactions between beam electrons, the beam can be assumed as a *system of non-interacting electrons*. The behavior of every particle in such a beam is controlling by electron optics rules. In such geometry optics the trajectory of every beam electron is similar to the light ray behavior in the light optics.

At increase of the beam electrons density the interaction energy due to the electrostatic forces, acting between the neighboring electrons elevate too, and particles behavior have a group character. The trajectories of a separated beam particle and the configuration of the beam envelope (boundary distribution) are function of common electric field, i.e. by the position of all adjacent beam particles in the studied time moment. This field is result of action of too many particles and is not controlling by exact position of the near neighbor electrons or by the exact corpuscle beam structure. These beams are called *intense beams* of electrons and the boundary between a beam of non-interacting particles and a intense electron beam is given by a perveance critical value of $10^{-7} - 10^{-8} \text{ A.V}^{-3/2}$ (see below for the definition of the perveance value the equation (28)).

In the case of a higher particles density in the beam, *the direct two-particles interactions between beam particles* take place. The electron group emitted from the cathode of the electron gun has a velocity distribution in the form of the Maxwell's distribution. In the course of formation of a fine electron beam, the current density of the beam increases, and the velocity distribution of the beam is broadened by energy relaxation due to the Coulomb's force acting between the electrons. This phenomenon known in the literature as Boersh effect [5], and the broadening rate of the velocity distribution of the beam is generally proportional to $j(z)^{1/3}$, when $j(z)$ is the beam current density on the beam axis.

All corporate effects (common electrostatic forces and two-body interactions of the beam electrons) lead to limitations of the beam minimal cross section as well as of the maximal density of the kinetic energy of the beam.

In many cases of technology applications, in the beam space there are also neutral or low energy charged particles. The interaction of the beam particles with these low energy atomic

particles is function of the relative velocity and nature of interacting corpuscles. In the potential gap, generating by the negative charge of the beam electrons, the newly generating by beam low energy ions are collected. This leads to neutralization of the beam space charge and in end case can shake off newly generated compensating ions from the space of the beam. Such *beam is overcompensated*. There is a possibility to have also only *locally neutralized beam* [6] (see below too).

In the case of higher densities of the low energy charged particles situated in the beam space (namely plasma) there is a group interaction between the beam electrons and the low energy plasma corpuscles. This leads to intensive transfer of beam particles energy to the plasma component and various effects of instabilities of the beam could be occurs. This phenomenon is directed to achievement a more stable equilibrium of the particles system namely the beam space extending and the smoothing its energy distribution.

The Electron Beam Macro-Characterization. Beam Integral Characteristics: Current, Energy and Diameter

The basic values that can characterize a beam of charged particles are number and energy distribution of the beam particles. For an intense electron beam the basic parameters, numerically determining the main integral characteristics are: the *beam current* I_0 [A], the *accelerating voltage* U_a [V] and the *dimensions of beam cross section* in a studied point along the beam axis and time. They characterize the mean number of passing through studied cross-section electrons, as the individual kinetic energy of these particles $E_0 \approx eU_a$ and the energy density of the beam, consisting of nearly mono-energetic electrons. The beam power $P_0 = U_a I_0$ [W] is characteristics of the beam average energy flow, transferred through studied cross-section per unit of time.

The measurements of the current and the accelerating high voltage (often the value of the beam current is assumed to be equal to the electrical source current) are technically resolved tasks. The characteristics describing the spatial distribution of the electrons and their energy in the beam are important when using the beams as sample treating instrument during technological processes. These characteristics are difficult to be measured due to the wide and smooth decrease of the distribution of the beam current in the beam boundary region (there is no clear limit between the beam and the surrounding area). That is why two approximate characteristic are used: the beam diameter (or two corresponding cross-section dimensions - width and length, when the beam is flat) and power density at definite cross-section, usually the one upon the processed material. The determination of the beam diameter as the dimension of the beam wide - in a general case is function of the sensitivity of measuring instrument or of a previously chosen limiting (minimal) value of the beam current density resolution.

It is convenient to characterize the beam current distribution across the beam (radial distribution in the case of axially-symmetrical case) by the maximal current distribution value and any other value, determined on a pre-chosen distance from the beam axis. For axially-symmetrical beams one can assume, that the beam diameter is distances where the beam current distribution is 1/2; 1/e or 1/20 of its maximal values. Respectively one can signify the beam diameters as $d_{0.5}$, d_0 , $d_{0.05}$.

In the many of practical cases a Gaussian distribution of the beam current in a beam cross section can be observed. In the case of a axially-symmetrical beam, that distribution can be written as:

$$j(r) = j(0) \exp\left(-\frac{r^2}{r_0^2}\right), \quad (1)$$

where r is distance to the studied point in the chosen cross section, measured from beam axis; $j(0)$ is the current density on the beam axis, r_0 is the beam radius at which the $j(r_0) = j(0)/e$, where $e \approx 2.72$ is the natural logarithm constant.

Integrating (1) between 0 and radius r , one can found the value of the current, transferred through such part of the beam cross section:

$$I_r = \pi \cdot r_0^2 \cdot j(0) \cdot [1 - \exp(-\frac{r^2}{r_0^2})] = I_0 \cdot [1 - \exp(-\frac{r^2}{r_0^2})], \quad (2)$$

where $I_0 = \pi \cdot r_0^2 \cdot j(0)$ is the beam current.

Than the current I_r , transferred through a part of the beam cross section of diameter $d_0 = 2r_0$; $d_{0.5}$ or $d_{0.05}$ (the indexes 0,5 and 0.05 means that there the beam current density is $j(r_{0.5}) = j(0)/2$ or $j(r_{0.05}) = j(0)/20$ respectively) is correspondingly 63%, 50% or 95% from the beam current I_0 (at Gaussian current density distribution).

In the case of the band like beam with coordinate axis x situated across the beam cross section and a uniform current distribution along the wider side of the beam cross-section (coordinate y) the respective current distribution will be:

$$I_x = I_0 \cdot \text{erf}(x/x_0), \quad (3)$$

Where $\text{erf}(x/x_0)$ is error function $\Phi(\alpha)$ called some times Integral of the error probability:

$$\text{erf}(x/x_0) = \Phi(\alpha) = \frac{2}{\sqrt{\pi}} \int_0^{\frac{x}{x_0}} \exp\left(-\frac{x}{x_0}\right)^2 \cdot d\left(\frac{x}{x_0}\right). \quad (4)$$

Function $\Phi(\alpha)$ is given in many handbooks in tabulated form. That function is given also on Figure 1.

Then through a gap with wide $2x_{0.5}$, $2x_0$ and $2x_{0.05}$, defined as $d_{0.5}$; d_0 or $d_{0.05}$, will be transferred current 75%, 84% and 98% from the beam current I_0 .

In the general case of axis-symmetrical cross section of a electron beam with a Gaussian distribution of the current (1), if the diameter of the beam defined at a level $1/a$ from the maximum beam density, the relationship between r_0 , defined at level $1/e$ and r'_0 , defined in this way is:

$$r'_0 = r_0 (\ln a)^{1/2}. \quad (5)$$

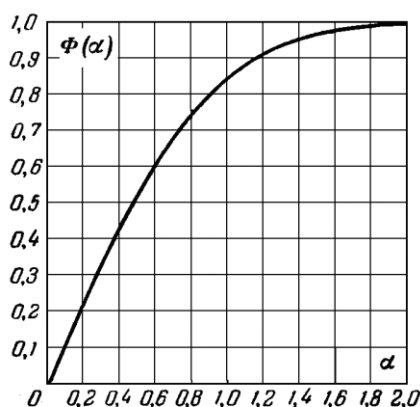


Figure 1. Error function $\Phi(\alpha)$ (from equation (4)) versus $\alpha=(x/x_0)$

Table 1. Beam parameters at various technology processes

EB Process	Typical parameters of the electron beam			
	Acceleration voltage, U_a [kV]	Diameter or width of the beam on the processed material $2r_0$, [mm]	Beam power P_0 , [kW]	Average power density, [W/cm ²]
EB surface thermal processing	115-150	0,1-1	1-15	104-106
EB melting and casting	15-35	5-80	10-5000	103-5.104
EB evaporation	10-30	2-25	0,1-100	103-105
EB welding	15-150	10-1-2	0,1-100	105-5.107
Electron radiation processing	50-5000	100-800	1-100	1-103
Thermal size processing	20-150	5.10-3-10-1	10-2-1	105-5.109
EB lithography	5-70	7.10-6-150	10-7-10-3	10-4-104
Electron microscopes, micro X-ray analysis and other methods of analysis with electron beams	1-1000	3.10-610-1	10-8-10-2	10-4-103

The power density, defined assuming a uniform distribution of the beam power in a spot with diameter $2r_0$, is another characteristic of the effect of power electron beams on the processed materials. Lots of the physical effects during this interaction depend directly on this characteristic's value. An idea for the numerical values of the mentioned characteristics of the *electron beams*, used for the material processing and analysis, is given in Table 1.

The use of the electron beam as a technological instrument in many technological processes is based on the possibility for a local interaction with the processed material. The diameter of the beam in the area of interaction in electron beam lithography and the other methods for analysis of materials with electron beam is around $(30-70) \cdot 10^{-10}$ m. The directed local interaction leads to better using of the energy of the electron beam at the use of the energy of the electrons transferred in thermal kinetic energy. The electron beam yields only to laser beams the reached power density, but they lead in efficiency of transfer of energy and

the possibilities for control of the process. There is no refractory or thermal-shock resistant material, which cannot be processed with electron beam. This is the basis for the thermal size processing (cutting, drilling, fixing exact sizes and values of resistivity of thin film resistors etc.), as well as the electron beam welding, evaporation etc. The high efficiency of transfer of the energy and the clean environment (the process is usually held in vacuum) made the use of powerful electron beams in metallurgy, for the fabrication and refining of high purity refractory metals and alloys, through electron beam melting and evaporation, a prospective industrial technology. Irradiation with beams of accelerated electrons is applied in many chemical processes of polymerization or treatment of food and medical supplies and instruments etc. Here the controlled effect on definite chemical bonds or biological structures makes the process more efficient energetically than the conventional thermal methods for treatment. The use of higher acceleration voltages leads to higher efficiency during the irradiation of thicker layers of the treated material.

Micro-Characterization of a Charged Particles Beam. Distribution Functions and Differential Characteristics of the Beams

Beams, as was mentioned, are composed from a big number of electrons. The beam state can be defined by an array of the coordinates and the impulse values of every particle in this composition. For characterization of an electron beam the number of particles in the elementary volume $d\vec{q}$ around the space coordinate $\vec{q}$ and the impulses $d\vec{p}$ around the impulse values $\vec{p}$, in moment t , that is connected with the *distribution function* $f(\vec{p}, \vec{q}, t)$ are used. This function of space and impulses distribution of the particles in the time t is normalized on total number of particles in the beam and is also called *phase density of beam electrons*:

$$dN(\vec{p}, \vec{q}, t) = f(\vec{p}, \vec{q}, t) \cdot d\vec{p} \cdot d\vec{q} \quad (6)$$

Usually instead (6) are written an equivalent equation, that for an axis-symmetrical beam is:

$$dN(\vec{r}, \vec{\Omega}, E, t) = f(\vec{r}, \vec{\Omega}, E, t) \cdot d\vec{r} \cdot d\Omega \cdot dE \cdot dt \quad (6)$$

Here $\vec{\Omega}$ is a vector unit in the particle velocity direction $\vec{V}$, and E is the kinetic energy of particles; dN and f are the number of particles and probability they to be in the volume of phase space $(\vec{r}, \vec{\Omega}, E, t)$. Then number of electrons dN , owning energy in the region $E \div E + dE$, and being in the elementary volume $d\vec{r}$, situated around the point $\vec{r}$, as well as

moving in the space angle $d\Omega$ around the vector-unit $\vec{\Omega}$, in the time moment t is given by equation (6).

In the case of interaction on beam particles with an outer field or after collisions between the particles, that change its impulses, the distribution function is unsteady. Opposite, for a beam of non-interacting particles the distribution function is not varying during the time. In the former case is applicable the Liouville's theorem for a beam of non-interacting particles, which states that particle density in 6-dimensional phase space of coordinates and impulses of the particles is value, that is invariant due to track length of the beam.

Using equation (6) one can find the corresponding particle's densities, depending by one or other parameter. Such are the space and energy particles distributions and the time dependent density of particles. For one chosen cross-section of the beam one can define the radial distribution of the particles, as well as - the angular particles distributions; the distributions of the particles energy and the time variations of the particles density in a point of the phase space).

Another characteristic of the beams is the values - stream, flow or flux of particles; stream of energy and stream of charges, propagating through a plane (beam cross-section) at one unit time. That information is applicable in the technology evaluations. In the case of r

becoming projection of the vector $\vec{r}$ in that cross-section-i.e. r is the distance from the axis to that point. Then if assuming a steady stream of charged particles through elementary area

dS (characterized by its normal vector $\vec{S}$) around a point with coordinate r , the differential particle flux, in which particles are with energy E , and the particles are moving in direction of vector $\vec{\Omega}$, one can write:

$$d\Phi(r, \vec{\Omega}, E) = (\vec{S} \cdot \vec{\Omega}) V f(r, \vec{\Omega}, E) d\Omega dE dS, \quad (7)$$

where $V = |\vec{V}|$, a $\vec{V} = V \cdot \vec{\Omega}$.

Let one define distribution function of the fluxes in the beam:

$$F_F(r, \vec{\Omega}, E) = V \cos(\vec{S} \cdot \vec{\Omega}) f(r, \vec{\Omega}, E).$$

Then, after suitable integrating one can find the streams of various groups of particles. As an example the *integral flux of particles* in the beam is given as:

$$F = \int_S \int_{\Omega} \int_E f_F(r, \vec{\Omega}, E) dS d\Omega dE; \quad (8)$$

The corresponding flux of charges is :

$$F_Q = I = \oint_S \int_{\Omega} \int_E q \cdot f_F(r, \vec{\Omega}, E) \cdot dS \cdot d\Omega \cdot dE; \quad (9)$$

and the flux of energy is respectively:

$$F_E = \oint_S \int_{\Omega} \int_E E \cdot f_F(r, \vec{\Omega}, E) \cdot dS \cdot d\Omega \cdot dE; \quad (10)$$

Besides the integral fluxes one can define the corresponding densities of the fluxes. As example the density of the particles flux can be written:

$$\varphi = \frac{dF}{dS} = \iint f_F(r, \vec{\Omega}, E) \cdot d\Omega \cdot dE. \quad (11)$$

An other value, finding wider application is the current (i.e. flux of charges):

$$\varphi = j = \frac{dF_Q}{dS} = \iint q \cdot f_F(r, \vec{\Omega}, E) \cdot d\Omega \cdot dE. \quad (12)$$

In an annalogical way is written the density of the energy flux.

In the cases when is needed to take in account the angular distribution of particle fluxes in the beam (as example - that is necessary at characterization of the sources of accelerated charged particles or in the case of deep penetration of the particles in irradiated material) the detail characterization of the beam can be given knowing the differential brightness in many concrete points. Measured by particles stream that differential brightness is:

$$b(r, \vec{\Omega}) = \frac{d^2 F(r, \vec{\Omega})}{dS \cdot d\Omega} = \int_E f_F(r, \Omega, E) \cdot dE. \quad (13)$$

The differential brightness measured by charge is:

$$b_Q(r, \vec{\Omega}) = \frac{d^2 F_Q(r, \vec{\Omega})}{dS \cdot d\Omega} = \int_E q \cdot f_F(r, \Omega, E) \cdot dE; \quad (14)$$

In an analogical way one can define the brightness of energy flux in the given point.

In the general case the density and fluxes are varying on the beam cross-section. Due to that very often are evaluated the average values of that parameters. For example if one use the mean value of particles flux, averaged on cross-section and space angle of the whole beam - it is found *the mean brightness of the particles propagation in the beam B* :

$$B = \frac{\oint dF(r, \vec{\Omega}).dS.d\Omega}{\oint_S \int_{\Omega} dS.d\Omega}. \quad (15)$$

Here is assumed, that axis, around which is measured the space angle $\vec{\Omega}_0$ is in coincidence with the beam axis. The mean brightness in equation (15) is identical with the photometry's brightness.

At characterization of electron beams is usual to utilize the electron brightness. They are defined by mean value of current, flowing through the one unit area of investigated cross-section in an unit of the space angle Ω .

$$B_Q = \frac{I}{S.\Omega}. \quad (16)$$

Due to gradual slur of the particles flux distributions in the beam envelope at the estimation of the beam brightness is necessary an exact concrete definition of integration limits in every case. Only in the beam regions where the particle flux distributions are with sharper boundaries (cross-over, focus) these values are more clearly defined. In all other cross-sections these values are done only after special assumptions for sensitivity of measurements or exactness of determination.

Between the energy densities of beam fluxes characteristics more wide use there is the value

F_E/S , called *power density of the beam* (please understand that there is the mean value in exact definition). This value there is not characteristics of the direction of particles and mean energy fluxes of the beam. The power density of the electron beam at most of the technological applications is desirable to be maximum. It is defined by the spacial density of the electrons in the beam and their kinetic energy. Mainly due to the electrostatic repulsion forces between electrons and also due to technical difficulties (high-voltage isolators, x-ray prevention etc.) and the relative effects at increase of the acceleration voltage, the power density of the beam cannot increase unlimitedly. Table 1 shows that at many technological processes the numerical values of the power density of the beam are considerable. The objective laws for the movement of electrons in such beams, called intensive electron beams, differ from those in beams with lower concentration of electrons (power density), such as the used in electron microscopes.

Emittance and Brightness

An ideal intensive electron beam is such laminar electron beam, in which the distribution of the velocities of the electrons is defined in every point, i.e. the trajectories of the electrons do not cross. In reality, the chaotic initial velocities of emission of the electrons, the aberrations of the forming electron-optic system and the non-homogenities lead to non-laminar movement of the electrons of the beam. In these cases for characterization of the

beams is used the characteristic *emittance*, signed ε . In one axial-symmetrical beam under use is the plane rr' and here every trajectory can be presented by a point of coordinates - radius r (namely distance between electron trajectory and beam axis) and divergence or convergence angle of trajectory to the normal of beam axis $r'=(dr/dz)$.

The *emittance* is the divided to π area of the region on the plane rr' where are situated the points, representing the particles of the beam (Figure 2).

The stationary particles distribution function in one monochromatic stream there four variables: x, y, x', y' . For the geometry presentation more suitable is to use two-dimensional projections xx' and yy' . Here the sign ' means the first derivative of corresponding value taken on the distance measuring along beam z ($x' = dx/dz$; $y' = dy/dz$). There projections, together with the beam cross section are able to give sufficient visual aid.

The emittance is a quality characteristics of the beams that determine the non-laminarity of the particle trajectories in the beam. Less emittance value means higher brightness of the beam. As general, the emittance diagram is elliptical and inclination of ellipse axis demonstrated the convergent or divergent beam trajectories. For real electron beams the emittance is always larger than 0. In these beams the beam region is not clearly limited, the distribution of the points of the diagram in the plane rr' is not uniform, and it has decreasing density near the boundary region. Then, for the definition of the emittance the area, which contain a certain part of these points, e.g. 90% is used.

Since the numerical value of the emittance depends on the velocity of the electrons V_z in the movement direction often it is used the characteristic *normalized emittance* [7,8]:

$$\varepsilon_n = \left(\frac{V_z}{c} \right) \varepsilon, \quad (17)$$

where c is the velocity of light.

From the Liouville's theorem considering the movement of particles in the phase space (the space of the coordinates and the impulses of movement of the particles) follows that the value of the normalized emittance should not change along the whole length of the beam. This is true only for ideal systems without aberrations and non-homogeneities, as well as without collisions between the electrons and the particles of the environment and interaction between separate electrons.

As were mentioned the emittance is connected with the *electron brightness*. The emittance and the electron brightness, considered as characteristic of the electron beam, have advantage on the mentioned current density (or the power density) because these parameters contain also information about the direction of the impulses of the separate electrons. In most cases in technological applications this is an important characteristic.

The appointed above disadvantage of the electron brightness as a characteristic of the gathering of moving electrons is that it is difficult to measure and mainly - the more difficult and no generally accepted choice of the limits of averaging in any unspecified cross-section of the beam. In the characteristic cross-sections of the beam: at the cathode, at the narrowest place in front of it called crossover, at the place of the image of the cathode and in the focus spot after the focusing lens, the electron beams are better outlined and the choice of the area and the space angle for determination of the average electron brightness are not so undefined.

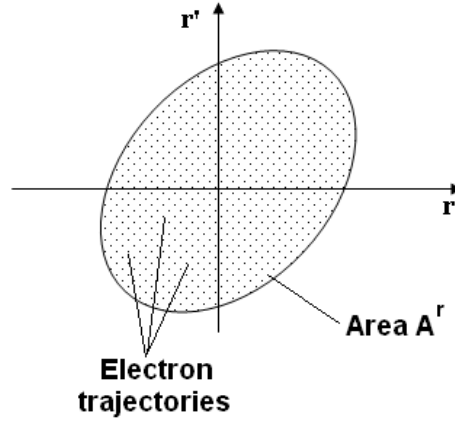


Figure 2. Diagram of the electron beam emittance

In order to avoid the difficulties when choosing the limits of averaging, it is accepted the following definition for the electron brightness:

$$B = \frac{\partial^2 I}{\partial s \partial \Omega}, \quad (18)$$

where ∂s and $\partial \Omega$ are small elements of the surface and the space angle. Here B characterizes the brightness in definite direction z ($\Omega=0$), and s is a corresponding normally placed surface. The brightness, corresponding to eq. (18) can be measured, by choosing and placing corresponding apertures and screens (Figure 3). Such brightness value is necessary for the determination and building of a more detailed emittance diagram in which areas with various brightness ranges could be distinguished.

In that a way, at differentiating the areas on the diagram with equal brightnesses, the respective beam parts can be considered as separated-independent sub-beams.

Beams with large brightness have small area at the diagram of the emittance, and this means small emittance value.

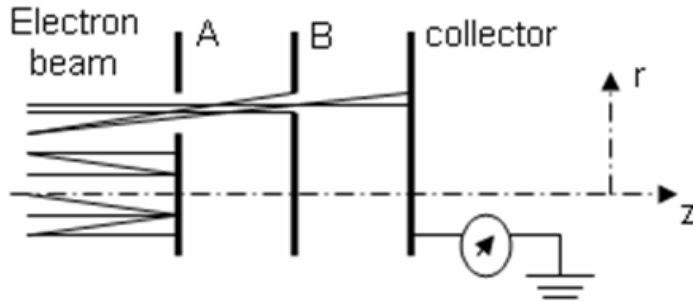


Figure 3. Scheme of emittance measurement of an electron beam in the plane. (The screen A is immovable; B moves. The position of the fissure in A defines r , and the one in B – the magnitude of r' for a given value of r)

An important characteristic of the electron guns and beams [9,10] is the *relative electron brightness* B/U , which is calculated as the electron brightness divided by the accelerating voltage. This characteristic corresponds to the normalized emittance and is constant along the beam in electron beam systems without aberrations. In real technological electron beam systems with intensive electron beams this invariability is a result also of partial or full compensation of the space charge of the beam. The knowledge of B/U gives possibility to compare electron beam systems, to choose highly effective emitters for them and to define the maximum possible current density or the power in the focus and the length of the active interaction zone.

Figure 4 presents data for the relative electron brightness B/U for some real electron beam welding systems.

The increase of the current of the beam leads to an increase of the radius of the cathode and of the crossover (the minimum cross-section of the beam in front of it), where the electron trajectories cross and the aberrations increase, as well as the electron brightness decreases. The increase in the space charge in the beam acts in the same direction. In the case of higher voltage guns the electron brightness is higher. Using the relative brightness B/U values and the data for the aperture angle in the crossover (the angle between the outer trajectories $2\alpha_m$), corresponding to the spatial angle $\Omega = \pi\alpha_m^2$, and maximal reachable power density in the focus $p_{\max}$ can be calculated by:

$$p_{\max} = \pi B U \alpha_m^2 = \pi \left(\frac{B}{U} \right) U^2 \alpha_m^2. \quad (19)$$

The initial chaotic velocities of emitting electrons, the aberrations, diaphragms and the collisions of the electrons of the beam with other elements of the electron optic system decrease the maximum density of the real electron beams.

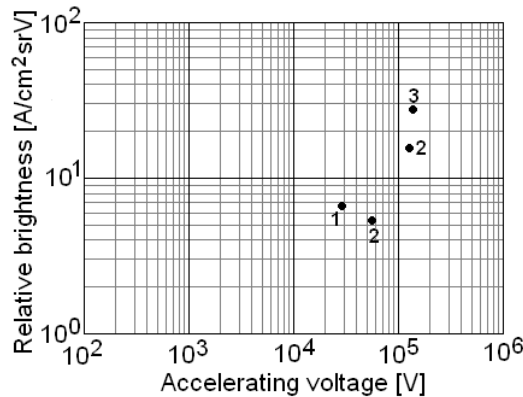


Figure 4. Data for the the relative brightness of electron optical systems for welding:

1. produced in EWI "Paton" of Ukr.AS;
2. produced in the Institute of applied physics, Dresden, Germany);
- 3- produced in Westinghouse Res. Laboratories, USA;

Effects of the Space Charge in the Intensive-Electron Beams

Intensive electron beams are those, in which the beam electrons have group behavior due to the perceptible interaction forces between them. The behavior of the electrons, moving in such an electron beam with high density of the particles in it, is defined to a considerable extent by the electrostatic interaction forces between them. The negative space charge influences are demonstrate mainly as a) emission of the current by a virtual cathode (current limited by the space charge) , and b) extension of the cross-section of the intensive electron beam.

With a big increase of the density of the particles in one unit volume of the beam, the energy distribution of the beam is changed due to two body interaction between neighboring electrons.

The particle's own electric field is not the only thing that affects the characteristics of the beam. Under certain circumstances (space charge compensation or relativistic electron velocities) and electrons' own magnetic field affects them. In presence of ionized particles from the residual gases or the vapors of the processed material in the technological vacuum chamber, wave movement of the electrons, plasma oscillations and beam instability are possible.

a) Current density, voltage and distance (cathode-anode) relation and limitations of the beam current by the beam space charge

The distribution of the electricity potential U in an intensive (dense) beam defines the velocity and the direction of movement of each electron, but at the same time depends on the space distribution of charges in the beam region. On account of this, instead of the Laplace equation, which is valid for beams with low density of electrons, here the distribution of the potentials is described by Poisson equation:

$$\nabla^2 U = -\frac{\rho}{\epsilon_0} . \quad (20)$$

Here ∇^2 is the Laplace differential operator, ϵ_0 is the dielectric constant of the environment and ρ is the density of the space charge. The vector of the current density $\vec{j}$ is connected with ρ and the velocity of the electrons $\vec{V}$ by:

$$\vec{j} = \rho \vec{V} , \quad (21)$$

which in the case of electrons is:

$$\vec{j} = -\rho \vec{V} . \quad (22)$$

Two other relations are also valid - the continuity equation and the conservation of energy law (the collisions between the particles of the beam and of the residual gases are neglected):

$$\text{div } \mathbf{j} = 0, \quad (23)$$

$$eU = \frac{mV^2}{2}. \quad (24)$$

Here e and m are the charge and the mass of the electrons, correspondingly. In such way for the distribution of the potential in intensive electron beams is obtained:

$$\nabla^2 U = \frac{1}{\epsilon_0} \left(\frac{m}{2e} \right)^{1/2} \frac{j}{U^{1/2}}. \quad (25)$$

Most strong influence has the space charge of electrons in the near-cathode area in all electron optical systems due to their slow motion. In the cases, when the cathode emits enough big quantity of electrons, the current is limited by their space charge. Equation (25) is easily integrated under the assumption for linear and laminar trajectories of mono-energetic beam of electrons, i.e. neglecting their initial velocities in flat parallel, coaxial cylindrical or spherical structure. For flat cathode and anode, after integration of eq. (25), the density of the current of the cathode, limited by the beam space charge is:

$$j = \frac{4}{9} \left(\frac{2e}{m} \right)^{1/2} \epsilon_0 \frac{U^{3/2}}{z^2}, \quad (26)$$

where U is the potential on a distance z from the emitting surface of the cathode. In this way, at distance $z=d$ between the electrodes and anode voltage U_a , the equation (26), known as Child-Langmuir equation or 3/2 power law, becomes:

$$j = 2,33 \cdot 10^{-6} \frac{U_a^{3/2}}{d^2}. \quad (27)$$

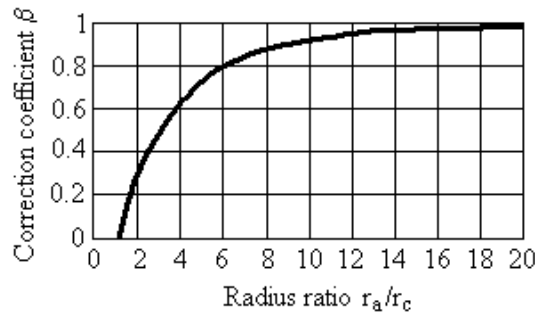


Figure 5. Correction coefficient β for a cylindrical coaxial diode as a function of r_a and r_c

In the cases of cylindrical and spherical two-electrode systems, as well as multi-electrode systems, the coefficient $2,33 \cdot 10^{-1}$ changes. For example, for cylindrical construction with length 1 m, from coaxial anode, including the cathode, the coefficient is $2,33 \cdot 10^{-6} \beta^2$ (when defining the density of the current on the anode). Here β is Langmuir correction coefficient, which is a function from the ratio between the anode radius r_a and the cathode radius r_c (Figure 5). From the figure it is seen that with the decrease of the ratio r_a/r_c the density of the current increases. At constant ratio between these radii with the decrease of r_a the intensity of the field in front of the cathode increases, which leads to considerable increase of the current, obtained from the cathode by such construction.

b) *Perveance*

The characteristic conductivity p , called *perveance*, is defined as:

$$p = \frac{I_0}{U^{3/2}}, \quad (28)$$

where I_0 is the current of the electron beam (in axially symmetric beam with radius r_0 and current density j), $I_0 = \pi r_0^2 j$. This characteristic is a measure for the influence of the space-charge on the properties of the beam. The experimental investigation and computer calculations of electron beams shows that the space charge influences the electron trajectories in good vacuum conditions at values of perveance $p > 10^{-7} \text{AV}^{-3/2}$, and that value of the perveance can be accepted as the limit between the intensive electron beams and the beams with low density of electrons. In the nowadays technology installation for welding the beam perveance values lay between $p = 10^{-8} \text{AV}^{-3/2}$ and $p = 2 \cdot 10^{-5} \text{AV}^{-3/2}$ (as example, a typical perveance value of EBW gun could be $5 \cdot 10^{-7} \text{AV}^{-3/2}$). Note, that there a correction of perveance value due to higher pressure in the draft space and the action of the effect of compensation of negative space of beam electrons by generating positive ions become appreciable.

The maximum value of the perveance, and consequently of the beam current, which can be obtained after the beam formation, is also limited by the space charge of the beam electrons. Due to the negative charge of beam electrons the potential in the space, occupied by the beam, decreases. For example, in unlimitedly wide electron beam going along the axis between two perpendicular to this axis equi-potential planes, situated on a distance l from each other, the potential distribution $U(z)$ has minimum in the middle between these planes. From integrating eq. (26) follows that with the increase of the current density the value of the potential in the minimum decreases, reaching $U_a/3$ for $j l^2 U_a^{-3/2} = 18,6 \cdot 10^{-6} [\text{AV}^{-3/2}]$. Further increasing of the current density leads to a jump of the potential in the middle point from the initial value to value, equal to 0, i.e. a virtual cathode is formed. This abrupt decrease of the potential is physically connected with slowing down of the electrons and considerable increase of the space charge. That is why with the decrease of the current density the potential in the minimum stays equal to zero until current densities corresponding to $j l^2 U_a^{-3/2} = 9,3 \cdot 10^{-6} [\text{AV}^{-3/2}]$ are reached, then the potential in the middle between the equi-potential planes jumps to $0,75U$ and the normal current flow is restored.

In the case of limited cylindrical electron beam, fully filling metal tube with potential U_a , the maximum value of the perveance is $32,4 \cdot 10^{-6} [AV^{-3/2}]$. In this case, the potential along the axis of the tube decreases to $U_a/3$. Near the axis of such a beam the electrons are moving slowly, the space charge increases, and the potential abruptly decreases. That is why the current density in the border part increases, the potential decreases, and the current flow is variable. The distribution of electron according their velocities in real beams leads to smoother transition of the beam to this unstable state. Characteristics of the different types of configuration of electron optical systems affect these two values of the beam perveance (the first - described unstable and gradually decreasing current flow and the second, where the normal flow is gradually restored).

c) Extension of the beam wide, due to the space charge of the electrons

Another (second) very important effect of the space charge is the action of the electrostatic repulsion forces between the beam electrons. They lead to difficulties in the focusing and to a widening of the beam cross section. The equation describing the movement of the electron in radial direction is:

$$m \frac{d^2 r}{dt^2} = -eE_r. \quad (29)$$

Here E_r is the radial intensity of the electric field created by the volumetric charge. Let us assume that outer accelerating, focusing and deflecting electric and magnetic fields are missing. Applying Ostrogradski-Gauss theorem for the field intensity vector flow through a cylinder with radius r , situated co-axially with the beam, and eq. (22), for the radial force is obtained:

$$F_{re} = -eE_r = -\frac{eI_0 r}{2\pi r_0^2 \epsilon_0 \sqrt{\frac{2e}{m} U_a}}. \quad (30)$$

Here r_0 is the radius to the border trajectories.

Differentiating by z in eq. (29), using $\frac{d}{dt} = \frac{dz}{dt} \cdot \frac{d}{dz}$, $\frac{dz}{dt} = V_z$ and substituting F_{re} with eq. (30), the boundary electron trajectory equation becomes:

$$\frac{d^2 r_0}{dz^2} = \frac{I_0}{4\pi \epsilon_0 r_0 \left(\frac{2e}{m}\right)^{1/2} U_a^{3/2}} = \frac{p}{kr_0}. \quad (31)$$

Again the importance of the perveance as a characteristic of the space charge in the beam is clear. Here $k=6,6 \cdot 10^{-4} [AV^{-3/2}]$. If the extending of the beam is limited by $\Delta = r - r_0$, which

are small compared to r_o , then r_o in the right part of eq. (31) can be accepted as constant and after integration the following equation is obtained:

$$r_o = a_{\min} + \frac{1}{2} \frac{p}{ka_{\min}} z^2. \quad (32)$$

Here $a_{\min}$ is the minimal diameter of the beam. If the perveance $p = 10^{-8} [AV^{-3/2}]$, the expanding Δ is not more than 1% from the length z of the beam, if the radius of the beam does not exceed 0,77 mm.

More precise integration of eq. (31) is proposed by Glazer.

It gives the universal relationship between the dimensionless radius $r_o/a_{\min}$ and the parameter

$$Z = 174 \frac{\sqrt{p}}{a_{\min}} z.$$

This relationship is shown on Figure 6. Here $a_{\min}$ is defined by:

$$a_{z_0} = a_{\min} \exp\left(\frac{k\alpha_o^2}{2p}\right), \quad (33)$$

where $\alpha_o = \left(-\frac{dr_o}{dz}\right)_{z=z_0}$ is the initial angle of shrinkage of the border electron trajectory, a_{z_0}

- the initial radius of the beam. When there is initially expanding beam, α_o is negative.

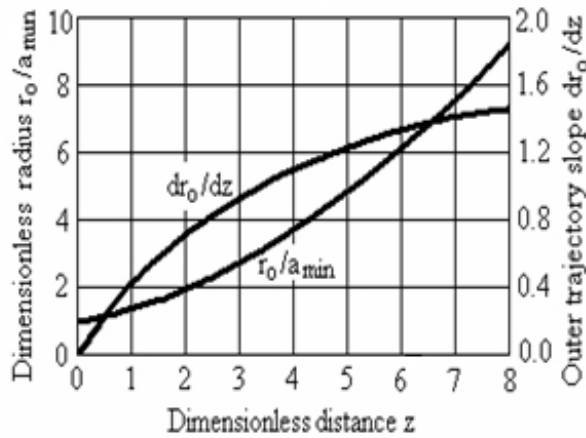


Figure 6. Universal relationships between the dimensionless radius $r_o/a_{\min}$, the angle of the slope $\partial r_o/\partial z$ and the dimensionless distance along the axis Z , characterizing the border trajectories in axially symmetrical electron beam

Compensation of the Space Charge of an Electron Beam with Ions. Magnetic-Ion and Ion Self-Focusing of Intensive Electron Beams

Besides their own electric field the moving electrons create also magnetic field. According Bio - Savar law, the magnetic induction B of the surrounding surface of a cylindrical beam with radius r_0 can be defined by:

$$B = -\frac{\mu_0 I_0}{2\pi r_0} \quad (34)$$

and the radial force influencing on the boundary electron towards the axis of the beam is:

$$F_{\text{m}} = -\frac{e\mu_0 I_0}{2\pi r_0}. \quad (35)$$

The summary radial force, which is a result of the mutual electrostatic repulsion of the electrons and the magnetic attraction of the lines of the current, is obtained by summing eq. (30) and eq. (35):

$$F_{r\Sigma} = \frac{eI_0}{2\pi r_0} \left(\frac{1}{\varepsilon_0 V_z} - \mu_0 V_z \right). \quad (36)$$

Keeping in view that ε_0 and μ_0 are connected with the ratio:

$$\varepsilon_0 \mu_0 = \frac{1}{C^2},$$

eq. (36) can be written as:

$$F_{r\Sigma} = \frac{eI_0}{2\pi \varepsilon_0 V_z r_0} \left(1 - \frac{V_z^2}{C^2} \right). \quad (37)$$

When $V_z \ll C$, the magnetic radial force is negligible and the action of their own magnetic field must be accounted only for relative electrons.

In the case of partial compensation of the beam space charge of the electrons with positive ions, created by the electron beam or imported from the outside, with f can be defined the relative space charge of the compensating ions:

$$f = \frac{\rho_{\text{ions}}}{\rho_{\text{electrons}}}. \quad (38)$$

Then the overall radial force, influencing the boundary electron is:

$$m \frac{d^2 r}{dt^2} = \frac{I_o}{4\pi\epsilon_o m \left(2 \frac{e}{m}\right)^{1/2} U_o^{3/2} r_o} \left(1 - f - \frac{V^2}{C^2}\right) \quad (39)$$

and the trajectory of the boundary electron:

$$\frac{d^2 r}{dz^2} = \frac{I_o}{4\pi\epsilon_o m \left(2 \frac{e}{m}\right)^{1/2} U_o^{3/2} r_o} \left(1 - f - \frac{V^2}{C^2}\right). \quad (40)$$

When $f < 1$, the influence of the partial compensating influence of ions is accounted and when $f > 1$ (overcompensated space charge) the effect of ion self-focusing of the beam by the positive ions, situated in the volume of the beam, is observed. In the case of $f = 1$ (full compensation) there is magnetic-ion self-focusing, which is a result of the combined influence of the ion compensation and the magnetic pinch-effect.

The radial distribution of the potential $U(r)$ for the ideal case of a beam with uniformly distributed volumetric charge, for which the radial forces are defined, as well as the case of a real electron beam are shown on Figure 7. It can be noted that as a result of partial or full neutralization the boundary current increases $(1 - f)^{-1}$ times. Often before reaching this limit other effects appear, for example plasma electron-ion oscillations and instabilities, which also define the limit value of the current.

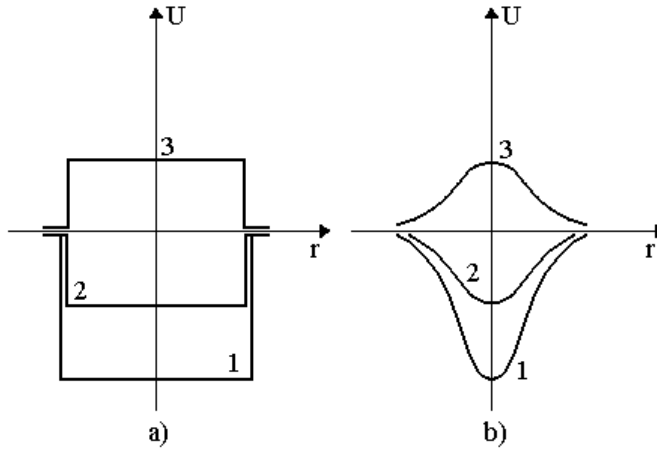


Figure 7. Distribution of the potential on radial direction of the cross-section of the electron beam: (a).Uniform current distribution along the cross-section of the beam; b) Gaussian distribution of the current density. There:

(b).1-intensive electron beam in ultra high vacuum; 2-partial ion compensation of the space charge; 3-overcompensation of the negative space charge of the electrons by the created in the transition zone ions

Generalized Influence of the Emittance, the Space Charge and Its Ion Neutralization upon the Configuration of an Electron Beam without Aberrations

Let an electron beam passes through a very small cross-section in an area without outer electric and magnetic fields (Figure 8). It is assumed that the influence of the space charge of the electrons of the beam and the included in it ions, as well as their own magnetic field is negligibly small. Because of this the shown trajectories of the separate electrons are straight lines. The beam is non-laminar i.e. its emittance is $\varepsilon \neq 0$. Some typical trajectories are shown on the diagrams of the emittance in the phase plane rr' having reference to some cross-sections. It is known that the points lying on an ellipse in a cross-section z , lie on an ellipse with the same area in the rest cross-sections. The orientation of the axis of the ellipses correspond of shrinking or expanding beam as it is seen from the diagrams related with the cross-sections I, II, III and IV. For practical purposes the boundary trajectory (drawn with dashed line) is important. The equation of this trajectory is an equation of a hyperbola with semi-width $a_{\min}$ and asymptotic angle $\varepsilon/a_{\min}$:

$$\frac{d^2 a}{dz^2} - \frac{\varepsilon^2}{a^3} = 1. \quad (41)$$

Assuming uniform distribution along the cross-section of the beam of the space charge of the electrons and partially compensating them ions, caught in the potential minimum is the beam space and in presence of outer axially symmetrical electric field, the equation of one paraxial boundary trajectory of the electron beam is:

$$a'' + a' \frac{U'}{2U} - \frac{U''}{4U} a - \frac{\varepsilon^2}{a^3} - \frac{\left(1 - f - \frac{V_z^2}{C^2}\right) I_0 m^{1/2}}{4\sqrt{2\pi\epsilon_0} U^{3/2} e^{1/2}} \cdot \frac{1}{a} = 0. \quad (42)$$

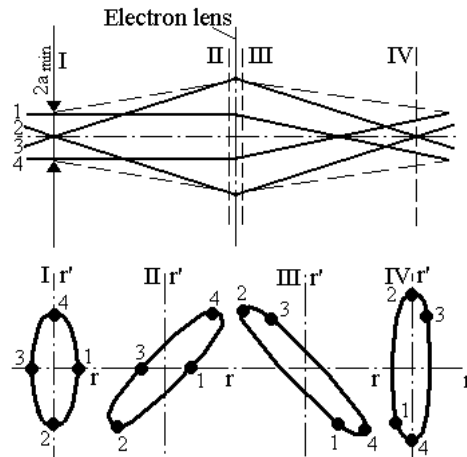


Figure 8. Trajectories and diagrams of the emittance in a non-laminar electron beam moving through a space without outer electric and magnetic fields

Here the indexes ' and '' are signed the operators $\frac{d}{dz}$ and $\frac{d^2}{dz^2}$. The first and the last term form the equation of expanding beam in a free of fields area. If a' and a'' are equal to 0, the Child-Langmuire law eq. (27) is obtained with potential $U \sim Z^{4/3}$. In order to define if the emittance or the volumetric charge prevail as a factor controlling the behavior of the beam with radius a , the forth and the fifth terms in eq. (42) are compared:

$$\left(\frac{m^{1/2}}{4\sqrt{2}\pi e^{1/2}\epsilon_0} \right) \left(1 - f - \frac{V_z^2}{C^2} \right) pa^2 \geq \epsilon^2. \quad (43)$$

If the dimension of ϵ is in [m.rad] and that of a is in [m], the numerical value of the constant in the first brackets is $1,5 \cdot 10^3$. Then for a current of 0,5 A, acceleration voltage $30 \cdot 10^3$ V and $f=0$, the emittance prevails at $1,2 \cdot 10^{-3}$, i.e. if $a > 80\epsilon$ everywhere in the beam the space charge is the main limitation of the minimal cross-section of the beam. In the cases when $a < 80\epsilon$, the limitation factor is the emittance. In nearly fully compensated beams ($f \approx 1$) the emittance is the main limitation for reaching high density, until the processes of collision, expanding of the energy distribution of the electrons, non-homogeneities and aberrations make its usage for characterization of the beam impossible. The ions in not fully compensated intensive electron beam are in a potential gap with depth proportional to the perveance. They oscillate and plasma oscillations and instabilities are possible to appear. The non-homogeneities of the cathode emission, the aberrations and other nonlinear effects lead to a loss of the beam structure, described by paraxial or other idealized equations. Further description of the electron beam can be made statistically, using as characteristic of the cross-moving of the electrons the temperature $T_e \approx T_c \cdot K$, where K is the compression by area of the beam, T_c is the temperature of the cathode, i.e. during focusing of the beam the electron temperature T_c increases and the electrons move with velocities stronger inclined towards the axis z . The generalized description of such beams is a difficult task. Only analyses of some special cases are known.

Electron-Optical Aberrations

Often in the electron optics the properties of the electron beams are analyzed through the behavior of separate electrons in accelerating and focusing electric and magnetic fields. Theoretical expressions exist allowing if the field distribution is known, to find the trajectory of the electron. Or the opposite task - to find the field necessary to ensure of a definite form of the electron beam. Widely used approximation in electron optics is the presentation of narrow near-axis paraxial beam. If a basic trajectory is given and the distribution of the components of the electric or the magnetic fields along its length is known, it is possible the neighboring trajectories to be found. The removing of the electrons from this basic trajectory (beam axis) and the angle between the axis and the calculated trajectory are accepted to be small. The simplest but met in almost all electron beam devices is the case of straight-lined axis and axis-symmetrical electric field. Usually the potential distribution along the axis of symmetry is

known (and most simple to define). In cylindrical coordinate system (z, r, φ) , the distribution of the potential in near axis area can be presented through the value of the potential of the axis $U_0(z)$. Due to the absence of relationship between the potential and the angle coordinate φ and the volume charge after applying Laplace equation for the potential of axi-symmetrical field is obtained the equation:

$$U(z,r)=U_0-\frac{1}{4}U_0''(z)r^2+\frac{1}{64}U_0^{IV}(z)r^4-....=(-1)^k\sum_{k=0}^{\infty}\frac{U_0^{2k}}{(k!)^2}\left(\frac{r}{2}\right)^{2k}. \quad (44)$$

With the indexes II, IV and $2k$ are signed the corresponding derivatives of the potential by z . The terms with odd powers in the series in eq. (44) are missing due to the equality of the potential in symmetric by the axis points. When the electron moves near the axis z , it is assumed that the axial ingredient of the field does not depend on the distance to the axis r , and the radial ingredient is proportional to r , i.e. only the lowest powers in the series in eq. (44) are used. The velocity of the electrons in the narrow near axis beam is defined by:

$$V \approx V_z = \sqrt{\frac{2e}{m}U_0(z)}. \quad (45)$$

The movement of the electron in radial direction is defined by:

$$m\frac{d^2r}{dt^2} = -\frac{1}{2}eU_0''(z)r. \quad (46)$$

After the elimination of the time t from eq.(45) and eq.(46) the trajectory of a paraxial electron at non-relative energies is described by the differential equation:

$$\frac{d^2r}{dz^2} + \frac{U_0'(z)}{2U_0(z)}\frac{dr}{dz} + \frac{U_0''(z)}{4U_0(z)}r = 0. \quad (47)$$

Since in (47) the charge and the mass of the particle are missing, the trajectory of each of the charged particles in the axi-symmetrical electrostatic field is equal. The difference is in the time of movement. The equation is uniform toward the potential, and that is why the simultaneous change of the potential in all the points of the field, the trajectory does not change.

The solution of eq.(47) is found as a sum of two partial linearly independent solutions:

$$r(z) = C_1r_1(z) + C_2r_2(z), \quad (48)$$

where C_1 and C_2 are constants, defined by the initial conditions. It is obvious, that if the field is not homogeneous at $U_0''(z) > 0$, it is possible the first partial solution to be 0 twice, i.e.:

$$r_1(z_A) = r_1(z_B) = 0. \quad (49)$$

At $C_2=0$ and fulfilled eq. (49) and eq. (48) give a group of trajectories with beginning at point $A(z_A,0)$, crossing in point $B(z_B,0)$ again on the axis, i.e. B is electron-optical image of point A . If $C_2 \neq 0$, but eq. (49) is fulfilled, all the trajectories at given C_2 and different C_1 go through points S and I (Figure 9), which do not lie on the axis. Correspondingly the point source $S[z_A, C_2, r_2(z_A)]$ is projected in the point electron-optical image $I[z_B, C_2, r_2(z_B)]$. Consequently, every non-uniform axially-symmetrical electrostatic field, in which $U_0''(z) > 0$ behaves like collector electronic lens. Analogous consideration is possible also for the axial-symmetrical magnetic field. The basic difference is in the fact that the magnetic field obtains azimuth velocity and the image is twisted at definite degree toward the object. The movement of the electron in axial-symmetrical magnetic field is described by the following system of differential equations:

$$\frac{d^2 r}{dz^2} = -\frac{e}{8mU_0(z)} B_0^2(z) r, \quad (50)$$

$$\frac{d\phi}{dz} = \sqrt{\frac{e}{8mU_0(z)}} B_0(z). \quad (51)$$

The angle of twisting depends on the direction of movement of the particle that is why the trajectories even for one and the same particles are irreversible. If there is a change of $U_0(z)$ n times, $B_0(z)$ must change correspondingly $n^{1/2}$ times in order to keep the trajectories the same. $U_0(z)$ represents the energy of the electron, i.e. the accelerating difference in potentials, but not the value of the electric potential in a corresponding point z . The analysis of the eq. (50) and eq. (51) shows, that non-uniform axi-symmetrical magnetic field in the near-axis area behaves like electronic lens. The short axi-symmetrical magnetic field always performs the role of collector lens, because in eq. (50) $B_0(z)$ is raised to the second power.

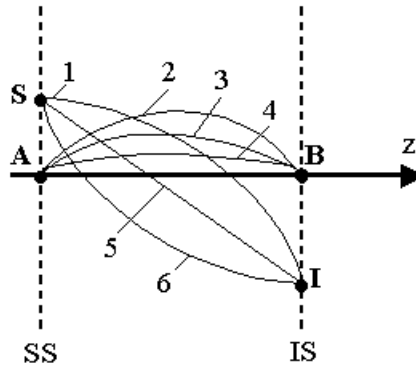


Figure 9. Electron-optical images I and B of points A and S . SS -Sample Surface; IS -Image Surface

Trajectories: 1- $C_1 r_1(z) + C_2 r_2(z)$; 2- $C_1^I r_1(z)$; 3- $r_1(z)$; 4- $C_1^{II} r_1(z)$; 5- $C_2 r_2(z)$; 6- $C_1^{II} r_1(z) + C_2 r_2(z)$

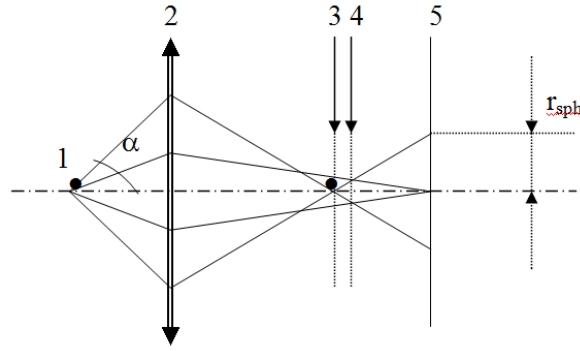


Figure 10. Trajectories of the electrons, explaining the appearance of spherical aberration: 1-source of electrons; 2-electron lens plane; 3-focusing plane of the outer (in the area of the lens) electrons; 4-minimum cross-section plane; 5-paraxial image plane

Condition for obtaining an ideal image in axi-symmetrical electric and magnetic field is the proportionality of the change of the angle of the slope of the trajectory raised to the first power from the radius. This condition is fulfilled only for the near-axis electrons. The real beams do not fulfill the requirements for being paraxial. Then in the equations are included the terms, containing the ingredients of the field of higher order. The electron-optical images are no longer ideal and become unclear. The deviations of the real image from the ideal (paraxial) image are called *aberrations*. When calculating of real electron-optical systems, usually are taken into account the aberrations of third order, i.e. those which are imported by the additional addends in the differential equations of the trajectories of terms, including r^3 , $r^2(dr/dz)$, $r(dr/dz)^2$ and $(dr/dz)^3$. There are several types of aberrations of the electron lens.

Spherical aberration. It appears due to the electrons, which after passing the outer part of the lens deviate stronger and cross the axis before the plane of the paraxial image (Figure 10). In this plane instead of a point appears a sphere of deviation with radius:

$$r_{\text{sph}} = \frac{1}{2} C_{\text{sph}} \alpha^3. \quad (52)$$

Here α is half of the angle at the apex of the cone, formed by the outermost trajectories of the electrons, forming the image, and C_{sph} - coefficient of spherical aberration of the lens. Usually C_{sph} is the product of a dimensionless coefficient K and the focus distance. K depends on the lens geometry. Lenses with short focus distance have smaller aberration. The spherical aberration is the basic type of aberration. It is essentially irremovable in real electric or magnetic lenses and it is impossible to remove from further electronic optical influence. That is why it is important to design of elements with minimum spherical aberration.

Astigmatism. This type of geometric aberration is caused by the beams, coming out of a point, situated remote from the electron-optical axis, pass through different parts of the electronic lens. The passing beams in the plane, in which lie the point and the axis, and those which lie in the perpendicular plane, cross at different distances from the lens. The cross-sections of the beam become elliptical with different orientation of the ellipse (Figure 11).

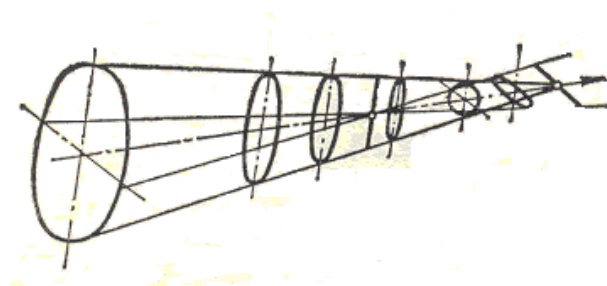


Figure 11. Scheme of electron trajectories and cross-sections of the beam, explaining astigmatism

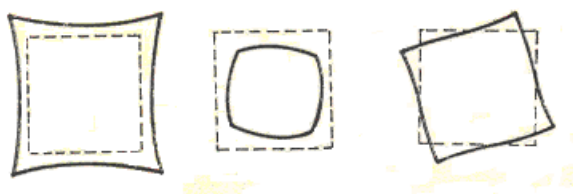


Figure 12. Twisting of the image of a square due to distortion of the electronic lens

Moreover a place can be found where the image has spherical shape (free of the astigmatism). The surface, on which these images lie is not flat and only osculate the plane of the paraxial image. Often this is considered as independent aberration, called twisting the surface of the images.

Coma. There is coma, when the image of the point not lying on the axis has comet-like shape with apex coinciding with the paraxial image.

Distortion. As the magnifying of the electron-optical system depends on the remoteness of the sample point from the axis, the image of the sample is twisted. Due to this the image of a square can look like a barrel or like a pillow (Figure 12).

Besides these aberrations the magnetic lens can have typical for them anisotropic aberrations due to the difference in the rotation of the image of differently remoted points from the axis (anisotropic coma, anisotropic astigmatism, anisotropic distortion). Generally the magnetic lenses, usually found outside the vacuum system, have bigger sizes and their aberrations are smaller.

Aberrations appear also when the axial symmetry of the fields, focusing the electron beam, is infringed. As a result even points of the sample lying on the electron-optical axis have images, which are ellipses or lines. Analogue mistakes are obtained also due to inexact assembly of the system.

Chromatic aberration. It appears due to the non-homogeneity of the velocities of the electrons of the beam. This type of aberration is observed also, when there is ideal paraxial beam. As the particles with lower velocities stay longer in the field of the electronic lens they deviate stronger. That is why the image of the point made by the slower electrons is closer than that made by the faster electrons of the beam. The effect of the pulsation of the supply pressure of magnetic electronic lenses is analogous.

The aberrations in contrast to the general analytical expressions for the trajectories of electron-optical systems are analysed for a particular system. In electron beam devices with high resolution (drilling electron devices for analysis, scanning systems for electron lithography) the aberrations are the limiting factor of the system capabilities.

Phase and Trace Volumes of the Beam and the Beam Emittance in Electron Beam Welding Machines

The process of electron beam welding is influenced by the beam energy space distribution, being a characteristic of the beam quality. Various methods for estimation of the electron beam quality were proposed. Measuring of the current distribution of powerful mono-energetic electron beams in a transverse cross section (called also the beam profile) was proposed and applied recently [11-15]. It is clear, that for prognostication of deep penetrating welding results one need from evaluation of the “parallelism” or “laminarity” of the beam (namely the angular distribution of beam particles) in the same time of evaluation the current radial distributions in the studied transverse cross sections along the beam axis. It were mention that, for description of collective behavior of the beam particles one need of a knowledge of the value of the particle density in the six-dimensional phase space (x, V_x, y, V_y, z, V_z), because t is excluded in the case of continuous electron beam. There x, y, z are coordinate axes and V_x, V_y and V_z are the respective velocity components. There z is the beam axis direction. It is important to note, that the phase volume of the beam in the 6D phase space (x, y, z, V_x, V_y, V_z) termed *6D hiper emittance*, as well as the related particle densities and/or these values in a 4D trace space ($x, y, dx/dz, dy/dz$), involving transverse coordinates and angles, are constant along the beam axis and in time, under ideal condition of a beam, particles of which are non-interacting with short – range forces. In cases of not coupled transverse dimensions is more practical to determine the projections of beam parameters in two 2D sub-planes: ($x, x' = dx/dz$) and respective ($y, y' = dy/dz$) plane. Together with the mentioned conditions - lack of collisions, which is required for conservation of volume of a non-relativistic beam phase (trace) space, is an additional requirement for excluding the frictional forces that depend on particle velocity. The thermal spread of the emitted electrons is a reason for non-zero value of the geometry emittance. Coulomb interaction lead to a “space-charge” effect causing increase of the beam phase volume and emittance; the non-linear elements of beam forming system lead to distortions and wrapping of the phase volume and a quasi-expanding of the beam effective emittance.

As was mentioned, the six-dimensional description for a beam in the drift space is usually split into two-dimensional (x, x') and (y, y') subspaces and a *geometry emittance* is defined there as the areas, occupied by all or a chosen part of the beam particles (current) in these two-dimensional spaces, dividing to π (Figure 3). For $x0x'$ plane:

$$\varepsilon^x = \frac{A^x}{\pi}, \quad (53)$$

where A^x is the area, occupied by the beam (respectively a beam part); the index x means, that parameter A and emittance are measures in the (x, x') sub-space. As example ε^x and ε^y signed

the emittances in the (x, x') and (y, y') subspaces. Conservation of ε^x and ε^y take place in the case that beam transport releases at not coupled sub-spaces, that is usual at electron beam welding optical systems. In case of characterization of part of the beam current $p = I/I_0$, where I is an investigated part of the total beam current I_0 , than a bottom index p is added to the ε^x and ε^y and ε_p^x and ε_p^y are the corresponding two-dimensional emittances.

In the case of accelerating of the electrons or at describing a relativistic beam the velocity V of beam particles is changed. At increase of longitudinal component of V , the divergence of beam gets smaller. Then the geometry emittance decreases too. A scaling velocity could be c , the speed of light in vacuum, that give a independent of beam energy emittance. So is introduced *normalized emittance*, which is invariant in the case of acceleration regions of the electrons of studied powerful beam. At assuming the relativistic Lorenz factor equal to 1 (or multiplying with him calculated value) it can be written:

$$\varepsilon_{p,n}^x = \frac{V \cdot \varepsilon_p^x}{c}. \quad (54)$$

In the case of usually assumed 2D Gaussian distribution of the beam current, the probability density N is:

$$N(x, x') = \frac{1}{2\pi\sigma_x\sigma_{x'}(1-r^2)^{1/2}} \exp - \frac{1}{2(1-r^2)} \cdot \left[\left(\frac{x}{\sigma_x} \right)^2 - 2r \left(\frac{x}{\sigma_x} \right) \left(\frac{x'}{\sigma_{x'}} \right) + \left(\frac{x'}{\sigma_{x'}} \right)^2 \right] \quad (55)$$

where $\sigma_x, \sigma_{x'}$ are the standard deviations of the particle coordinates and angles x and x' , and r is correlation between these random quantities. At $r=0$ (no correlation) the probability density N could be presented by the product of two normal distributions and the boundary of the projection of phase space on xOx' takes place of an ellipse in a canonical position (namely its main axes coincide with x and x' axes). In the case of $r=1$ the ellipse becomes a straight line $x'=(\sigma_x/\sigma_{x'})x$.

The use of 2D normal distribution (55) leads to elliptical shapes of the boundaries of the particle distribution diagram, given in the xOx' plane that coinciding to the elliptical trajectories of particles in the phase plane.

The equation of emittance ellipses could be written as:

$$\gamma x^2 + 2\alpha x x' + \beta x'^2 = \varepsilon_p \quad (56)$$

There ε_p is the emittance for part p of the beam current, containing in respective ellipse; α, β and γ are so called Twiss (or Courant-Snyder) parameters that obey:

$$\beta \cdot \gamma - \alpha^2 = 1, \quad (57)$$

and are given on Figure 13. Note, that (57) is just the geometrical properties of an ellipse.

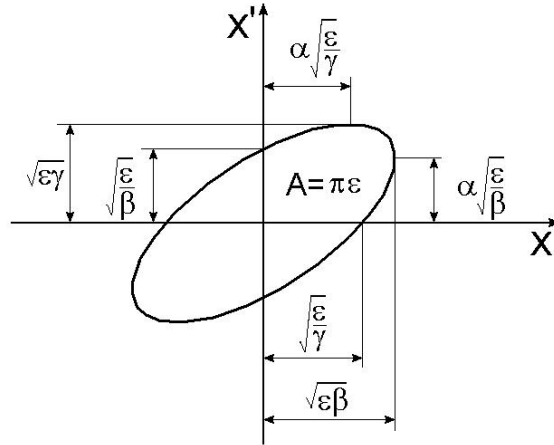


Figure 13. Determination of emittance ellipse by Twiss parameters

Coefficient (or Twiss parameter) β characterizes changes of the beam envelope. Its definition could be written in terms of second order moments of distribution function:

$$\beta_x = \frac{\langle x^2 \rangle}{\epsilon_x}. \quad (58)$$

There the brackets $\langle \rangle$ means an average value, performed over the beam particles distribution.

Respectively γ is a measure of the average declination of electron trajectories from the beam axis:

$$\gamma_x = \frac{\langle x'^2 \rangle}{\epsilon_x}, \quad (59)$$

and the Twiss coefficient α is determined as:

$$\alpha_x = \frac{\langle x, x' \rangle}{\epsilon_x}. \quad (60)$$

In the case of a more complicated beam distribution the area, occupied by particle points in x, x' or y, y' planes, could have a not easily defined shape (Figure 14). The effective root-mean-square (r.m.s.) emittance $\bar{\epsilon}$, the definition of which is based on the concept of “equivalent perfect beam”, is applicable in that case.

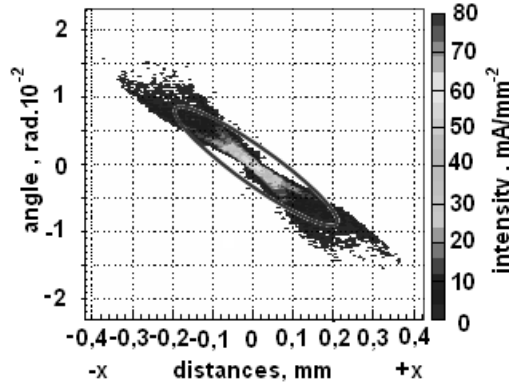


Figure 14. Effective root-mean-square (r.m.s.) emittance $\bar{\varepsilon}$ and the concept of “equivalent perfect beam”

It can be shown to be:

$$\bar{\varepsilon}_x = 4[\langle x^2 \rangle \langle x'^2 \rangle - \langle x.x' \rangle^2]^{1/2}, \quad (61)$$

This is taken as a definition of the effective r.m.s. emittance in general (at assumption to contain about 0.9 of the beam current).

The correlation coefficient r in eq.(55) could be defined as:

$$r = \frac{\langle x.x' \rangle}{\sqrt{\langle x^2 \rangle \langle x'^2 \rangle}}, \quad (62)$$

and the Gaussian (normal) distribution (55) can be rewritten as:

$$N(x,x') = \frac{\exp\left\{-\frac{\gamma x^2 + 2\alpha.x.x' + \beta.x'^2}{2\varepsilon}\right\}}{2\pi\varepsilon}. \quad (63)$$

Beam Radial Intensity Profile Monitors

The emittance of a beam is not measured directly parameter. It can be inferred by beam current profile in the transverse cross-section (radial intensity profile) and by angular distributions of beam particles in that transverse position, evaluated or measured (see below).

A beam profile monitor placed in the beam path convert the beam flux density in a measurable signal that is a function of positions towards the beam axis. A schematic presentation of radial profile monitor is shown on Figure 15.

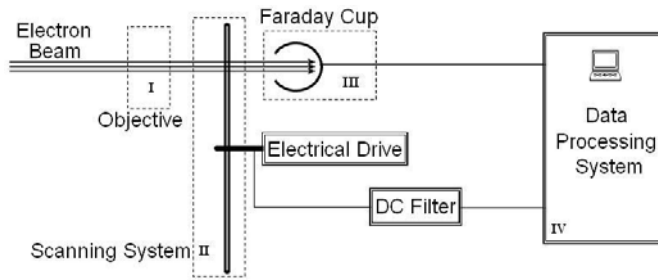


Figure 15. Block-scheme of beam current distribution (or radial profile) monitor.

There: I is objective (usually part of electron gun); II is scanning (modulation device); III is Faraday cup and IV is data processing and display system

When measuring beam profile of a intense beam (that power excess of 1kW and are going to tens or hundreds of kW), the beam has enough energy to deteriorate most sensors or current collectors, that might be placed in the beam path. So, a sampling assembly, often consisting of a scanning (rotating, moving) wire, pinhole, drum or disc containing a knife-edge or slit, permits to measure passed or absorbed part of the beam using one collector, Faraday cup or sensor, irradiated with this small beam part at any time.

An example pinhole method is shown on Figure 3. This technique is difficult for direct use in case of characterization the powerful beams, due to destroying the first screen A by intense beam heating. Various approaches and apparatus for determining of charged particles beam characteristics (beam configuration, diameter, energy peak, current density, spot size and edge width-see as example [12-17]) could be used as base for quantitative characterization of EB.

Quantitative diagnostics of beam profile in one cross-section could be done by a rotating wire device (Figure 16). This early method is simple. The device operate by scanning a thin electrically conductive wire crossing through the beam to sample the beam current and could estimate roughly the diameter of EB (the periphery of the EB current distribution). There output signal is the dependence of the wire collected current on coordinate x , coinciding with the wire movement at crossing the beam studied beam cross-section $j(x)$. In the same time instead $j(x)$ (or exactly $j_{xw, yw}$ that is the integrated value of the beam current along the wire), of interest is $j(r)$.

Analogically at use instead wire a slit the slit signal is integrated along this sampling slit and the detailed information for current density distribution in every point of the beam cross section, as that is in the case of rotating wire, is need to be calculated. To get current density in a point instead its value, integrated along that line (slit), one could do *inverse integral transformation of Abel* [20] after the assuming *axis-symmetrical beam current distribution*. This transformation is partial case of Volterra integral equation of the first kind and is typical for not correct formulated (or ill-conditioned) mathematical problems, that solution is unstable at small changes in the input data. In all cases, due to neglecting the signal at big distances from the beam axis, there is probability that a false minimum on the beam axis to be observed [18,20].

Some other reasons for errors are the beek scattered secondary electrons (and electron emission) from the heated wire or slit edges. Limitation of that method is the poor heat dissipation from the wire.

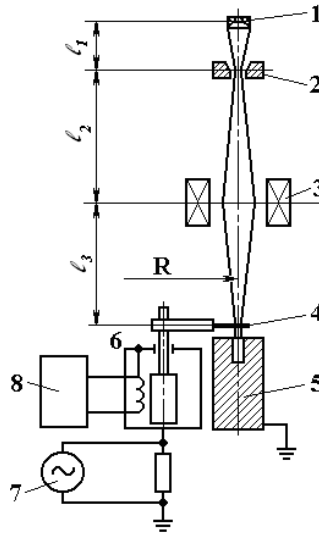


Figure 16. Scheme of an rotating wire measurement of beam profile. The signs are: 1-cathode, 2-anode, 3- focusing coil, 4-rotating wire, 5- collector, 6- electric motor, 7-oscillograph, 8-power source and control of movement

Similar difficulties exist in the analogical to mentioned method utilizing a sharp edge, where the relative movement between the beam and measuring element play the role of rotating wire.

It is interesting to mention that design of a number of EB profile measuring devices and signal formation there could be analyzed on base of the space-frequency characteristics consideration [18] . A matrix of 32x32 sufficiently short sampling impulses and transfer rate twice higher than the maximum spectrum frequency can create adequate image of the beam current distributions along any coordinate.

A new approach to use the modified rotating wire method is shown (for not very powerful beams) on Figure 17. There multi-wire sensor, consisting of thin refractory metal wires, situated on distance d one of other, rotates in the plane of studied beam cross-section. The measurement of the collected currents on every wire is executed on M steps, situated one to another on a rotating angle increment of $\Delta\theta$. Every set is a beam projection (see tomography measurement below). The measurement of projections is finished at $M \cdot \Delta\theta + 180^\circ$. From these projections a 2-D beam profile could be reconstructed by computer tomography algorithm and without difficulties of Abel transformations.

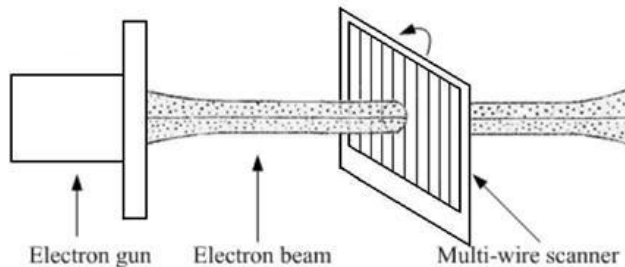


Figure 17. Rotating multi wire measurement of beam profile

A development of a modified pinhole method as general way to measure the current density distribution of the beam in a point of its cross-section and due to this to overlay the difficulties of Abel transformation is shown on Figure 18. By scanning through a rectangular raster the EB cross-section with a pinhole, done by relative movement of two slits [14]. There the beam current density distribution in studied transverse cross-section is studied using regularly spaced intervals of measurement. Signal-to-noise ratio in that device is enough high. An example of pinhole measurement of three electron beam profiles in one EBW machine, done by relative movement of two slits (shown on Figure 18a) in three cross sections of the studied beam, are given in Figure 19.

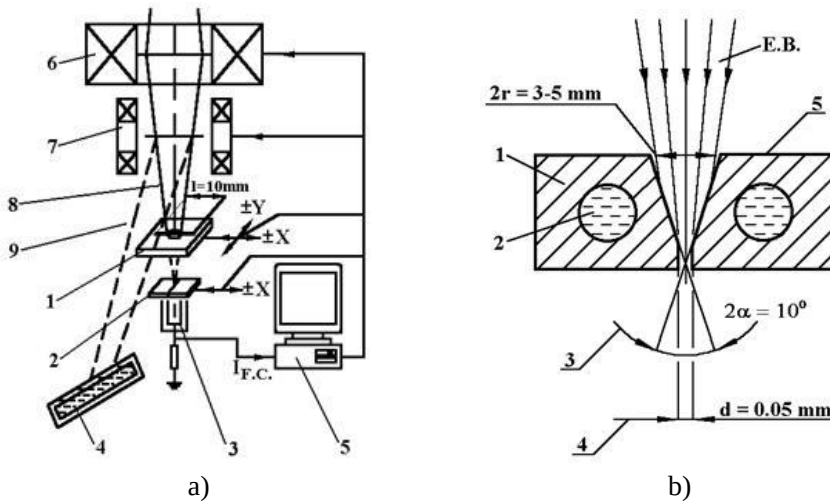


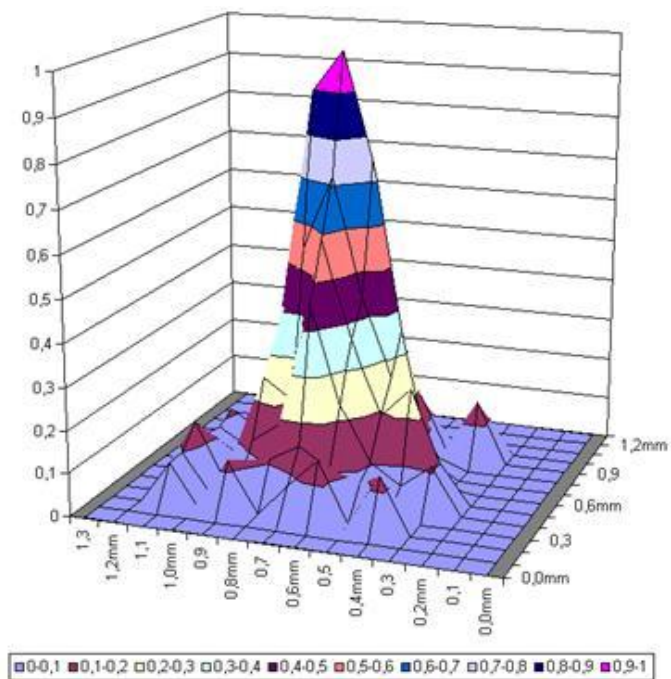
Figure 18. a) Scheme of modified pinhole beam profiler. The signs are: 1- input first water cooled plate; 2- second analyzing plate; 3- Faraday cup; 4- collector of deflected EB; 6- focusing coil; 7- deflecting coils, b) Design of measuring slit in the first water cooled plate and position of EB during the measurement

In the Figure 20 are shown the approximations of these distributions as Gaussians, need for calculation of emittances.

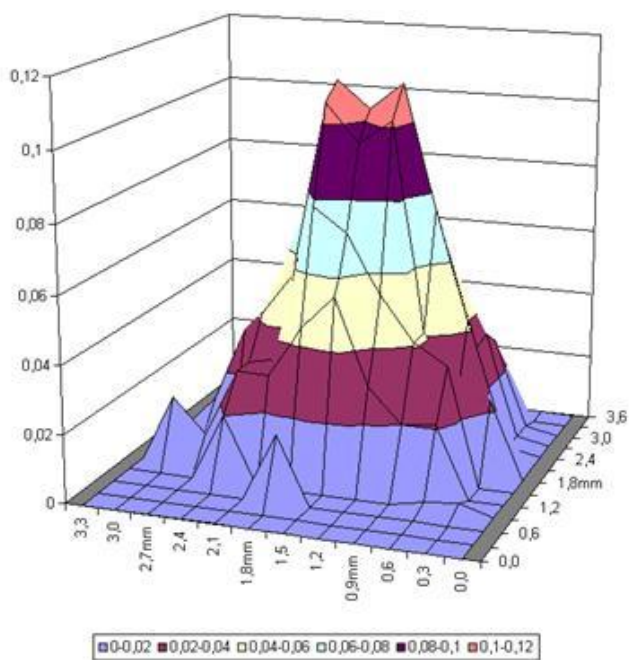
In the case of tomographic reconstruction of the beam profile [11,12,16,17] as was mentioned the expected distribution of beam current density is not need to be assumed and possible non-correctness of the beam profile analysis are waived. But in such a case more axes of the beam cut are need to be created. One example of tomography measuring approach were shown on Figure 17. Another possibility is use of rotating drum (see Figure 21) could obtain data for projections of the beam current density at up to seven different cut axes (changing slit angle toward the movement direction), but never in the cut axes coinciding with the direction of movement. Changing position of a rotating wire profiler around the beam cross-section ($\theta = \text{Var}$) needed projections (Figure 23) for tomographic reconstruction of beam current density distribution can be collected.

Another excellent proposal for use of a Faraday cup and few radial slits in a disc on which the monitored electron beam is rotated (see Figure 24) had been given by Elmer and co-workers [11,16]. This technique measures the electron beam profile by integrating the current passing along these thin slits in projections of the beam intensity, taken at equally spaced angles around beam. A non-uniform slit width of one slit is provided especially for

orientation of the sampling disc modulator towards the technology chamber. The side walls of slits are with a inclination to the vertical plane of 5° .



a)



b)

Figure 19. (Continued)

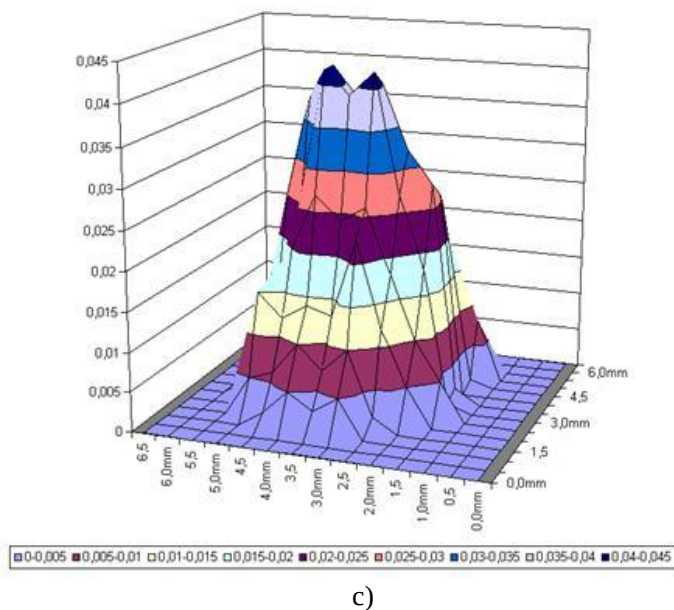


Figure 19. Experimental measured beam current distributions in different cross-sections: a) $z=320\text{mm}$; b) $z=245$ and c) $z=170\text{ mm}$

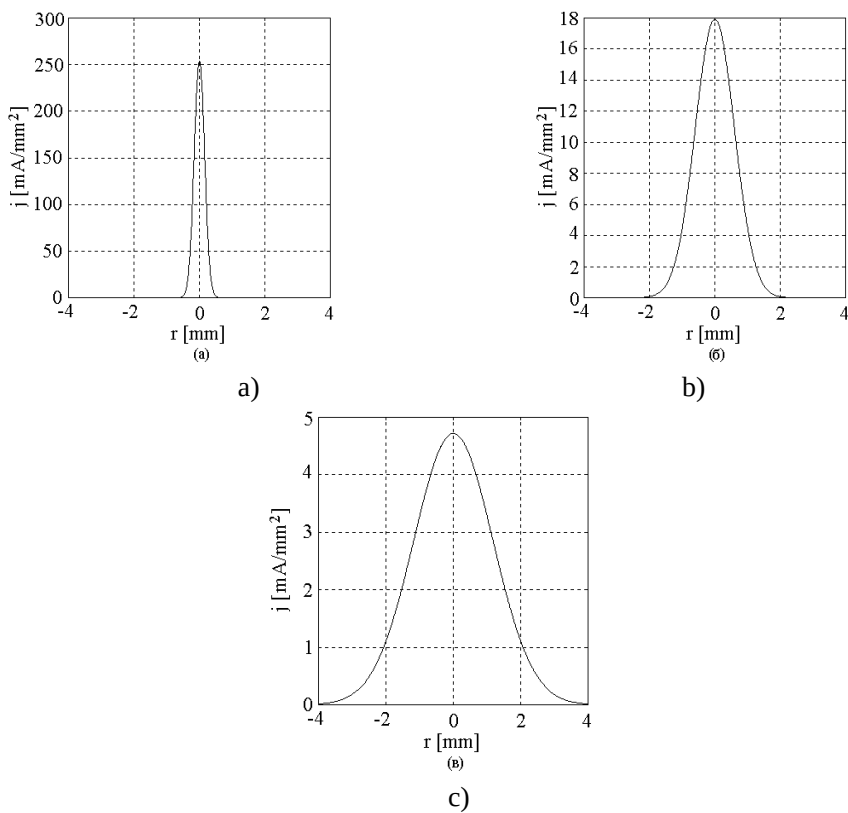


Figure 20. Approximated current distributions in the same cross-section as on Figure 19

Tomography is the technique of reconstruction a two dimensional object image from a set of its one dimensional projections, measured as an array of line integrals (or slices) of the studied object

The technique of tomography reconstruction of suitable projections is widely used in sciences, starting from medical applications and material sciences. There a Fourier transformation from real to frequency plane (Figure 25) and a consequent back Fourier transformation permits to reconstruct beam cross section current density image (beam radial intensity profile) with their asymmetry features.

On Figure 22 is shown a modified Faraday cup signed as a, b is an isolator, refractory sampling disc with radial slits is c (see Figure 24), d is the measuring set body, e is the signal output contact and f is the grounding screw; g is inner diameter of the Faraday cup. Figure 26 is presentation of the positions of the space domain points where signal is reconstructed by back Fourier transformation of frequency domain signal approximation.

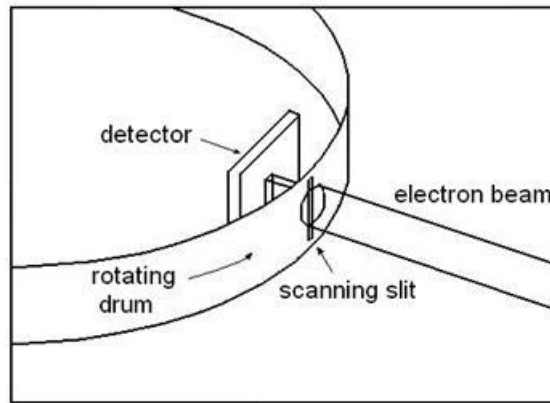


Figure 21. Measurement of a projection of beam current density distribution using rotating drum. Obtained signal in every moment is a integral of the beam current passed along slit. The projection is the sum of all integrated by slit line density signals, measured during a cut of beam

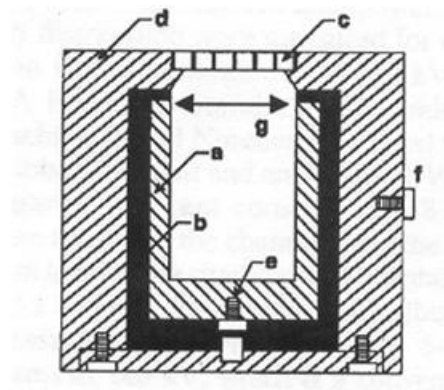


Figure 22. Set with modified Faraday cup and refractory disc, proposed from Dr.Elmer, for sampling the beam current distribution projections suitable for a tomography reconstruction of the beam profile

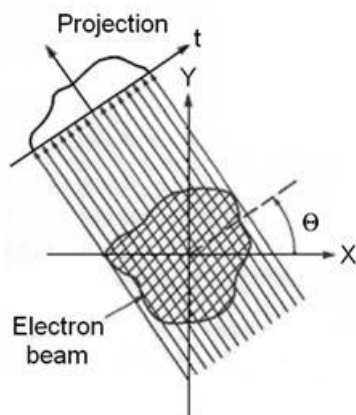


Figure 23. Scheme of measurement of projection under a direction of beam cut on angle θ . For tomography reconstruction are need a lot of such projections at various cut directions (namely θ)

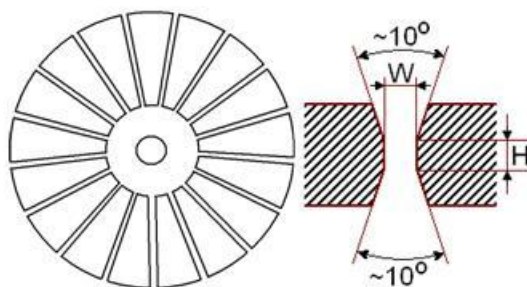


Figure 24. Tungsten disc with radial slits (see inserted cross-section too) on which beam is rotated. The part of beam passing through these slits is measured by Faraday cups

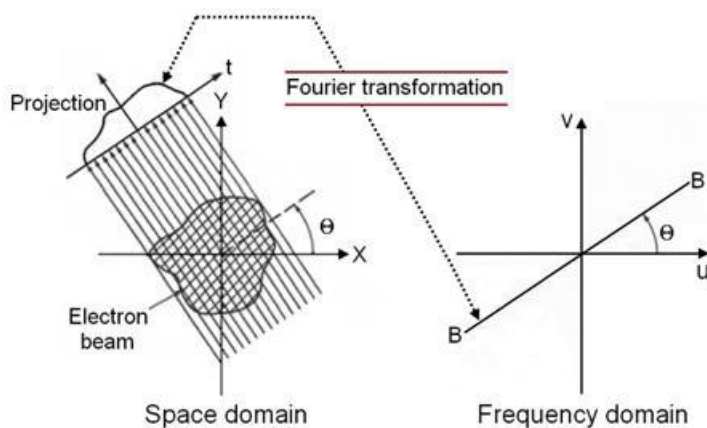


Figure 25. Schematic presentation of fast Fourier transformation of a projection of the beam particles distribution

The modified Faraday cup, proposed in [18], for the measurements of projections of beam profile, is shown on Figure 27.

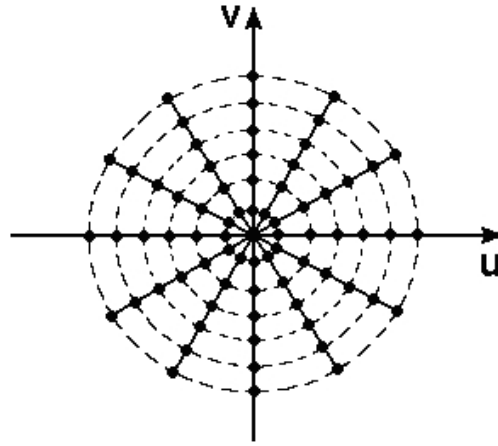


Figure 26. The obtained beam profile in frequency plane after Fourier transform of six projections

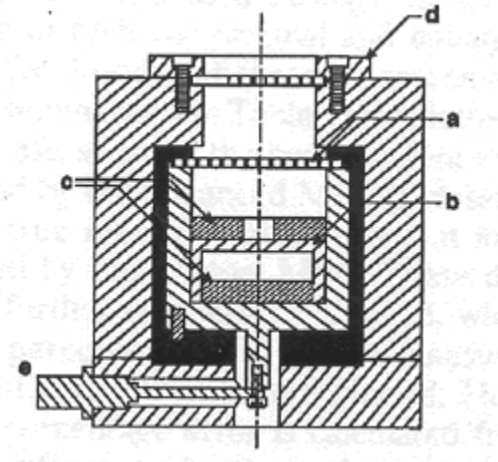


Figure 27. Modified Faraday cup: a) additional shield for back-scattered electrons with wider slits, b) second Faraday cup for calibration (measuring whole beam at its centering), c) carbon disc for minimization of back scattering electron and improved heating stability, d) clamp for pressing the measuring disc

Measurement of Angular Distribution of the Beam Particles and Calculation the Beam Emittance

The base way to measure angular distribution of beam electrons is use of two movable pin-hole plates and one collector electrode (Figure 3).

Pinhole method, shown on Figure 3, is difficult for direct use in case of characterization the powerful beams, due to destroying the first screen by beam heating. Note that as result of mention above analysis in ref. [18] one could evaluate, that about 10^6 sufficiently short sampling impulses and transfer rate twice higher than the maximum spectrum frequency can create adequate detailed image of the beam angular distributions. This means that for enough

adequate analysis a signal, collected from lot of measuring positions of both plates must be transferred and treated. This is too long for testing angular distribution in a production welding machine

More practical way for evaluating the beam angular distribution (and estimation of the beam emittance) for powerful electron beams, based on the multiple beam profile measurement, were proposed in [21-23]. In [21-23] emittance calculation by: a) the measurement of two beam profiles and a known focusing plane position or b) by three measurements of the beam current density profiles at three locations along the beam axis was proposed.

The emittance ε_p and the standard deviations are related:

$$\varepsilon_p = C \cdot \sigma_x \cdot \sigma_{x'} \quad (64)$$

where the coefficient C could be calculated as (see Figure 29):

$$C = [-2 \ln(1-p)]^{1/2} \quad (65)$$

The relations between the emittance ε_p and the product of σ_x and $\sigma_{x'}$ at a radial symmetrical beam for various beam current parts p are given in Table 2.



Figure 28. Photography of the set for measuring radial current distribution of EBW beam utilizing the method of Dr. Elmer. The tungsten sampling disc have 7 radial slits

Table 2. Relation between the values of the emittance and the part p of the beam current

p	0,63	0,78	0,86	0,99
ε_p	$2\sigma_x \cdot \sigma_{x'}$	$3\sigma_x \cdot \sigma_{x'}$	$4\sigma_x \cdot \sigma_{x'}$	$9\sigma_x \cdot \sigma_{x'}$

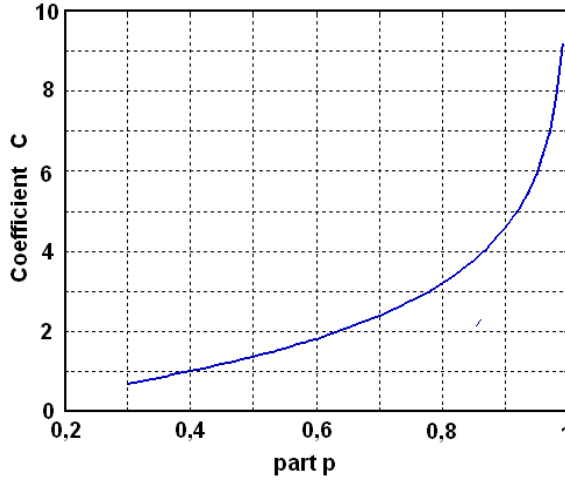


Figure 29. Plot for obtaining the coefficient C from beam current part p

The transformations of coordinate x and x' in the drift space (that is free from external to the beam forces) are given in a matrix expression as:

$$\begin{pmatrix} x \\ x' \end{pmatrix}_2 = \begin{pmatrix} 1 & L \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ x' \end{pmatrix}_1. \quad (66)$$

There index 1 stands for the cross-section at $z=z_1$ before the draft region with length L and the index 2 – at $z = z_2$.

On the base of the theorem for the dispersions of the sum of two random quantities and a zero value of the co-variance between x_0 and x_0' due to the canonic position of the emittance diagram in the cross-over image plane (called usually “focus” or “waist” of the beam) and using eq. (66) a system of three equations can be written:

$$(\sigma_{x1})^2 = (\sigma_{x0})^2 + (L_{0-1})^2 (\sigma_{x0'})^2, \quad (67)$$

$$(\sigma_{x2})^2 = (\sigma_{x0})^2 + (L_{0-2})^2 (\sigma_{x0'})^2, \quad (68)$$

$$(\sigma_{x3})^2 = (\sigma_{x0})^2 + (L_{0-3})^2 (\sigma_{x0'})^2. \quad (69)$$

There indices 0-1, 0-2 and 0-3 are respectively the differences between z of the mentioned cross-sections ($L_{0-1} + L_{1-2} = L_{0-3}$ and vice versa). At measured values of σ_{x1} , σ_{x2} and σ_{x3} and known L_{1-2} and L_{1-3} , the “focus”(or “waist”) parameters L_{0-1} , σ_{x0} and $\sigma_{x0'}$ can be found. In the case of known position of the beam “focus”(or “waist”) two equations (or measurements of the beam profile) are necessary.

The data evaluated from the beam profiles shown in Figure 19 and Figure 20 by that method are shown in Table 3. The signs are: p is the part of the beam current I_m normalized by the total beam current I_0 . The values a_r and b_r are the ellipse axis values of the respective parts, including chosen part of the beam current. Index p defines the evaluated emittance and relative brightness.

Table 3. Evaluated data of the studied EBW gun with a bolt cathode

$P=I_m/I_b$	-	0.39	0.63	0.78	0.86	0.99
K	-	1	2	3	4	9
a_r	mm	0.222	0.313	0.384	0.444	0.666
b_r	mrاد	10.92	15.4	18.9	21.84	32.76
ϵ_p	mm mrad	2.42	4.85	7.27	9.7	21.8
ϵ_{np}	m rad	1.17	2.35	3.52	4.7	10.56
$(B/U)_p$	$10^5 \text{A/m}^2 \text{rad}^2 \text{V}$	8.87	3.56	1.96	1.22	0.277

Another method for the calculation of emittance using slits and a deflected beam with a changing place of the beam “focus” (“waist”) were proposed in [18,22]. This method was applied for evaluation of emittance in x_0x' and y_0y' planes. For that aim the beam was crossing through two perpendicular slits and two measured signals of passing electrons at continuously changed focusing coil current was measured. Let see the signal use for calculation of one emittance ϵ_x .

In the investigated cross-section is situated water-cooled input plate with a narrow slit. The beam is deflected across that slit. From a previous investigation the relations between some values of the focusing coil current and the focusing length of the electron gun magnetic focusing lens f , knowing also the corresponding positions of beam “waist”(or so called “focus”) planes z_{bf1} , z_{bf2} and ... z_{bfi} are known. Please, do not mix the focusing length f of the electron lens with the distance between central plane of focusing lens and crossover image plane (namely beam “waist”, called usually also as beam “focus” plane).

The base electron lens equation is used:

$$\frac{1}{|z_{co} - z_{fl}|} + \frac{1}{|z_{bf} - z_{fl}|} = \frac{1}{f}. \quad (70)$$

There z_{co} is the cross-over place on the beam axis; z_{bf} is the place of the beam “focus” (image plane) and z_{fl} is the central plane position of the magnetic focusing lens of the electron gun (see Figure 30).

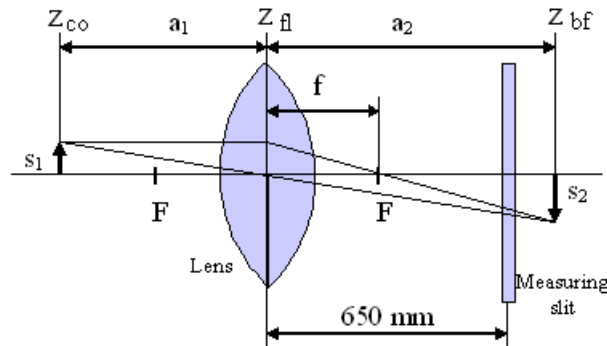


Figure 30. Measuring the beam current distribution by changing the position of the focal plane

For the calculation of the standard deviations of the normal distributions of electrons at the beam “focusing” planes (images of the cross-over) at various focusing lengths $\sigma_{i0}, \dots, \sigma_{n0}$ the coefficients of magnifications k_i are calculated by :

$$\frac{z_{fl} - z_{co}}{z_{bfi} - z_{fl}} = k_i, \quad (71)$$

and

$$\sigma_{0i} = \sigma_0 \cdot k_i. \quad (72)$$

Then, using the equation:

$$(\sigma_{xi})^2 = (\sigma_{x0i})^2 + (z_{0i} - z_0)^2 (\sigma_{x'0i})^2, \quad (73)$$

written at a condition of zero value of the co-variance between x and x' in the canonic position of the emittance diagram, one can find $\sigma_{x'0i}$ at measured σ_{xi} .

In [24] was proposed a third method of emittance observation through adding a second thin focusing lens, that transforms the angular beam distribution in radial one. The studied beam cross-section before lens is crossing by a moving slit along x . The output signal, that is a transformation of x' to x , obtained by output slit in suitable position after the lens is given on y axis of an oscilloscope (on x is given signal, produced by x movement of movable slit. The emittance diagram is observable directly on the oscilloscope screen.

In the all shortly discussed methods where a slit is applied for sampling a line integral of beam current distribution the parameters of: i) slit wide W , ii) modulator slit thickness H (in its narrower part, see Figure 24) and iii) angles between slit walls φ_{in} φ_{out} in input or output orifices of slit channel, as well as iv) the distance between two neighbor slits L_s have to be optimized for the certain value of the emittance to be measured. The following criteria have to be fulfilled for a correct emittance evaluation.

Angular acceptance of the slit must be significantly bigger than the maximal beam divergence. Distance L have to be enough big. So:

$$\varphi_{in} \geq 10^\circ; \varphi_{out} \geq 10^\circ. \quad (74)$$

$$\frac{2H\varepsilon}{W\sigma_{xi}} \ll 1; \quad \frac{\sigma_{xi}}{2} < L_s \quad (75)$$

where σ_{xi} is the beam size at the slit center .

Experimental Results and Calculation of the Current Distribution at Change of the Focus Position

The measuring device used is shown in Figure 31. During the experiments the ‘focus’ position of beam changes. Two scans are made – along X-axis and after that along Y-axis (see Figure 31). The measured current distributions represent a set of linear integrals of the current distributions along the other axis. They are presented on a single bitmap for different focus positions on Figure 32. There, each line corresponds to the integral current distributions for different cross-sections and for different focus positions (see the distance a_2 in Figure 30)

The empiric formula, which gives the connection between the distance of the focus position from the main axis of the focusing lens f and the corresponding number of the bitmap line N_L for the studied EBW gun is:

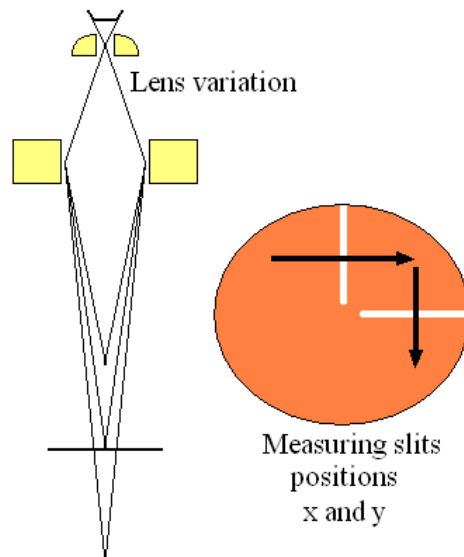


Figure 31. Experimental measuring device

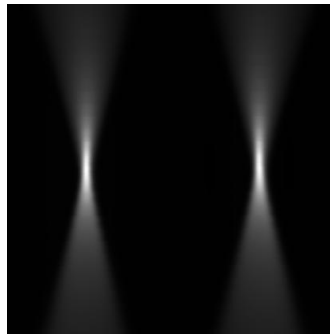


Figure 32. Two measurements of the integral current distributions along X- and Y-axes

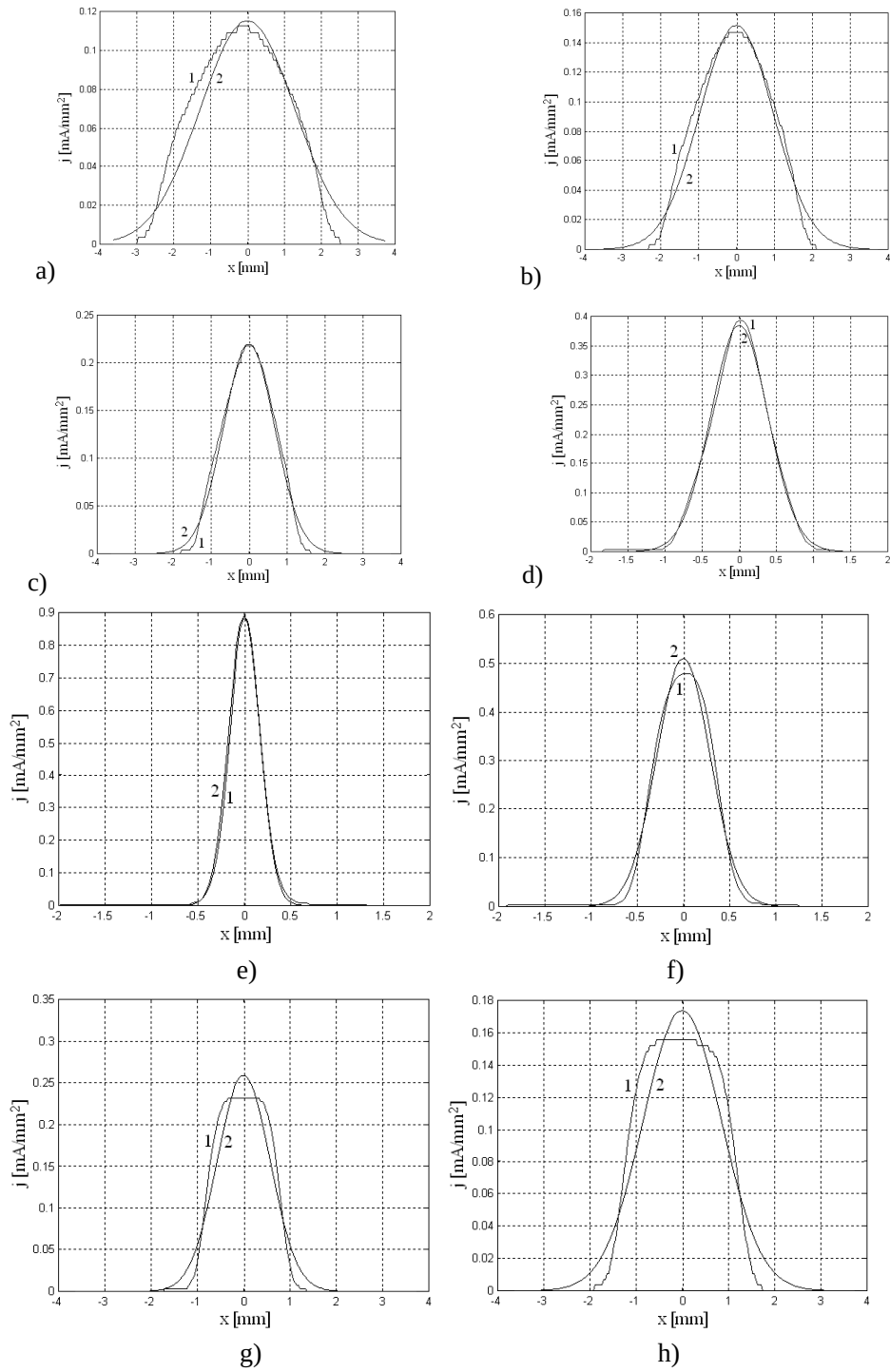


Figure 33. (Continued)

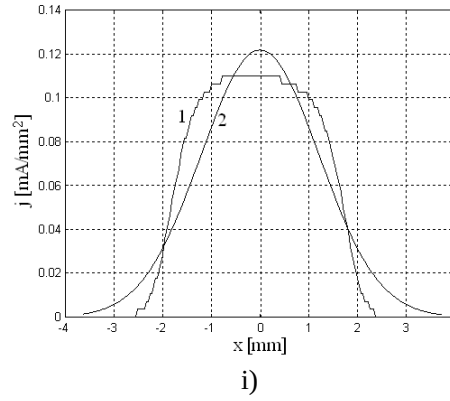


Figure 33. Experimental data (curves 1) and (curves 2) fitted to normal integral beam current distributions at different “focus” positions from X-axis scans (see Table 4)

$$f = HV / (0.1714 + HV 4.142 \cdot 10^{-6} (N_L - 200)) \quad (76)$$

where $HV = 60$ keV is the energy of the electrons, N_L is the number of the line, presenting in the bitmap, shown on Figure 32, the beam. This bitmap shows the converted into light current distribution transferred through the slit at beam scan at two slits, as this is displayed at the insertion of Figure 31. The beam current is 10 mA.

The integral current distributions in nine cross-sections from the X-axis and respectively the Y-axis distributions, corresponding to equal focus positions, are investigated.

On Figure 33 are presented the results, obtained on the base of experimental data (curves 1), from fitting the measured integral current distributions to normal distributions (curves 2) at nine “focus” plane positions from the X-axis scans. They are fitted using the least squares method. The calculated values of the focus position, the number of the bitmap line and the estimated standard deviation are given in Table 4.

It can be observed, that with the decrease of the beam diameter the accuracy of the approximation of the current distribution with a normal one increases. The deformations (deviations of measured current distributions from the normal one) are a result of aberrations and beam ion generation as well as non-uniformities in the beam transport track.

Then, using this formula the distance a_1 (Figure 30), which is constant, can be calculated from the measurement, when distance to the image of the crossover coincides with the distance to the measuring slit (the shaded one in Table 4, signed with letter ‘e’ with the smallest diameter), as:

$$a_1 = \frac{f \cdot a_2}{a_2 - f} = 746.8734 \text{ mm.} \quad (77)$$

Then the distance a_2 to the image s_2 (Figure 30) for different focal length of the lens can be calculated by (see Table 4):

$$a_2 = \frac{f \cdot a_1}{a_1 - f} \quad (78)$$

One can write the following relations based on the optical theory (Figure 30):

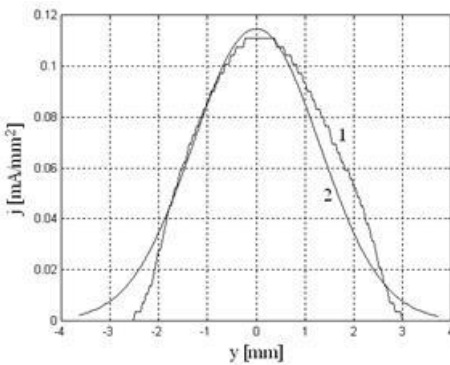
$$\frac{\sigma_2}{\sigma_1} = \frac{s_2}{s_1} = M = \frac{a_2}{a_1} \quad (\text{see Table 4}) \quad (79)$$

The value of the standard deviation of the image in the focus $\sigma_2 = s_{x0} = 0.1661$ [mm] is estimated from the experimental data (Table 4-e)), consequently the variance $\sigma_1 = \sigma_2 / M = 0.1908$.

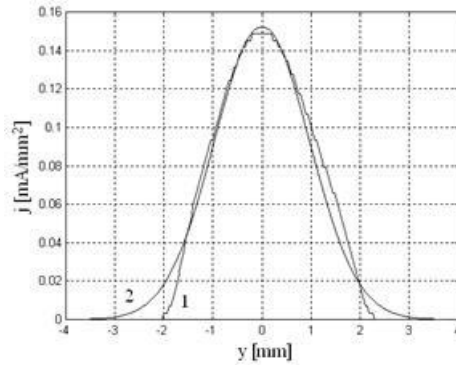
Calculations are made for the respective cross-sections from the Y-axis. The estimated normal integral beam current distributions together with the experimental ones are presented on Figure 34. The values of the focus position f are the same as the ones given in Table 4 for the X-axis cross-sections. The values of a_2 and M are also the same. The value of the standard deviation of the image in the focus $\sigma_2 = s_{y0} = 0.1713$ [mm] is estimated from the experimental data (Table 5)), consequently the variance $\sigma_1 = \sigma_2 / M = 0.1968$.

Table 4. The parameters of the beam current distribution along X-axis

Figure 33	N_L	f [mm]	s_{xi} [mm]	s_{xi}^2 [mm ²]	a_2 [mm]	$M = a_2/a_1$	$\sigma_{210} = \sigma_1 M$
a)	0	493.0318	1.2828	1.6456	1450.6383	1.9423	0.3706
b)	50	447.3539	0.9749	0.9504	1115.5091	1.4936	0.2850
c)	100	409.4222	0.6727	0.4525	906.1652	1.2133	0.2315
d)	150	377.4209	0.3837	0.1472	762.9821	1.0216	0.1949
e)	205	347.5388	0.1661	0.0276	650.0000	0.8703	0.1661
f)	250	326.3956	0.2897	0.0839	579.7600	0.7762	0.1481
g)	300	305.7294	0.5705	0.3255	517.6114	0.6930	0.1322
h)	350	287.5243	0.8510	0.7242	467.4968	0.6259	0.1194
i)	415	266.8662	1.2121	1.4692	415.2339	0.5560	0.1061



a)



b)

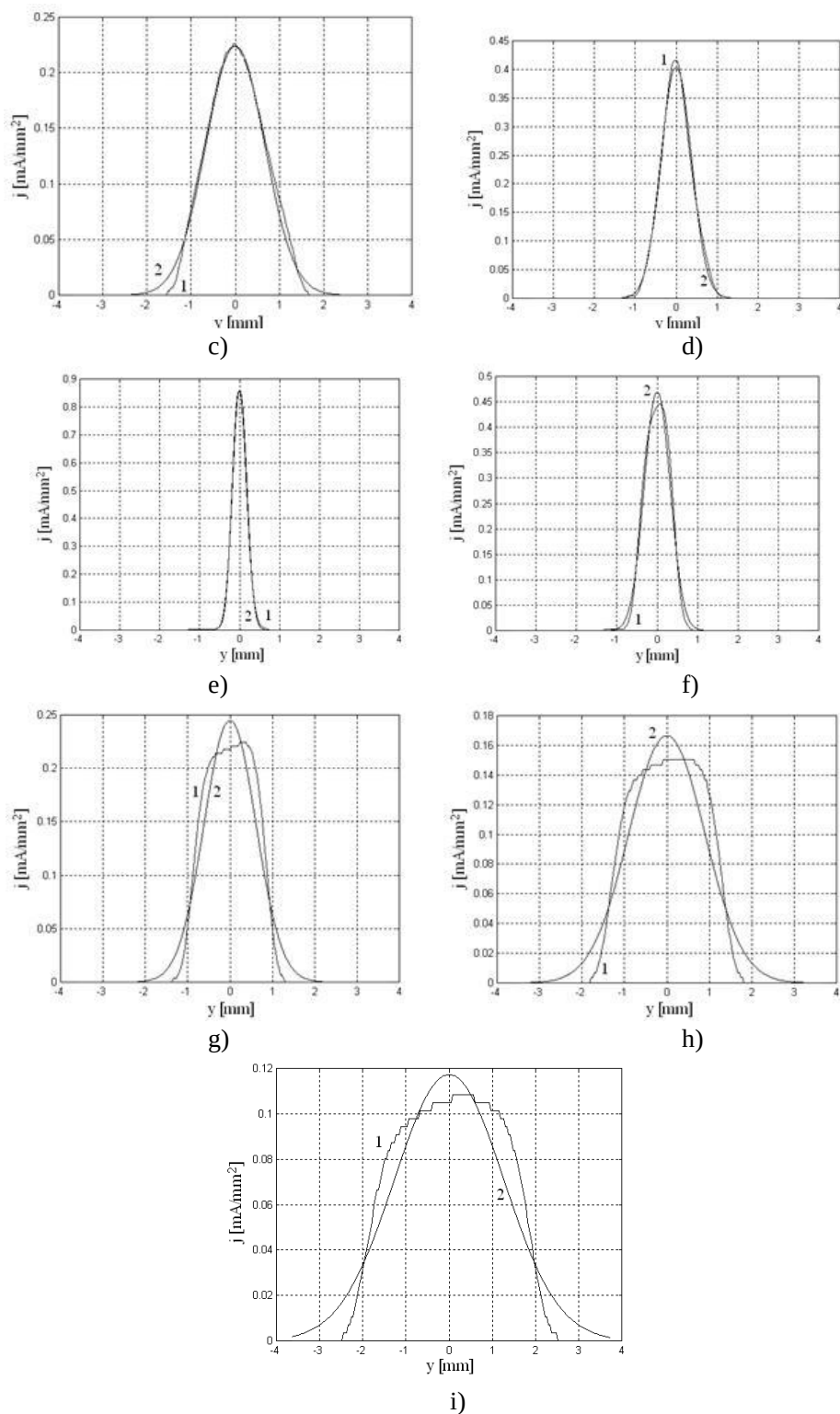


Figure 34. Experimental (curves 1) and fitted to normal (curves 2) integral beam current distributions at different “focus” positions from Y-axis scans

Table 5. The parameters of the beam-current distribution along Y-axis

Figure 34 4	N _L	S _{yi} [mm]	S _{yi} ² [mm ²]	σ _{2i0} = σ ₁ M
a)	0	1.2121	1.4692	0.3822
b)	50	0.9698	0.9405	0.2939
c)	100	0.6597	0.4352	0.2388
d)	150	0.3656	0.1337	0.2011
e)	205	0.1713	0.0293	0.1713
f)	250	0.3145	0.0989	0.1528
g)	300	0.6038	0.3646	0.1364
h)	350	0.8873	0.7873	0.1232
i)	415	1.2602	1.5881	0.1094

To characterize the beam quality through the values of beam emittance could be used the equation:

$$(\sigma_{xi})^2 = (\sigma_{x0i})^2 + (z_{0i} - z_0)^2 (\sigma_{x'0i})^2, \quad (80)$$

written at a condition of zero value of the co-variance between x and x' in the canonic position of the emittance diagram, one can find $\sigma_{x'0i}$ at measured σ_{xi} . The parameters of the beam current distribution, calculated on the base of experimental data, for the mentioned nine positions of the “focus” plane along X-axis are given in Table 6.

Since $\sigma_i^2 = \sigma_{0i}^2 + \Delta_{0-i}^2 \sigma_{i0}^2$, then:

$$\sigma_{0i}^2 = \frac{\sigma_i^2 - \sigma_{i0}^2}{\Delta_{0-i}^2}. \quad (81)$$

Then the emittance $\varepsilon = \pi \sigma_{i0} \sigma_{i0}'$ [mm.mrad], the mean value of ε is 0.557037 [mm.mrad].

The obtained results for the beam emittance along Y-axis are presented in Table 7. The values of $|\Delta_{0-i}|$ are the same as those in Table 6. The mean value of ε is 0.578951 [mm.mrad]. The canonical presentation of the emittance diagram can be calculated using the ellipse equation:

$$\left(\frac{x}{a_x}\right)^2 + \left(\frac{x'}{b_x}\right)^2 = a_x b_x = \varepsilon_p = \pi \cdot \sigma_{i0} \cdot \sigma_{i0}'$$

From the obtained results is concluded, that the current distribution of the beam is very close to an axis-symmetrical one, which reveals its good adjustment. Contour plots of the canonical view of the emittance is calculated for the investigated 8 cross-sections (without the beam focus) by finding the mean values of σ_x and $\sigma_{x'}$ from the data for x and y assuming the case that in a rotation symmetric beam they are identical. In this way a transition is made to rr' coordinate system (instead of xx' and yy'). The mean values are given in Table 8.

Table 6. The beam emittance along X-axis from the investigated cross-sections

N_L	$ \Delta_{0-i} = a_{i0}-a_{650} $	$ \Delta_{0-i} ^2$ [mm ²]	σ_{i0}^2 [mm ²]	$\sigma_{oi}'^2$ [mm ²]	σ_{oi}' [mm]	ε [mm.mrad]
0	800.6383	641021.6874	0.1373	$2.3529 \cdot 10^{-6}$	$1.5339 \cdot 10^{-3}$	0.5685
50	465.5091	216698.7222	0.0812	$4.0111 \cdot 10^{-6}$	$2.0028 \cdot 10^{-3}$	0.5708
100	256.1652	65620.6097	0.0536	$6.0789 \cdot 10^{-6}$	$2.4655 \cdot 10^{-3}$	0.5708
150	112.9821	12764.9549	0.0380	$8.5547 \cdot 10^{-6}$	$2.9248 \cdot 10^{-3}$	0.5701
205	0	0	0.0276	*	*	*
250	70.2400	4933.6576	0.0219	$12.5667 \cdot 10^{-6}$	$3.5450 \cdot 10^{-3}$	0.5250
300	132.3886	17526.7414	0.0175	$17.5731 \cdot 10^{-6}$	$4.1920 \cdot 10^{-3}$	0.5542
350	182.5032	33307.4180	0.0143	$21.3136 \cdot 10^{-6}$	$4.6167 \cdot 10^{-3}$	0.5512
415	234.7661	55115.1217	0.0113	$26.4519 \cdot 10^{-6}$	$5.1431 \cdot 10^{-3}$	0.5457

Table 7. The beam emittance along Y-axis from the investigated cross-sections

N_L	σ_{i0}^2 [mm ²]	$\sigma_{oi}'^2$ [mm ²]	σ_{oi}' [mm]	ε [mm.mrad]
0	0.1461	$2.0641 \cdot 10^{-6}$	$1.43669 \cdot 10^{-3}$	0.54910380
50	0.0864	$3.9415 \cdot 10^{-6}$	$1.98533 \cdot 10^{-3}$	0.58348751
100	0.0570	$5.7630 \cdot 10^{-6}$	$2.40063 \cdot 10^{-3}$	0.57327149
150	0.0404	$7.3058 \cdot 10^{-6}$	$2.70293 \cdot 10^{-3}$	0.54355978
205	0.0293	*	*	*
250	0.0233	$15.3136 \cdot 10^{-6}$	$3.91326 \cdot 10^{-3}$	0.59794645
300	0.0186	$19.7410 \cdot 10^{-6}$	$4.44308 \cdot 10^{-3}$	0.60603641
350	0.0152	$23.1817 \cdot 10^{-6}$	$4.81474 \cdot 10^{-3}$	0.59317538
415	0.0120	$28.5971 \cdot 10^{-6}$	$5.34762 \cdot 10^{-3}$	0.58503005

Table 8. The mean values for X- and Y-axes of σ , σ' and ε

N_L	f [mm]	σ_{r0}	σ_{r0}' [mm]	ε [mm.mrad]
0	493.0318	0.37640	0.0014853	0.567994
50	447.3539	0.28945	0.0019967	
100	409.4222	0.23515	0.0024331	
150	377.4209	0.19800	0.0028139	
250	326.3956	0.15045	0.0037291	
300	305.7294	0.13430	0.0043175	
350	287.5243	0.12130	0.0047157	
415	266.8662	0.10775	0.0052454	

On Figure 35 are presented the plots of the dependencies between the main axes of the ellipse of the emittance (σ_{r0} and $\sigma_{r0}', 100$) and the focus position from the main axis of the focusing lens f .

In order to calculate easily the values of these axes as a function of the focus position value, regression equations are estimated:

$$\sigma_{r0} = -0.22068 + 0.0024082 f - 0.0000067632 f^2 + 0.00000000879 f^3; \quad (82)$$

$$\sigma'_{r0} = 0.0153310 - 0.000049513 f + 0.00000004367 f^2. \quad (83)$$

The continuous curves on Figure 35 represent the functions (82) and (83), while the dots show the calculated emittance ellipse axes values for the investigated cross-sections (Table 8).

The relation between the distance of the focus position from the main axis of the focusing lens f and the corresponding number of the bitmap line N_L calculated by eq. (76) is shown on Figure 36.

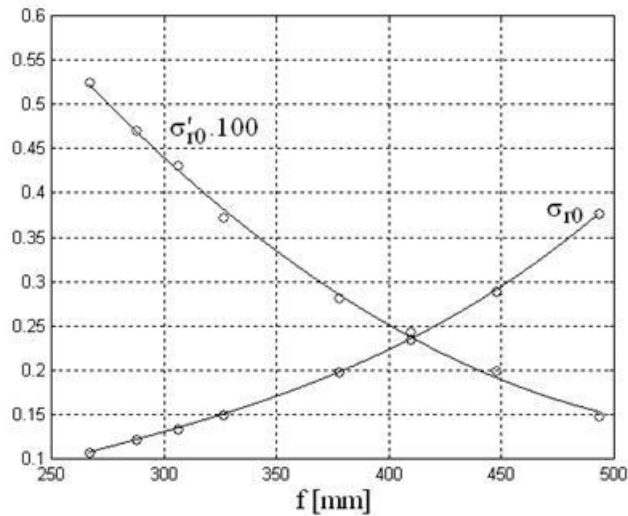


Figure 35. Dependencies between the main axes of the ellipse of the emittance (σ_{r0} and $\sigma'_{r0.100}$) and the focus position from the main axis of the focusing lens f

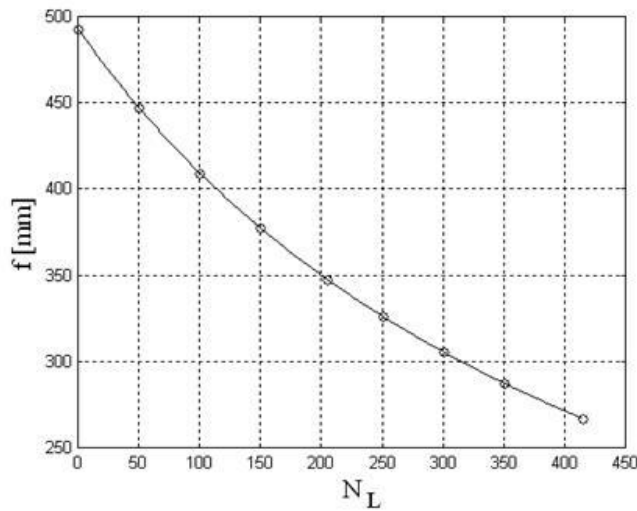


Figure 36. Relation between the distance of the focus position from the main axis of the focusing lens f and the corresponding number of the bitmap line N_L - eq. (76)

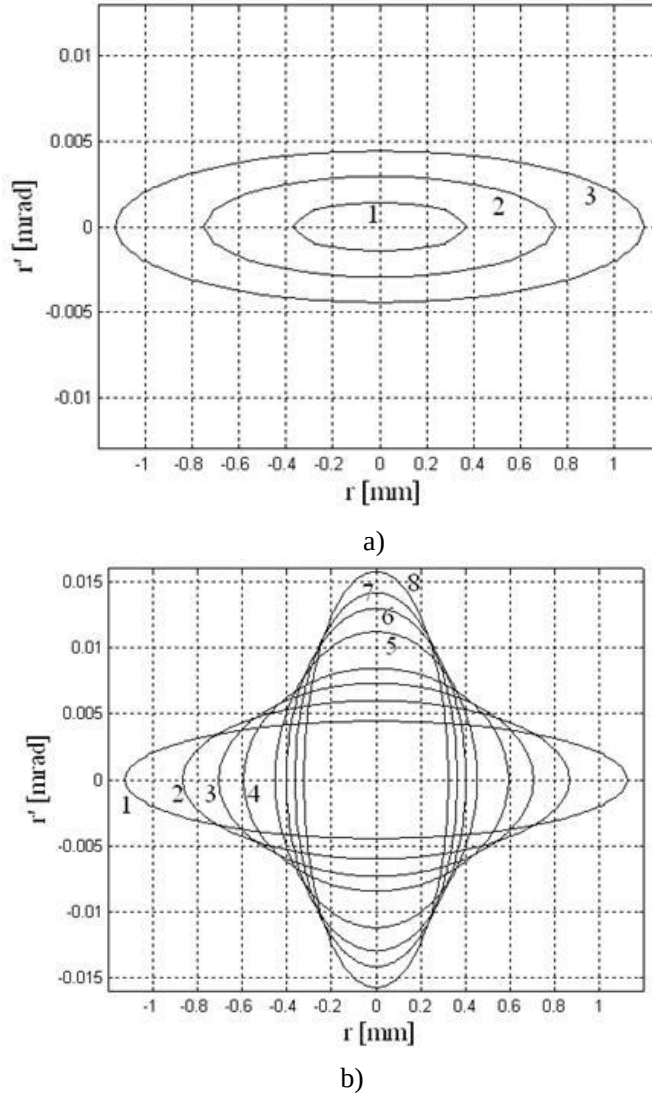


Figure 37. Contour plots of the emittance in canonical view for different focus positions and parts of the beam current p :

a. ellipses: 1 is calculated for $p=0.39$; 2 – for $p=0.86$; 3 – for $p=0.99$; $N_L=0$, b) $p=0.99$ and ellipses: 1 – for $N_L=0$; 2 – for $N_L=50$; 3 – for $N_L=150$; 4 – for $N_L=250$; 5 – for $N_L=300$; 6 – for $N_L=350$; 7 – for $N_L=415$. b. ellipses position in r, r' plane for $p=0.99$

On Figure 37a are presented the contour plots of the emittance in canonical view for the cross-section $N_L=0$. The contours are evaluated for parts of the beam current: $p=0.39$, 0.86 and 0.99 .

On Figure 37b is given the emittance canonical view of all the investigated cross-sections for $p=0.99$.

The current density distribution in the phase plane can be defined as particle flow per $\text{mm} \times \text{mrad}$. It is calculated for the first cross-section (Table 8) assuming its normal distribution and axis-symmetrical beam. 2D and 3D view of this distribution is presented on Figure 38 a,b.

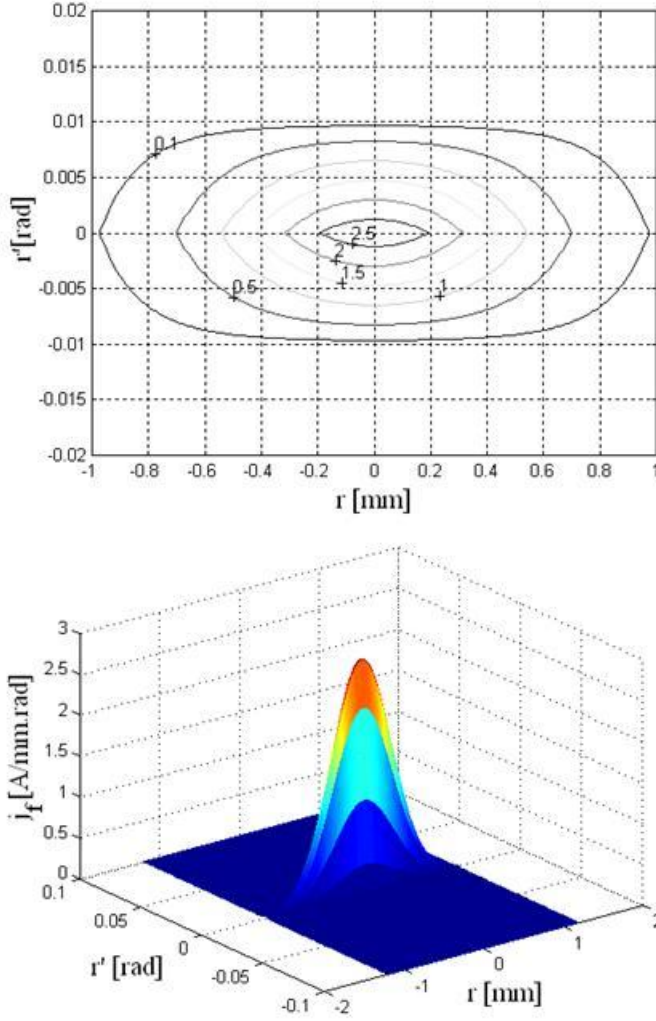


Figure 38. 2D and 3D presentation of the calculated current density in the phase plane from the first cross-section (NL=0)

$$j_r(r, r') = \frac{10}{2\pi\sigma_r\sigma_{r'}(1-\rho^2)^{1/2}} \exp\left(-\frac{1}{2(1-\rho^2)} \cdot \left[\left(\frac{r}{\sigma_r}\right)^2 - 2\rho\left(\frac{r}{\sigma_r}\right)\left(\frac{r'}{\sigma_{r'}}\right) + \left(\frac{r'}{\sigma_{r'}}\right)^2 \right]\right). \quad (84)$$

Figure 39 shows 2D and 3D view of the calculated current density in the beam focus.

Another invariant, besides the emittance, the beam brightness per volt accelerating voltage is:

$$(B/U)_p = 2I_p / (\pi^2 \epsilon_p^2 U). \quad (85)$$

There B/U is the average value for emittance ellipse, through which the part I_p of the beam current is transferred.

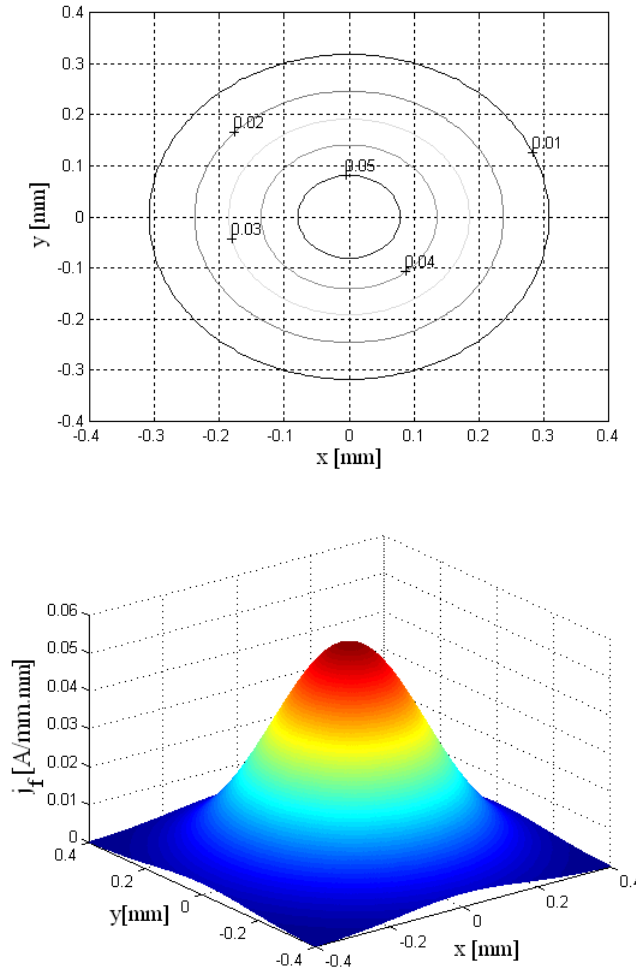


Figure 39. 2D and 3D view of the calculated current density in the beam focus (NL=205). Note that the coordinate

The values of $(B/U)_p$ calculated for some parts of the electron beam current are given in Table 9. The obtained from experimental data values for the brightness differ slightly for the different cross-sections for different parts of the beam current. Their mean values presenting theoretically invariant $(B/U)_p$ are calculated.

The power density distribution is calculated assuming 2D normal distribution for the different focusing positions, corresponding to the explored 9 cross-sections. The obtained results are presented on Figure 40. The formula used is:

$$P_0(x,y) = \frac{600}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{1}{2} \cdot \left[\left(\frac{x}{\sigma_x}\right)^2 + \left(\frac{y}{\sigma_y}\right)^2\right]\right), \quad (86)$$

where correlation $\rho=0$ is assumed.

On Figure 41 is presented 3D view of the power density distribution, calculated for the beam focus – case e) on Figure 40.

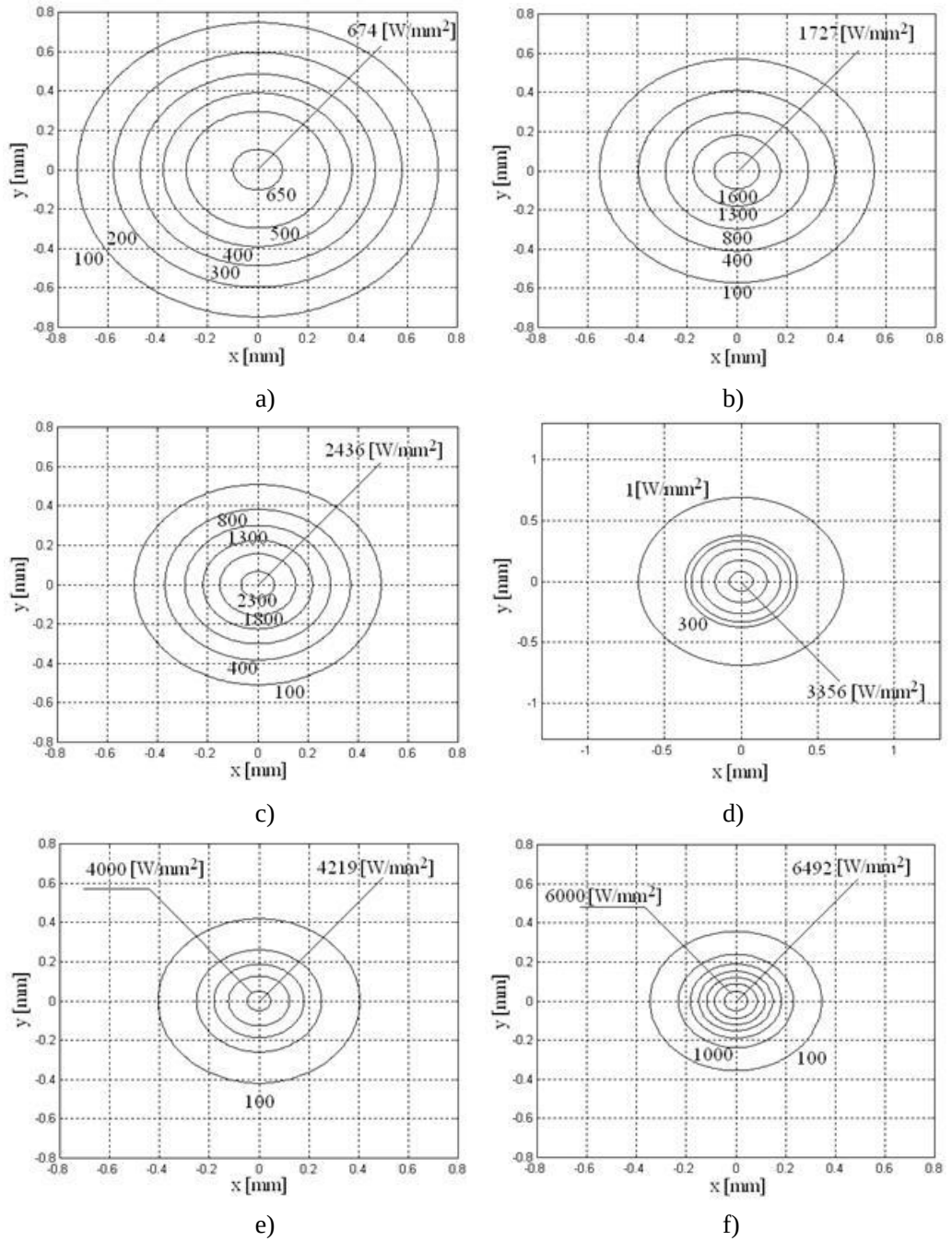
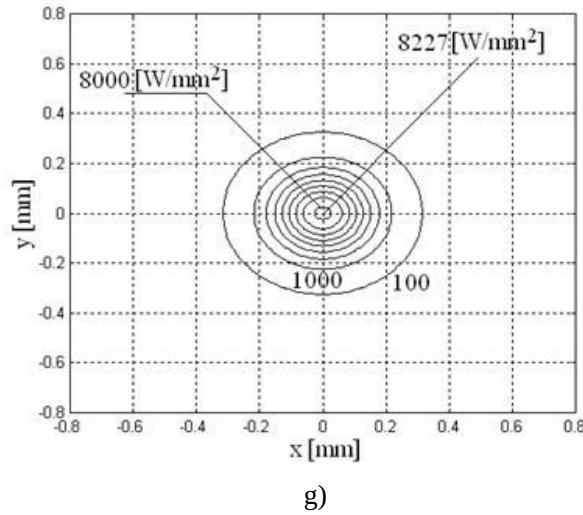


Figure 40. (Continued)



e) The contours that are not signed have levels: $P_0 = 500, 1000, 2000, 3000$ [W/mm²]

f) The contours that are not signed have levels: $P_0 = 800, 1300, 1800, 2300, 2900$

h) The contours that are not signed have levels: $P_0 = 2000, 3000, 4000, 5000$ [W/mm²]

i) The contours that are not signed have levels:

$P_0 = 2000, 3000, 4000, 5000, 6000, 7000$

Figure 40. The power density distribution P_0 for the different focusing positions. Signatures a)-i) correspond to $N_L=0$ to $N_L=415$. The contours represent 2D presentation of a given constant level of the function $P_0(x,y)$

Table 9. Brightness per volt accelerating voltage (B/U)_p. The index p determines the calculation in the given part of the beam current p

$p=I_m/I_b$		0,39	0,63	0,78	0,86	0,92	0,99
ϵ_p		$\sigma_x \cdot \sigma_{x'}$	$2\sigma_x \cdot \sigma_{x'}$	$3\sigma_x \cdot \sigma_{x'}$	$4\sigma_x \cdot \sigma_{x'}$	$5\sigma_x \cdot \sigma_{x'}$	$9\sigma_x \cdot \sigma_{x'}$
(B/U) _p	$N_L = 0$	$4.2185 \cdot 10^{10}$	$1.7036 \cdot 10^{10}$	$9.3744 \cdot 10^9$	$5.8139 \cdot 10^9$	$3.9805 \cdot 10^9$	$1.3220 \cdot 10^9$
	$N_L = 50$	$3.9474 \cdot 10^{10}$	$1.5941 \cdot 10^{10}$	$8.7720 \cdot 10^9$	$5.4403 \cdot 10^9$	$3.7247 \cdot 10^9$	$1.2371 \cdot 10^9$
	$N_L = 100$	$4.0279 \cdot 10^{10}$	$1.6266 \cdot 10^{10}$	$8.9508 \cdot 10^9$	$5.5512 \cdot 10^9$	$3.8006 \cdot 10^9$	$1.2623 \cdot 10^9$
	$N_L = 150$	$4.2475 \cdot 10^{10}$	$1.7153 \cdot 10^{10}$	$9.4390 \cdot 10^9$	$5.8540 \cdot 10^9$	$4.0079 \cdot 10^9$	$1.3311 \cdot 10^9$
	$N_L = 250$	$4.1888 \cdot 10^{10}$	$1.6916 \cdot 10^{10}$	$9.3085 \cdot 10^9$	$5.7731 \cdot 10^9$	$3.9525 \cdot 10^9$	$1.3127 \cdot 10^9$
	$N_L = 300$	$3.9216 \cdot 10^{10}$	$1.5837 \cdot 10^{10}$	$8.7147 \cdot 10^9$	$5.4048 \cdot 10^9$	$3.7004 \cdot 10^9$	$1.2290 \cdot 10^9$
	$N_L = 350$	$4.0297 \cdot 10^{10}$	$1.6274 \cdot 10^{10}$	$8.9548 \cdot 10^9$	$5.5537 \cdot 10^9$	$3.8024 \cdot 10^9$	$1.2629 \cdot 10^9$
	$N_L = 415$	$4.1275 \cdot 10^{10}$	$1.6669 \cdot 10^{10}$	$9.1723 \cdot 10^9$	$5.6886 \cdot 10^9$	$3.8947 \cdot 10^9$	$1.2935 \cdot 10^9$
MEAN		$4.0886 \cdot 10^{10}$	$1.6511 \cdot 10^{10}$	$9.0858 \cdot 10^9$	$5.6349 \cdot 10^9$	$3.8580 \cdot 10^9$	$1.2813 \cdot 10^9$

Regression equation giving the dependence between the maximum value of the beam power density distribution $P_{0\max}$ and any focus position from the main axis of the focusing lens f is estimated:

$$P_{0\max} = 116668 - 934.43 f + 2.9571 f^2 - 0.0043089 f^3 + 0.00000241 f^4 \quad (87)$$

This function - $P_{0\max}(f)$, together with the calculated data from the investigated cross-sections (signed with dots) are presented on Figure 42.

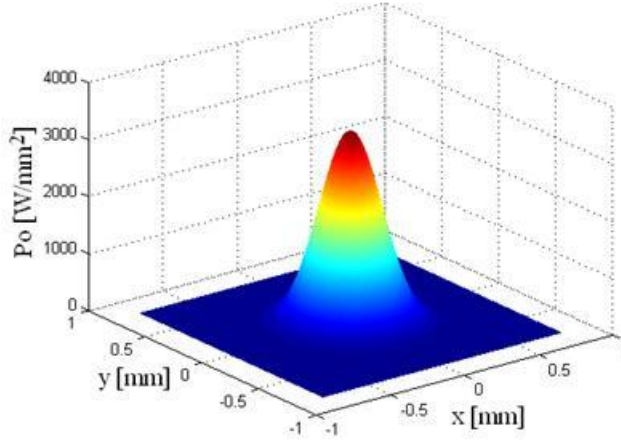


Figure 41. 3D view of the power density distribution in the beam focus (NL=0)

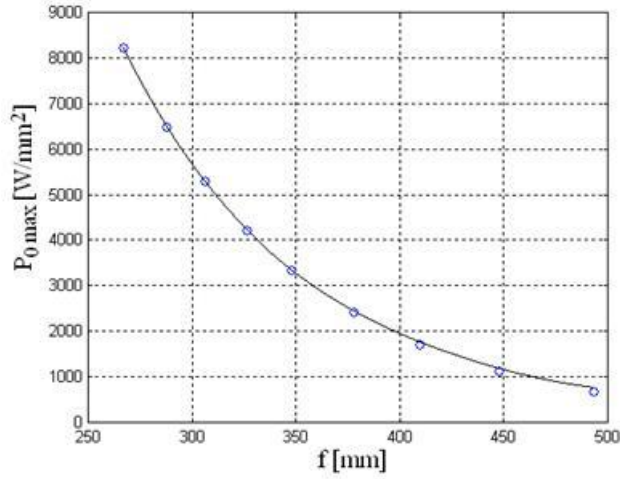


Figure 42. The maximum value of the beam power density distribution P0max and any focus position from the main axis of the focusing lens f

Analysis of Medium Current (or Partially Commenced) Electron Beams where the Space-Charge and Emittance Effects Are Comparable

To calculate beam divergence θ and beam emittance ε in that case, the equation (42) for an axial symmetrical nonrelativistic beam could be applied. Let us discuss a beam propagating in vacuum (no compensation of space charge occurs) in drift region. Then the differential equation (42) could be rewritten

$$\frac{d^2 R}{dz^2} + \frac{\eta B^2}{8U} R = \frac{I}{4\pi\epsilon_0 \cdot U^{3/2} \cdot \sqrt{2\eta}} \cdot \frac{1}{R} + \frac{\varepsilon^2}{R^3} \quad (88)$$

where I is the beam current; ε_0 is the dielectric permittivity; U is the acceleration voltage; B is the axial magnetic field; η is the electron charge-to-mass-ratio; ε is the beam emittance. Numerical solution of eq. (88) give the beam envelope R dependence (i.e. an evaluation of "some averaged" beam radius due to the distributed on R beam current) on the magnetic lens field intensity (or on the distance lens-focal plane or most usable focusing coil current) at constant U and z . Here, the concept of the rms (root mean square) emittance could be used, if the corresponding values of the beam divergence θ and the beam radius r are defined as second moments.

The measurements are described on Figure 31 and Figure 32 at use of an static measuring plate of refractory material with two perpendicular narrow slits. Beam envelop diameter are calculated statistically for all scans, presented as bitmap lines on Figure 32 during the variation of beam focus position from $f_0 - \Delta f$ to $f_0 + \Delta f$. This measurement is able to determine and correct beam astigmatism and beam misalignment additionally. After scan the two orthogonal slits and measuring line integrals the beam jumps back to the starting point of x scan on the first slit very fast. A small increment of focusing coil changes the beam focus position. Than x and y scans are fulfilled again. As a result a number of line integrals of beam intensity are collected as this is shown on Figure 32. There bright shows a high power density, dark present a low intensity. The generated bitmap consists of two beam profiles, which represents the beam dimension in x and y directions as function of the focal lengths. The recording of such a bitmap with a resolution of 400X400 Pixel could be realized for about 100 ms. During bigger part of the measuring time the beam is defocused. Only for the short time when the focal length is coincident with the central focal length f_0 (of order of 1-2 ms) high power intensity is deposited on the measuring sensor. Thus his destruction can be avoided.

The analysis starts with the determination of the beam diameter for every single focal length. According to ISO 11146 one has to find the centroid of the intensity distribution first. Knowing X, Y of the distribution centre, the beam diameter can be calculated the second moment of beam width is:

$$\langle x^2 \rangle^{\frac{1}{2}} = 4 \left[\frac{\iint (x - X)^2 I(x, y)}{\iint I(x, y)} \right]^{\frac{1}{2}}. \quad (89)$$

where x is the current position of the pixel; X : centroid of the intensity distribution; I : intensity.

Note, that it is important to subtract the background noise very carefully, because in (89) the term $(x-X)^2$ overemphasizes small signals located far away from the centre of the intensity distribution. With the calculated position of the centre X and Y and the related beam dimensions $\langle x^2 \rangle^{\frac{1}{2}}$ and $\langle y^2 \rangle^{\frac{1}{2}}$ it is possible to specify beam astigmatism, beam alignment and exact focal position. Beams with power up to 2 kW can be measured continuously. Then it is possible to correct astigmatism and misalignment of the beam automatically by using the obtained data to control the corrector coils of the EB-gun. The beam can be focussed exactly on the surface of the work piece by determining the position of the minimum beam diameter.

To calculate the beam divergence θ and the beam emittance ε a more advanced analysis of the data is necessary. The propagation of a charge carrying particle beam is described by the following equation

$$\frac{d^2 R}{dz^2} + \frac{\eta B^2}{8U} R = \frac{I}{4\pi\epsilon_0 \cdot U^{3/2} \cdot \sqrt{2\eta}} \cdot \frac{1}{R} + \frac{\varepsilon^2}{R^3} \quad (90)$$

where $R = (\frac{1}{2} \langle x^2 \rangle)^{\frac{1}{2}}$; I is the beam current; ϵ_0 is dielectric constant; U is beam acceleration voltage; η is charge-to-mass-ratio for the electron; ε : rms beam emittance; B : axial magnetic field.

Here, the concept of the rms (root mean square) emittance is introduced (See Figure 14, and eq. (61). The divergence θ and beam radius r values are defined as second moments (see e.g. equation (77)).

Beam parameters in a not very powerful EB welding machine is typically in a range, that the influence of the space-charge on beam propagation could be neglected ($U_A > 50$ kV; $I_b < 200$ mA). Therefore the dominating effect on the beam envelope beside external electric and magnetic fields is the emittance.

In that case and for field-free space, equation (88) and (90) are reduced to

$$\frac{d^2 R}{dz^2} = \frac{\varepsilon^2}{R^3} \quad (91)$$

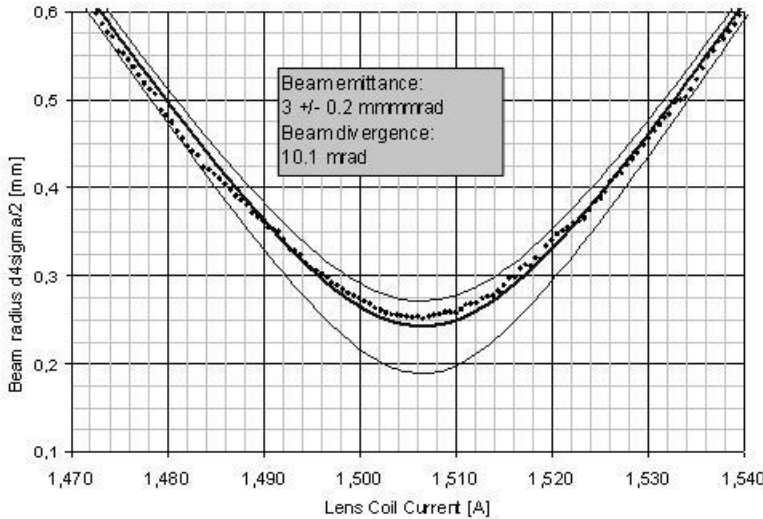


Figure 43. Measured (dots) and calculated (solid line) beam radius on the sensor at different current through the magnetic lens. Determined emittance is 3.0 mm.mrad, evaluated divergence of the beam is 10.1 mrad

It is possible to solve this differential equation for a converging beam (focussing with a magnetic lens). If the radius R is determined at a fixed position z_0 , while beam divergence θ_0 , starting radius R_0 and emittance ε are parameters, the solution of (91) has the following form:

$$R(z_0)^2 = (z_0 - \frac{\theta_0 \cdot R_0}{T1})^2 \cdot T1 + \frac{\varepsilon^2}{T1} \quad (92)$$

with $T1 = \theta^2 + \varepsilon^2/R_0^2$; and θ , ε , R_0 defined as rms values.

With the right choice of θ , ε , R_0 the graph of (92) can be fitted to the measured beam diameter (see Figure 43). Thus divergence and emittance of the studied beam is given. The big amount of the measured beam diameters (several hundreds) leads to a very reliable result.

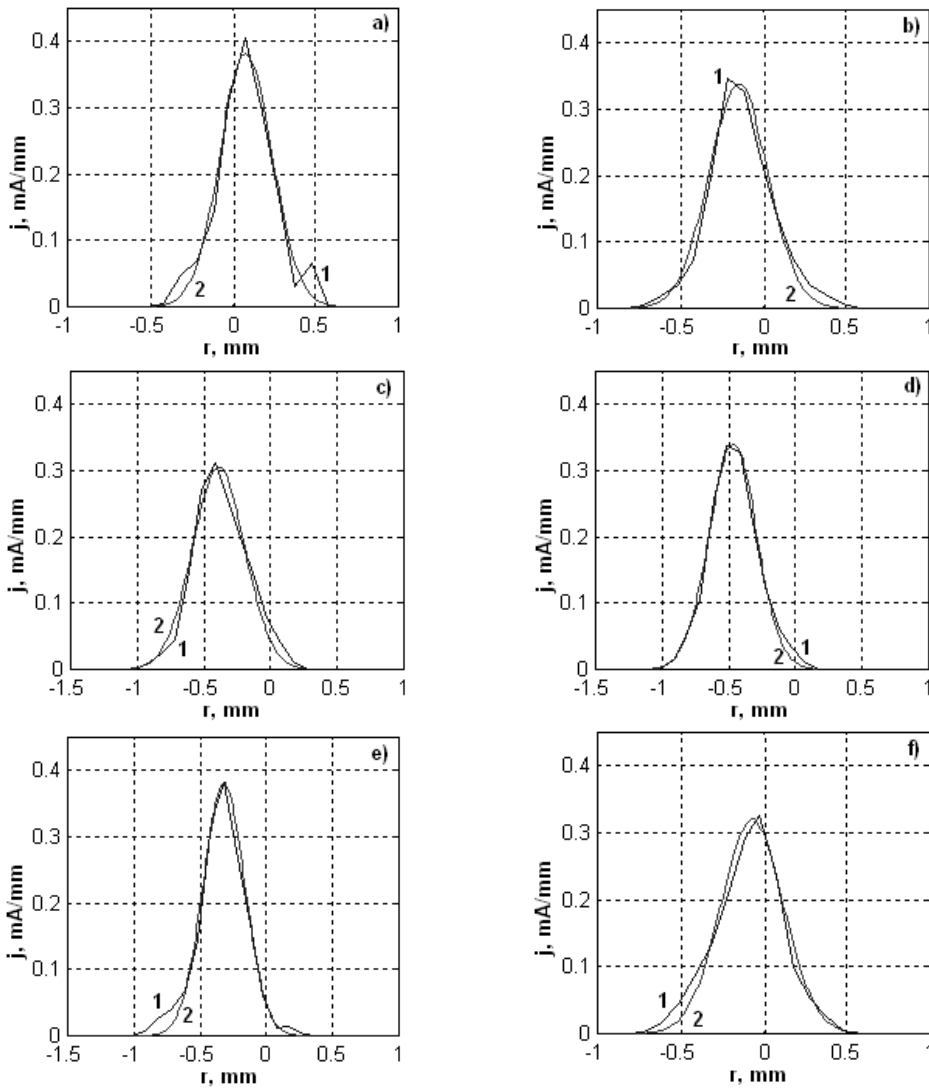


Figure 44. Continued

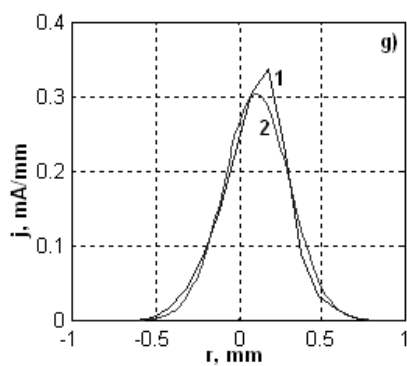


Figure 44. Integral current densities at beam focus (at distance 320 mm from the focal winding) at different angles:

a) $\theta = 0^\circ$; b) $\theta = 51^\circ$; c) $\theta = 102^\circ$; d) $\theta = 153^\circ$; e) $\theta = 204^\circ$; f) $\theta = 255^\circ$; g) $\theta = 306^\circ$. 1 – Experimental; 2 – Approximated

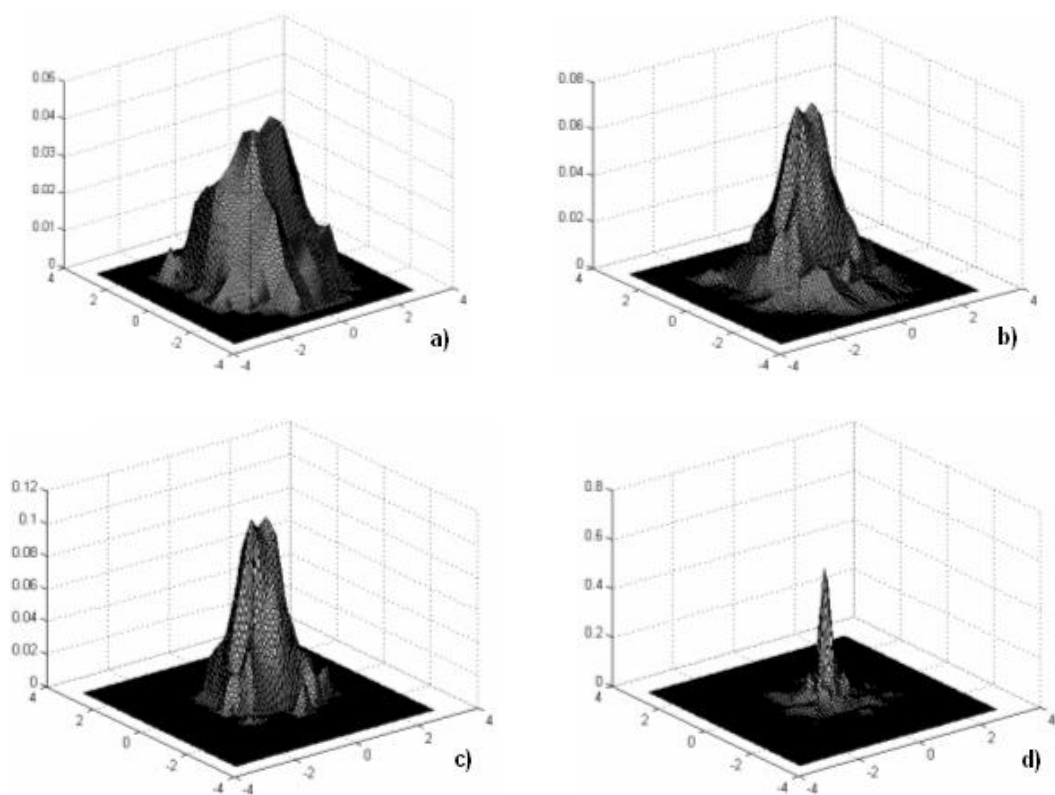


Figure 45. Continued

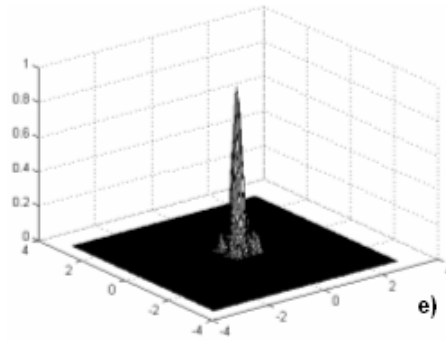


Figure 45. Reconstructed radial current density distributions [mA] depending on x and y [mm] coordinates in five cross-sections of the beam at different distances from the focal winding: a) $z = 170$ mm; b) $z = 207.5$ mm; c) $z = 245$ mm; d) $z = 282.5$ mm; e) $z = 320$ mm (focus)

Tomographic Approach – Measurement of Integral Current Densities at Different Angles and Obtaining Emittance Values

The tomography reconstruction of a two-dimensional beam profile from a set of its one-dimensional projections, measured as an array of line integrals (by wire probe collector or proposed from Elmer modified Faraday cup with radial slits) of a cross section of the beam could be applied to get the beam emittance values too.

There a Fourier transformation from real to frequency space and a consequent back Fourier transformation permits to reconstruct the beam cross-section current density distribution image (beam radial intensity profile) with their asymmetry features and without need to assume a theory beam distribution prior calculation.

The estimation of the beam emittance is performed using the methods described previously in that chapter. An example of such estimation is given on the base of 7 projections, sampled by 7 radial slits (with wide 0,1mm and placed at 51° from each on the refractory disc (Figure 24) in the modified Faraday cup (see Figure 22 and Figure 28).

The experimentally measured voltage signal is stored by a digital storage oscilloscope with sampling rate up to 250 MS/s, at the beam moving in circle.

The values σ_x , $\sigma_{x'}$ are the radial and the angular standard deviations. The transformations of coordinate x and x' in the drift space (that is free from external to the beam forces) are given in a matrix expression (66).

At measured values of σ_{x1} , σ_{x2} and σ_{x3} and known L_{1-2} and L_{1-3} , the “focus” (or “waist”) the parameters L_{0-1} , σ_{x0} and $\sigma_{x0'}$ are found from system equations (67-69).

In Table 10 are presented the calculated results for the variances, standard deviations of the radial and angular distributions and the calculated emittance values.

The beam emittance, as well as the beam profile are significant and appropriate characteristics of the beam quality. The measurement of these characteristics will: (i) help standardization of electron optical systems, (ii) provide adequate conditions for welding production quality control by keeping a high reproducibility of the welds (iii) support the attempts to transfer the concrete technology from one welding machine to another and (iv) at creating expert systems for an operator choice of suitable regimes for gaining desirable welds.

Table 10. The variances, standard deviations of the radial and angular distributions and the beam emittance

	$z = 320 \text{ mm}$ (focus)	$z = 320 \text{ mm}$ (focus)	$z = 245 \text{ mm}$	$z = 245 \text{ mm}$
θ°	m [mm]	$\sigma_{i0}^2 \text{ [mm}^2\text{]}$	m [mm]	$\sigma_{i0}^2 \text{ [mm}^2\text{]}$
0	0.08	0.16	0.08	0.60
51	-0.15	0.18	-0.18	0.57
102	-0.39	0.20	-0.38	0.53
153	-0.47	0.18	-0.45	0.56
204	-0.32	0.16	-0.31	0.58
255	-0.07	0.19	-0.07	0.56
306	0.11	0.20	0.10	0.51
	$z = 245 \text{ mm}$	$z = 245 \text{ mm}$	$z = 245 \text{ mm}$	$z = 245 \text{ mm}$
	$\sigma_{oi}^{\prime 2} \text{ [rad}^2\text{]}$ $(\sigma_{245,0}^2 - \sigma_{i0}^2) / 75^2$	$\sigma_{oi}' \text{ [rad]}$	$\varepsilon \text{ [mm.mrad]}$ (C=4) (P=0.86)	$\varepsilon \text{ [mm.mrad]}$ (C=9) (P=0.99)
0	$7.8222 \cdot 10^{-5}$	0.0088	14.1	31.7
51	$6.9333 \cdot 10^{-5}$	0.0083	14.1	31.7
102	$5.8667 \cdot 10^{-5}$	0.0077	13.8	31.0
153	$6.7556 \cdot 10^{-5}$	0.0082	13.9	31.3
204	$7.4667 \cdot 10^{-5}$	0.0086	13.8	31.0
255	$6.5778 \cdot 10^{-5}$	0.0081	14.1	31.8
306	$5.5111 \cdot 10^{-5}$	0.0074	13.2	29.8
	MEAN		13.857	31.186

2. DESIGN AND OPTIMIZATION OF THE HIGH BRIGHTNESS ELECTRON OPTICAL SYSTEMS FOR WELDING

Beams of accelerated electrons are widely used in various fields of pure and applied physics as well as in the key technologies of machine building, electronics and manufacturing of advanced materials. It can be noted, that the requirements for electron beam utilizing as thermal source for welding are specific and bellows will be discussed main features of its design, characterization and optimization.

The device in which appropriate electron beam is produced and shaped is termed *electron sources* or electrostatic part of *electron optical system* (EOS). Often the completed EOS additionally to the electrostatic part contains a set of other electron – optical elements, which constitute the beam transport system. These are the focusing and the deflection coils. In the case of EBW guns, designed for beam operation in open air have diaphragm, chambers and systems for pumping intermediate vacuum and input of He as shielding gas around the beam. For propagation of beam through small diaphragms there additional magnetic lenses could be utilized. The whole EOS, generating, shaping and transporting the beams for technology applications, is called *electron guns*. The quality of the beam is connected with: (i) the

thermionic emission of the beam electrons (or ionization and extraction in the case of plasma emitter gun) and (ii) the beam formation by self-consistency of the particle trajectories and the existing electrical and magnetic fields in the electron gun. Usually in the technology applications the beams are continuously operated, but there are also pulsed beams. The electrodes of electrostatic part of a electron guns are placed in vacuum to avoid (or strongly decrease) the collisions of the beam electrons with the molecules of gasses in acceleration and first part of beam transportation spaces. Thermionic emitter offer high currents and have low requirements on the vacuum ($p \leq 10^{-2}$ Pa) and compromise life time (tens, at the most hundreds hours). In concrete implementations for welding of metals the specific requirements to the electron gun are: a) low emittance, high brightness, small aberrations; b) high concentration of the beam energy in the zone of the beam interaction with the welded material and c) stable and reliable operation. Some additional requirements could be: easy change of cathode; low beam current losses (i.e. negligible quantity of the beam electrons reaching the gun electrodes); simple configuration of the electrodes; smooth control of the beam current over a wide range of its operational values; quick-operating vacuum valve situated between the accelerating space of EOS and the space of the welding chamber (that permit at the change of joining parts by operator by opening technology chamber, the hot cathode ensemble to be in vacuum).

Electron Emission in Electron Guns Utilizing Thermionic Cathode

The electrons in the welding electron guns are emitted usually by a thermionic cathode, which supply the free electrons. The current density j_e at thermal electron emission from a cathode heated to temperature T_c is given by Richardson-Dushman equation:

$$j_e = A \cdot T_c^2 \exp\left(-\frac{e\varphi}{kT_c}\right), \quad (93)$$

where $e\varphi$ is the work function of the emitter (namely φ is the potential of the surface gap of cathode material electrons in free electron observation), e is the elementary charge of an electron, k is the Boltzmann constant being equal to $1.38 \cdot 10^{-23}$ J.deg⁻²; A is a constant, depending from the material of the cathode and construction of the electrodes. Theoretical value of A is $120 \text{ A}/(\text{cm}^2 \cdot \text{K}^2)$

In a diode emitting system (simplest two-electrode construction in which emitted current can be generated) the current density j_e will be observed only in the case of enough big potential drop applied on the cathode/anode space. Observation of a saturation of the emitted current collected by the anode at various voltages and given T_c could be seen (Figure 46); at lower voltages the current-voltage characteristics is controlled by Child - Langmuir law (exponent 3/2 law) as this is shown on Figure 46. The Figure 46-a is idealized case, and Figure 46b is the real observable current-voltage characteristics of a vacuum diode generator of electrons.

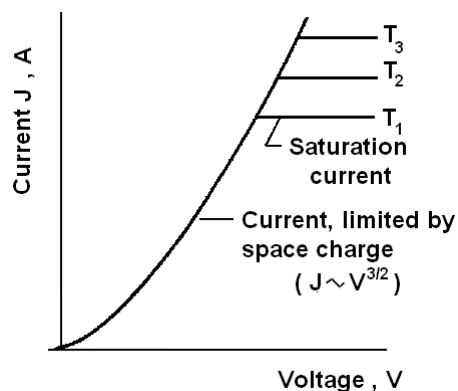


Figure 46a. I-V characteristics of an idealized vacuum diode; temperatures $T_3 > T_2 > T_1$

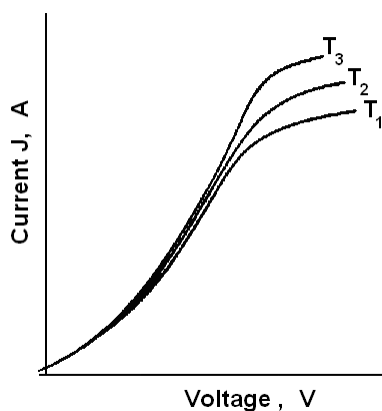


Figure 46b. I-V characteristics of a real vacuum diode

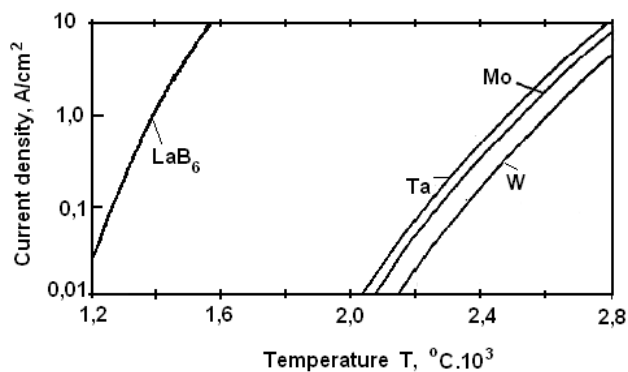


Figure 47. Current density vs. temperature of the cathode

To obtain desired high beam current density (or energy flow density) the current is emitted from cathodes obeying higher emission ability – as example Tungsten, Tantalum or LaB_6 . The choice of that materials is done as compromise between the emitted current densities and evaporation rates at given temperature and low ion sputtering yield (see Figure

47 to Figure 49). These factors limited the life of the emitter. Properties of emitter material after heating of the emitter(changes of crystalline grains; selective evaporation and/or activation etc., and workability are also important at that choice. An attempt to compare mentioned emitter materials is shown in Table 11. Additionally metals able to be used as emitters are Rhenium and Niobium. Rhenium obeys a similar behavior as the Ta at higher temperatures.

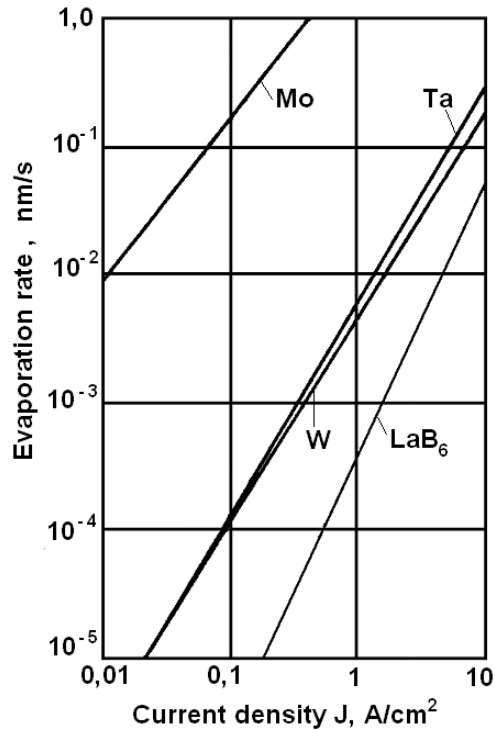


Figure 48. Evaporation rate vs. current density of cathode

Table 11. Emission properties of cathode materials

Property	Tungsten	Tantalum	Molybdenum	LaB ₆
ϕ	4.52	4.07	4.15	2.86 (2.36*)
T_c [°C]	2300-2700	1950-2150	1800-2000	1000-1600
A [A/cm ² ·°C ²]	60(70)	60(55)	55	73 (120*)
j_e [A/cm ²]	1-10	0.1-0.5	0.00083at 1600 °C	1-50
Ion bombardment stability	Very good	poor		good
Changes after heating	Becomes brittle	Remains soft		Active surface (improved emission at 1600 °C)
Workability	poor	good	good	Extremely poor

* data published in [39]

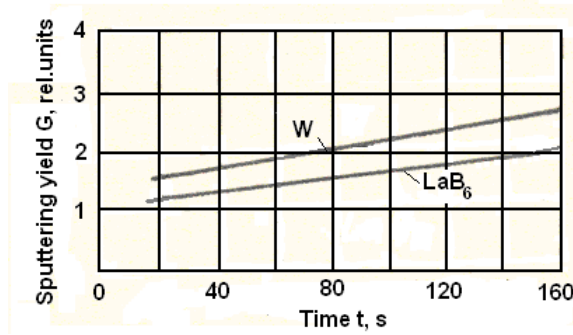


Figure 49. Relative ion sputtering yields of W and LaB₆ (abscissa-time; ordinate-weight losses)

From pure metals W is excellent as emitting ability and low erosion at ion bombardment. Tantalum is deformable and better workability. Fabrication of filaments and spherical segments in the user place is easy to be realized from tantalum. For technological guns as emitter material often the choice is LaB₆. Their not very high working temperature is advantage for a decrease of the heating power, but condensation of the evaporated or sputtered refractory metal on the emitter surface decreases its electron emission. As a result LaB₆ emitters are not implemented in EB welding systems for joining refractory metals.

The real diode system is demonstrated voltage current characteristic, different than shown on Figure 46a (see Figure 46b). At values of voltage $U_a \leq 0$ there are currents (one can see region of initial currents, due to Maxwell distribution of the velocities of the emitted electrons). At big voltage values are observed so called Shottky effect at which the emission of electrons is controlled by decrease of $e\phi$ due to the outer electric field.

This is not auto-electron emission (at electrical field of order $10^8 - 10^9$ V/m this is possible only on a tip – than the potential barrier is too narrow and tunneling transition of free electrons become possible; i.e. at auto-electron emission not need of emitter heating). Due to smooth transition between regions of “3/2 law” and of “saturation of thermo-emission current” in the voltage-current characteristics (mainly as result of Shottky effect, but also due to the real roughness and to the non-uniformity of the emitter surface) the real characteristics of emission current not obey exactly the theoretical equation (93) for saturated emission current density.

Due to Shotky effect the equation (93) can be written as

$$j_e = A \cdot T_c^2 \exp\left(-\frac{e\phi}{kT_c}\right) \cdot \exp\left[\frac{e}{kT} \sqrt{\frac{eE}{4\pi\epsilon_0}}\right] . \quad (94)$$

In the region of control of space charge the voltage-current characteristic also do not satisfied exactly the Child equation, that for flat diode (with distance between electrodes d and initial velocity of emitted electrons $V_0=0$) can be written as:

$$j = \left(\frac{4\epsilon_0}{9}\right) \sqrt{\frac{2e}{m}} U_a^{3/2} \cdot \frac{1}{d^2} = 2.334 \cdot 10^{-6} \frac{U_a^{3/2}}{d^2} . \quad (95)$$

Here d is measured in [cm] and j is calculated in [A/cm^2]. This equation is exact for emission of mono-energetic particles which initial velocities are equal to zero. The space charge of emitted electrons significantly affects the potential distribution near the cathode surface and could produce a potential minimum in the vicinity of the emitter. The maximal emitted current, limited by space charge is that, which is limited by potential distribution drop between diode electrodes, leading that on the emitter plane the potential gradient (i.e. electrical field) have zero value instead the uniform gradient of potentials between these two flat electrodes if the diode is situated in vacuum. In the case of emission of electrons with distributed initial velocities some electrons will be able to go to the anode at 0 or at stopping electrical field in front of the cathode. So a difference of the real emitted current take place. The problem in the case in which the charged particles are emitted with Maxwell velocity distribution had been solved by Langmuir [25].

For evaluation of universal function of potential distribution in front of a cathode one can assume dimensionless coordinates. Let dimensionless potential is:

$$\eta = \frac{U_a - U_{\min}}{kT_c / e}, \quad (96)$$

where $U_{\min}$ is the minimum of the potential. The dimensionless distance from the cathode can be written as:

$$\xi = 2\beta(z - z_{\min}). \quad (97)$$

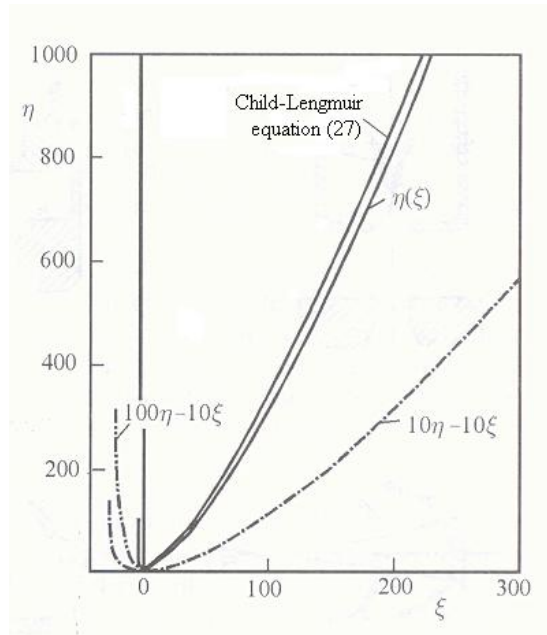


Figure 50. Potential distribution in the case of limited by space charge electron beam are emitted with Maxwell distribution of velocities of the electrons. The function is given in dimensionless parameters potential vs. distances $\eta(\xi)$ as they are defined in eq.(96) and eq.(97)

There z_{min} is the distance between cathode and potential minimum; β is a function of current density j and emitter temperature T_c , given by equation

$$\beta^2 = \frac{1}{2\varepsilon_0} \sqrt{\frac{\pi \cdot m}{2e}} \cdot j \cdot \left(\frac{kT_c}{e}\right)^{-3/2}. \quad (98)$$

The function $\eta(\xi)$ is tabulated (and/or available in the form of approximations) for two regions $\xi \leq 0$ and $\xi \geq 0$ (distances before and after the potential minimum-see Figure 50).

Current (limited by temperature), can be assumed as part of the maximal emitted current, evaluated by (95). If in the front of the cathode exists a stopping electrical field generated by a potential U_r due to the Maxwell velocity distribution of the emitted electrons, the current density is:

$$j = j_s \exp(eU_r/kT_c). \quad (99)$$

The electrons height of potential barrier is $U_c - U_r$ and using (96) one can find:

$$j/j_s = \exp(-\eta_c) \quad (100)$$

and

$$\eta_c = \ln(j/j_s). \quad (101)$$

Equation (9) is an initial condition. From η_c and $\eta(\xi)$ for negative values of ξ one can find ξ_c , than β from (98) and at known distance anode-cathode the dimensionless position ξ_a :

$$\xi_a - \xi_c = 2\beta \cdot d, \quad (102)$$

as well as the dimensionless potential η_a .

In that way at assumed part of saturated current density the anode potential is found. Assuming many values of j one can calculate the corresponding values of U_a and draw the voltage-current characteristics.

Geometry of the Welding Electron Gun Electrodes

The electron optical systems have been developed from the times of the design of the first electron microscope (E.Ruska,1931) and X-ray devices, as well as of the various electron tubes, being used in television, telecommunications and radars. The electron guns used for welding are characterized with an power of range 1-100kW and bed vacuum conditions in the

draft region, where is transferred the generated beam. But beams must be narrow and transfer energy to a considerable distance.

For creating an intense beam thermionic cathode is heated directly or indirectly by thermal radiation or by electron bombardment. Sometimes the thermionic cathode is replaced by a cold, secondary emitting cathode or by a plasma boundary, from which are extracted plasma electrons, but current density is lower and there plasma emission will be not discussed.

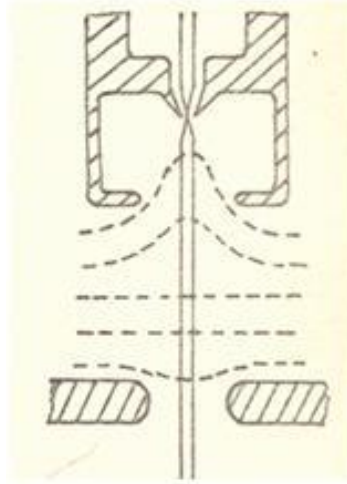


Figure 51a. Diode gun creating convergence electron beam (design of Steigerwald)

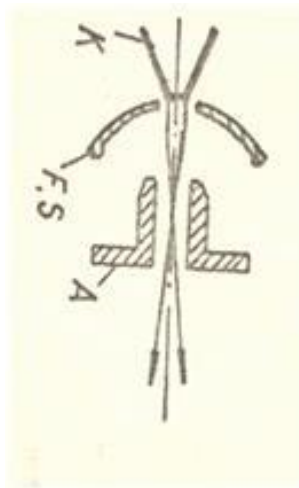


Figure 51b. Electron gun (design of Rogovsky). K and A are cathode and anode; F and S are Wehnelt or Control electrode. If the potential on S is changing-the gun is triode, if $U_f = U_k$ there are a diode gun and electrode is signed as F

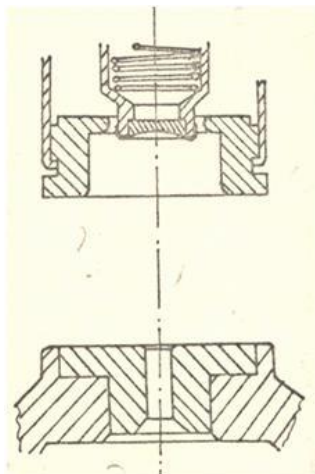


Figure 51c. Triode electron gun with LaB6 emitter and heater (tungsten spring)

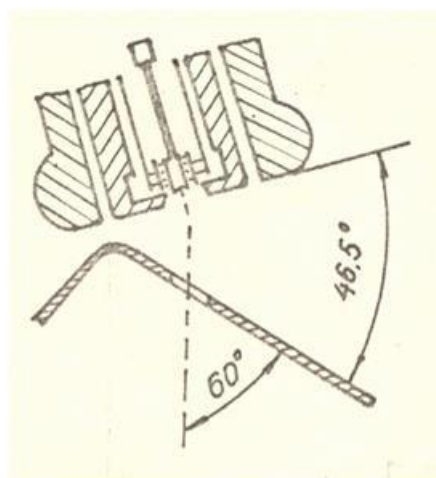


Figure 51d. Diode gun (design of Bas) Tilted electrodes protect emitter (front part of the bolt cathode) from ion bombardment.

The acceleration of the ejected from the cathode free electrons (owning negligible initial velocities) and the formation of the beam are fulfilled from an electrical field being generating in front of the cathode surface. Dependently from the potential distribution (number of metallic electrodes with various potentials), creating that field, one can distinguish diode or three-electrode electron guns – see Figure 51. That part, accelerating the electrons and shaping the beam in electrical field, as were mentioned, is called electrostatic part of the electron optical system of the electron gun. Additional parts for assembling real operating welding electron guns are magnetic focusing and deflection coils. Usually it can be achieves the required beam property with one magnetic focus lens. Additional coils are used for adjustment the geometry and electromagnetic axes of gun, to avoid aberrations and asymmetry. The control of electron beam spot on work piece is done by deflection coils. High frequency of deflection coils is advantage. In the technology applications are used usually electron guns with high value of the perveance.

A base approach for design of the electron gun, generating intense beam is proposed by Pierce (Figure 52). This gun obeys straight line electron trajectories. Idea of that approach come from observation of an initially unlimited beam (formed in a parallel planar diode), or convergent beam (generated in a cylindrical or a spherical electrode configuration).

If one chose a part of these beams to work as an actual beam and the outer parts of the virtual unlimited initial beam are replaced by the electrodes with suitable potentials and positions, that not change the balance of electrostatic forces and the boundaries of the chosen part of the beam. So the potential distribution on boundary of designing beam will be the same as in the unlimited beam; the derivative of the radial component of the potential there will be equal to zero (that means lack of extension of the beam) and the distribution of the potentials out of beam will be controlled by Laplace equation. The electrode configuration, obtained at such approach are given on Figure 53.

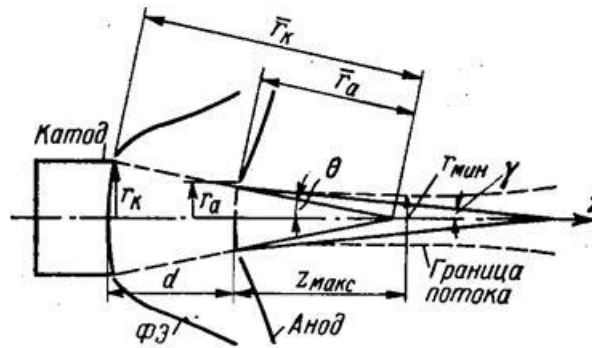


Figure 52. Geometry of the electrodes and the beam in a Pierce electron gun

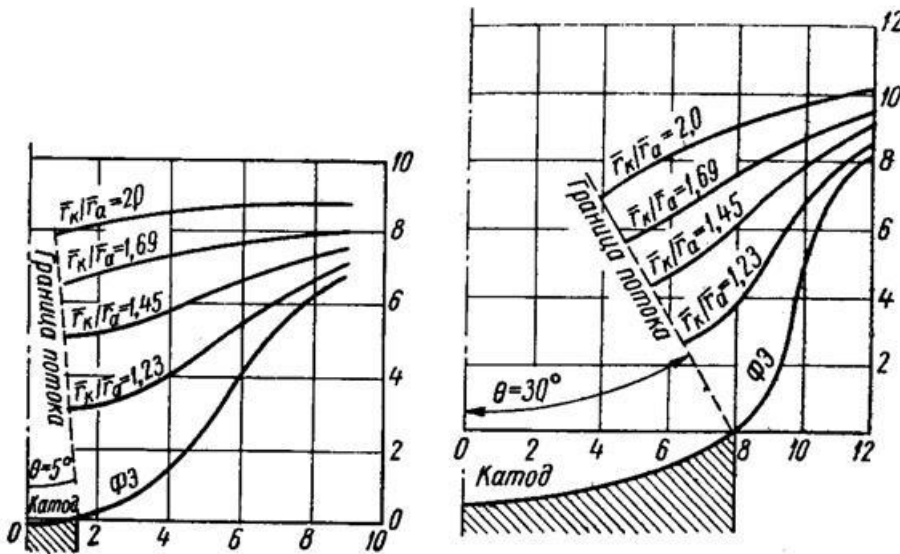


Figure 53. Electrode profiles of electrostatic part of Pierce electron guns with angle of beam convergence 50 and 300

There the convergent beam is forming in a part of a spherical diode, outer electrode of which is cathode, and inner electrode-anode. The angle of boundary (envelope) electron trajectories of designing beam is 2Θ . Such a beam is obtained as a cone part with space angle 2Θ at tip. The effect of space charge of removed part is replaced by a electrostatic field , generating of a focusing electrode and an anode that are suitable shaped. The profiles of that electrodes, presented in Figure 53 are obtained through modeling in an electrolyte bas at chosen two angles Θ and various ratios between the radii of the spheres of the emission surface R_c and the anode surface R_a . It can be noted that angle between boundary trajectory and non-emitting part of cathode (usually called focusing electrode) is about 67.5° . As a rule these difficult for machining profiles are replaced by approximating cone segments.

The relation between the parameters of electrostatic part of designing electron-optical system Θ and R_c/R_a and beam current I_b as well as the accelerating voltage U_a are given by equation for a spherical diode:

$$I_b = 29.34 \cdot 10^{-6} \frac{\sin(\Theta/2)}{[\alpha(R_c/R_a)]^2} U_a^{3/2} \quad .(103)$$

Here $\alpha(R_c/R_a)$ is a function, shown on Figure 54. The de-focusing effect of anode diaphragm is evaluating as that of defocusing lens (Figure 54). On the next figure (Figure 55) are shown the dependence of the angle of boundary (envelope) electron trajectories at the out of electrostatic part of the gun γ on the θ . It can be seen that at $R_c/R_a > 1.45$ at output of electrostatic part of the gun will be formed convergent beam with $\gamma < \theta$. That angle, as is seen on Figure 56 is function also of the perveance value p , because p is defined by θ and (R_c/R_a)

The behavior of the boundary (envelope) trajectories after anode, in the case of absence of electrical and magnetic fields there, are function of θ and the ratio (R_c/R_a) . The maximum distance to the minimal cross-section of beam (crossover) can be obtained at $(R_c/R_a) \approx 2.2$. The optimal angle of convergence of boundary (envelope) trajectories in such a gun is $\Theta \approx 0.37 \sqrt{p}$, if it is measured in radians or $\Theta \approx 21 \sqrt{p}$ at measuring that angle in degrees. Respectively $\gamma_{opt} \approx 0.16 \sqrt{p} [rad]$ or $\gamma_{opt} \approx 9.15 \sqrt{p} [deg]$. In that case $z_{min} \approx R_c$ and $r_{min} \approx 0.2 R_c$.

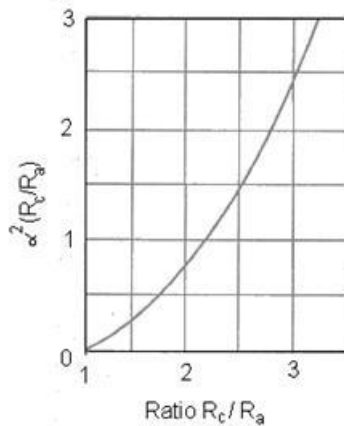


Figure 54. Dependence of $\alpha^2(R_c/R_a)$ vs. ratio (R_c/R_a)

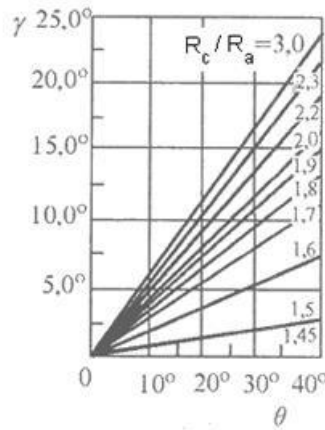


Figure 55. Dependence $\gamma(\theta)$ at various (R_c/R_a) from 1,45 to 3

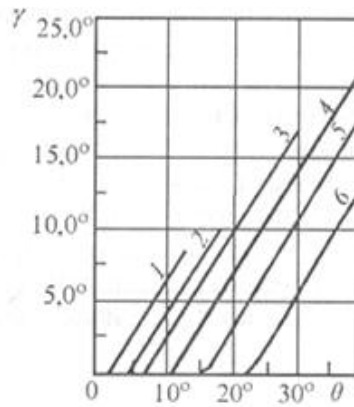


Figure 56. Dependence $\gamma(\theta)$ at various p : 1-0,063, 2-0,316, 3- 0,732, 4-1,58, 5-3,16, 6-7,32

At increase of perveance at constant convergence angle the ratio R_c/R_a decrease-so at constant R_c the accelerating electrode (anode) must come close to cathode. At that the defocusing effect of the anode diaphragm increases. When anode diaphragm come nearer, the electrical field in vicinity of cathode changes and distribution of emitted current become non-uniform, being lower in the central part of the cathode. Effect is stronger at bigger angles of convergence of envelope electrons 2θ . In the same time the number of electrons bombarding anode and the aberrations of the anode diaphragm are increased. The actual perveance of such gun is less than the calculated one. It is assumed that limiting ratio $d/2R_a$ is 0.7 for the applicability of the Pierce approach. In the powerful technological electron guns, the perveance of which is $p=1-2 \cdot 10^{-6} \text{ A} \cdot \text{V}^{-2/3}$ that ratio is of order of that limit and the mentioned no desirable effects take place. Aiming to avoid the non-uniformity of cathode current emission and losses of the electrons bombarding the anode, the shape and the distances of the electrodes can be corrected. As a result the profile of the beam is not as was calculated and the beam trajectories are not straight lines.

Much more universal are the approach of heuristic choice of electrode design and analysis of beam parameters by computer simulation, that will be discussed bellow, after the next paragraph.

The Design of the Electron Gun Additional Parts

High working temperatures of these emitters lead to necessity of considerable powers for heating of cathode. There is involved also design requirements for obtaining a low heating of the current inputs. The result is in the heavier cases of EB welding and melting guns to be applied electron bombardment for the heating of a high power and high brightness beam emitter. This lead to need of additional high voltage (1-3kV) current input and high potential ($V=V_a$) power source.

To improve radiation losses some constructive elements of the cathode construction are used as radiation screens. Heating of such block-cathodes is done by tungsten filament, analogically to radiation and conduction heating of the indirect heated cathodes. The shape and dimensions of heating spiral are determined by requirement for uniform distribution of temperature on emitting surface. A simplified evaluation of radiation losses can be done using the data for specific heat radiation values versus working temperature, given by eq.(104). For LaB₆ surface the radiation efficiency η_t (namely reduction coefficient of radiation losses in comparison with the heat radiation losses of black body is approximately 0.7; for tantalum surface that value is 0.426). In evaluation is used upper limit of working temperature for chosen emitter and outer surface of cathode block. The radiation screens can be estimated by Stephan-Boltzmann equation for radiation heat losses:

$$P_{\text{rad}} = \eta_t \cdot 5.64 \left[\left(\frac{T}{1000} \right)^4 - \left(\frac{T_0}{1000} \right)^4 \right]. \quad (104)$$

Here T_0 is the temperature of surrounding parts (namely radiation screen).

At calculations only the outer surfaces of cathode block are taken in the account, assuming that the inner walls radiation is adsorbed by opposite walls and that heat losses by thermal conductivity of electrical inputs and assembling elements are negligible.

The design criteria are minimal desired power for obtaining the working temperature and uniform distribution on emitting surface.

In the case of indirect heating of beam emitter, the heating filament is calculated similarly to the directly heated emitters. The ideal heater must be with uniform physical properties, chemical composition and exploration conditions. The use of high temperature electric isolation materials at working temperatures of LaB₆ and especially of pure emitting metals is practically impossible. If the role of more cold ends of the heating filament are negligible (at long filaments) and if is assumed equal temperature along whole filament length l_h in the case of tungsten wire of diameter d_h , one can write that power radiation P_h from such filament is:

$$P_h = P_1 \cdot d_h \cdot l_h, \quad (105)$$

and the heater resistance R_h is:

$$R_h = R_l \frac{l_h}{d_h^2}, \quad (106)$$

where P_l and R_l are respective values, evaluated for a cylinder of diameter 1cm and length 1cm. The current of heating filament is:

$$I_h = I_l d_h^{3/2} \quad (107)$$

and voltage on its ends is:

$$U_h = U_l \cdot \frac{l_h}{d_h^{1/2}}. \quad (108)$$

The current emitted by such filament is:

$$I_s = I_{s1} l_h \cdot d_h. \quad (109)$$

For evaluation the rate of evaporation M of such a heated wire, measured in [g/s] one can write:

$$M = M_1 \cdot d_h \cdot l_h. \quad (110)$$

In Table 12 are given the data of W filament, designed as cylinder of diameter 1cm and length 1cm.

Than, using data from Table and choosing the working temperature one can calculate filament with any power or emission current.

The lifetime of Tungsten filament can be evaluated as:

$$t = 8.45 \cdot 10^{-3} \frac{d}{M_1} \frac{1 - q^\beta}{\beta} [\text{h}], \quad (111)$$

where q is the ratio of diameter of filament in the end of lifetime to the initial diameter, β is coefficient defining by temperature of filament and exploitation conditions. In the case of keeping the constant temperature during all time of exploitation $\beta=1$. Usually constant is one of electrical parameters is keeping constant, and lifetime is limited by evaporation (5%-10% decrease of the diameter) or by decrease of emitted current (up to 80% of its initial value) in the case of cathodes with electron bombardment heating. The values for β are given in Table 13 (for working temperatures in range 2300-2600K).

Table 12. Data for design of electron bombardment heating filament from W

		2400K	2500K	2600K	2700K	2800K
1	P_1 [W.cm ²]	181.2	219.3	263	312.7	368.9
2	R_1 [10 ⁶ .Ω .cm]	89.65	94.13	98.66	103.22	107.85
3	I_1 [A.cm ^{3/2}]	1422	1526	1632	1741	1849
4	U_1 [10 ³ .V.cm ^{-1/2}]	127.5	143.6	161.1	179.7	199.5
5	I_{s1} [A.cm ⁻²]	0.364	0.935	2.25	5.12	11.11
6	M_1 [g.cm ⁻² s ⁻¹]	1.37.10 ⁻⁹	6.23.10 ⁻³	2.76.10 ⁻⁸	9.95.10 ⁻⁸	3.51.10 ⁻⁷

Table 13. Data for determination of coefficient β at various exploitation condition

Filament parameter, kept constant	β	Relative lifetime normalized to regime T=Const.		
		q=0.95	q=0.9 Is t /	Is = 0.8
Voltage of the filament	-5.46	1.18	1.43	0.244
Current of the filament	33.9	0.49	0.286	-
Power of the filament	9.14	0.82	0.68	0.218
Emission current	2.63	0.96	0.92	-

The real heating filaments have ends with decreased temperature. This change the real heater parameters - the current increase and the voltage decrease. Also the actual emitted current and the energy losses by radiation are lowered. In such a calculation unfortunately are not taken in the account the re-crystallization of the filament material, as well as local superheating due to other reasons.

Computer Simulation of Technological Electron – Optical Systems

1. Trajectory analysis of the beam formation in electron guns

The progress in electron beam technologies requires further improvements to the design as well as optimization of the electron guns, producing intense beams. In this respect the computer simulation of formation of the beams is a powerful means to analyze and optimize electron-optical systems of the technology electron guns.

In most of computer programs a general algorithm is used (see Figure 57) enabling the potential field, electron trajectories as well as the space charge distribution to be self-consistently obtained. Its basis steps are: (i) Dividing of discrete parts of the appropriate boundary conditions and the space of gun for calculation of electrostatic potential distribution by means of suitable mesh system; (ii)solution of Laplace's equation; (iii)calculation of the emission current density applying the law of Child-Langmuir to the virtual elementary diodes in the vicinity of the cathode emitting surface;(iv) calculation of a finite number of electron trajectories through the obtained electric field; (v)allocation of the space charge carried by separated trajectories to the grid nodes; (vi)solution of Poisson's equation for the newly determined space charge distribution; (vii)reiteration of above procedure from step (iii) to

step (vi) until a self consistent solution(stable values of potentials, electrical field and current and position of separated trajectories) is obtained.

Recently several authors proposed a number of improvements concerning these basic steps. Kasper [26,27] developed a space charge allocation method based on analytic formula for the space charge density and local divergence or convergence of the beam. Kumar and Kasper [28] proposed incorporation of new version of finite-difference method and interpolation procedure for calculation of the electric field in space charge limited electron beam. The thermal velocities of electrons and possible distinct appearance of the potential minimum in front of the cathode (virtual cathode) are also included in their theory. Weber [29], Ninomiya [30] and Monro [31] improved taking into account the thermal velocity effects on beam formation. Van den Broek [32] developed a method in which the cathode current is evaluated using Langmuir's law instead of Childs law and the space-charge density is calculated with a fitting technique. All these improvements substantially increase the accuracy and adequacy of the simulations.

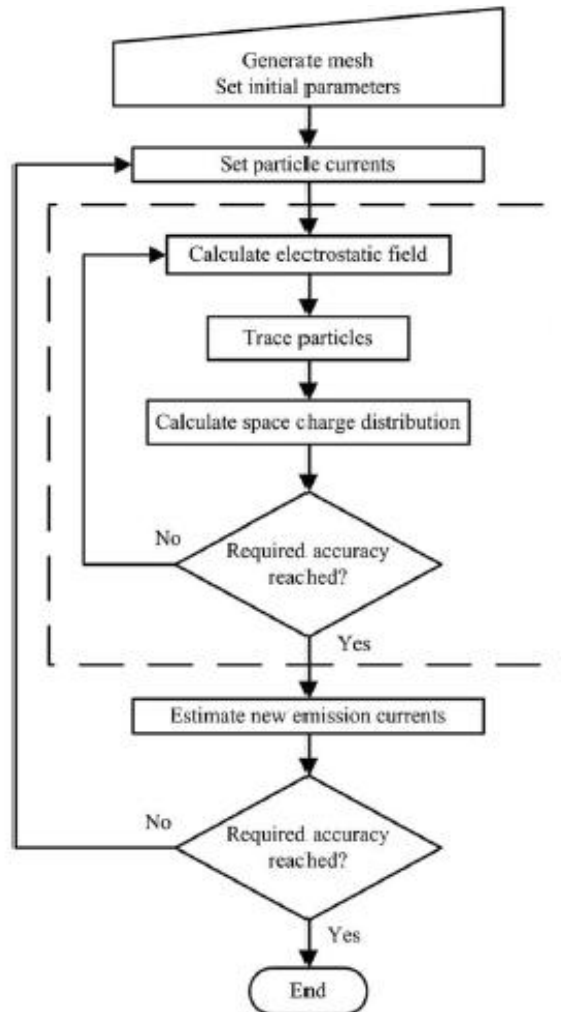


Figure 57. Flow chart of beam trajectory and current computer simulation

Information of such numerical experiments and interpretation of the data is performed mainly by analyzing the trajectory (ray)tracing. The adequacy obtained results are determined considerably by choice of region for calculation of the potential distribution and boundary conditions, division of the region of calculations on sub-regions and accepted net steps values.

Mathematically, the trajectory analysis models can be described by the following basic equations:

Poisson's equation governing the electrostatic potential U relatively to the space charge density ρ in axially-symmetrical beam is given by:

$$\nabla^2 U = - \frac{\rho}{\varepsilon_0} , \quad (112)$$

where ∇^2 is the Laplace operator in cylindrical coordinate system, ε_0 is the vacuum dielectric permittivity. Due to cylindrical symmetry, the potential need only to be determined in a half plane of the gun electrode configuration (from $r=0$ to $r=r_{\max}$).

Two types of boundary conditions in addition to (112) render the problem well posed:

Neumann boundary condition along the axis of the region considered (i.e. the radial

component of the potential distribution $\frac{\partial U}{\partial r} = 0$) and Dirichlet boundary conditions

along the rest of the boundary (potential in the end points of the mesh is potential of electrodes U_j ; in the gaps between electrodes the potential is assumed to be distributed logarithmically in radial direction and linearly if boundary is chosen parallel to the axis z).

The motion of the electrons in these conditions is given by the Newton equations:

$$\frac{d}{dt}(m \cdot \dot{q}_i) = -eE_i , \quad (113)$$

where q_i are the coordinates (namely $q_1 = x$, $q_2 = y$, $q_3 = z$); $\dot{q}_i = \frac{dq_i}{dt}$ are the electron velocities

and $E_i = - \frac{\partial U}{\partial q_i}$ are the components of the electrical field in that point, evaluated in the

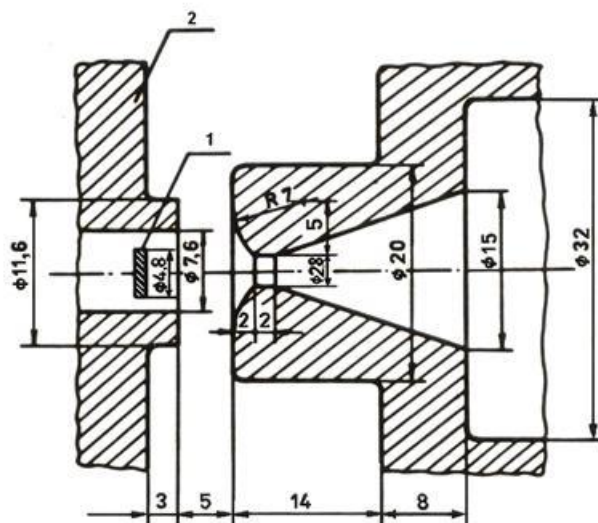
directions ($i=1,2,3$). In (113) are assumed (i) that the beam is non-relativistic ($m=\text{Const}$) and (ii) self-magnetic field of the beam is negligible.

The solution of differential equation (21) after exclusion of t to obtain trajectories of particle motion can be find using standard Runge-Kutta methods. In same cases for the increasing the accuracy of calculations near the cathode and/or electrodes a net with more fine pitch is required.

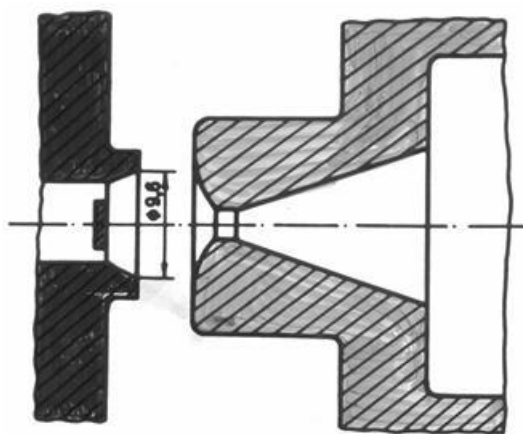
Instead many thousands electron trajectories (due to the limited computer resources) the calculated tracks are usually restricted to some tens. For that are used virtual big charged "particles" containing the charge of emitted by a cathode segment current (methods of cell or current tube method [26,27]). The current, obeying Child-Langmuir's equation is determined

for every near to flat plane diode in vicinity of a chosen in that way cathode segment. After that is carried out allocation of the space charge transferred by the calculated trajectories to the net nodes. The next step is solution of the Poisson's equation for the newly determined space charge distribution. At repeatedly reiteration of above procedure are obtained beam simulation results, describing complex electron gun characteristics, utilizing as an base at experimental improvement of its design.

As one example let us show two EBW guns (electrostatic parts) with indirect heating LaB₆ cathodes and very similar electrode geometry—see Figure 58.



a)



b)

Figure 58. Geometry parameters of two electrostatic partsof EBW guns, The difference is only cylindrical or conical inner wall of the control electrode

The emitter is from La_6B tablet. All dimensions of electrodes(emitter, control electrode and anode) are the same. The difference is only in shape of control electrode shape-in variant a) this wall is a cylindrical one, as well as in the case b) there are a conical shape.

Results of trajectory analysis of generated beam at various voltages on control electrode M are shown on Figure 59 and Figure 60. The accelerating voltage K-A is 30 keV in all cases.

At comparison of beams shown on Figure 59 and Figure 60a one can understand qualitative character of the trajectory analysis. Every trajectory presented carry different electron current and exact comparison of beams after mixing the trajectories originating from central emitter area and from emitter periphery is impossible. In ref. [33] are calculated statistical values of emittance and brightness at distances z equal to 3,4 and 5 cm from the emitter surface (for three control voltages : 0,-500 and -1500 V) and definitively the second case of gun (caseFigure 58b) with conical inner wall of control electrode was chosen due to lower emittance values and bigger brightness.

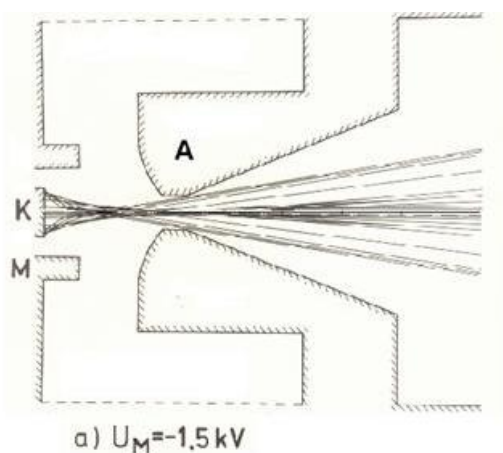


Figure 59. Trajectory analysis of electrostatic part of EBW gun shown on Figure 58 a) at control electrode(M) potential -1,5 kV to the emitter electrode (K)

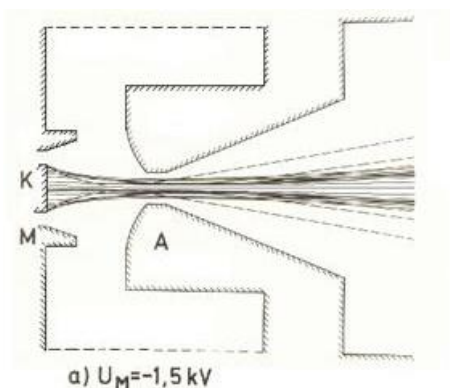


Figure 60a. Trajectory analysis of electrostatic part of EBW gun shown on Figure 58 a) at same conditions as Figure 59.

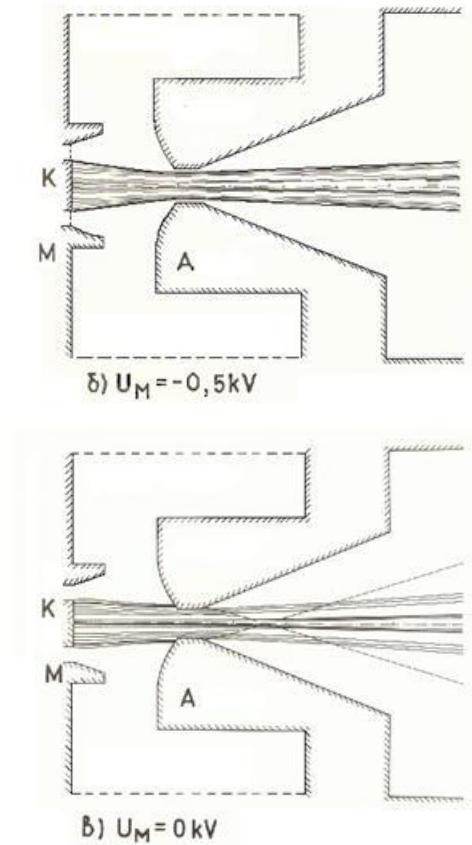


Figure 60b,c. Trajectory analysis of the generated beam in electrostatic part of EBW gun shown on Figure 58 b) at control electrode voltage -0,5 kV and 0 kV

The welding system used at Leybold-AG, Hanau, (now PTR GmbH, Dörnigheim) is a triode electron gun with two focusing coils and a deflection system [34]. The gun itself has a small square directly heated cathode, situated in the circular aperture of a Wehnelt electrode, which is biased negatively with respect to the cathode with voltages of -300 to -3000V . At -300V a maximum of current is drawn, while at -3000V the electron emission is suppressed completely.

The basic approximation is the assumption that a round cathode in the simulations will give results, which agree well with experimental results obtained with a square shaped cathode. While this has turned out to be true, the explanation may be seen in the 3D interpenetration of electrical fields, which is stronger at the edges of the square cathode, hence reduces there the emission. By this effect the cathode will be “effectively round”, which then becomes obvious by the shape of the beam spot on the work piece.

The second problem is that due to the limited mesh resolution the cathode becomes invisible and the results are questionable. The simulation of equipotential lines shown on Figure 61 is used to calculate a field-line (black and dashed) between Wehnelt and Anode in order to cut out the cathode part, using this field line as a slanted and curved Neumann boundary for the simulation in Figure 62.

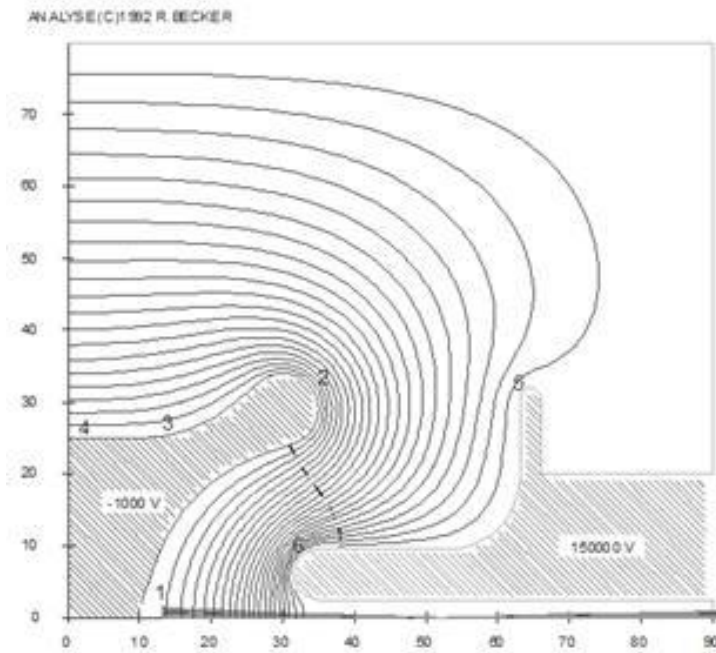


Figure 61. Calculation of electric field distributions in the famous Steigerwald electron gun, used in former times as EBW gun [34]

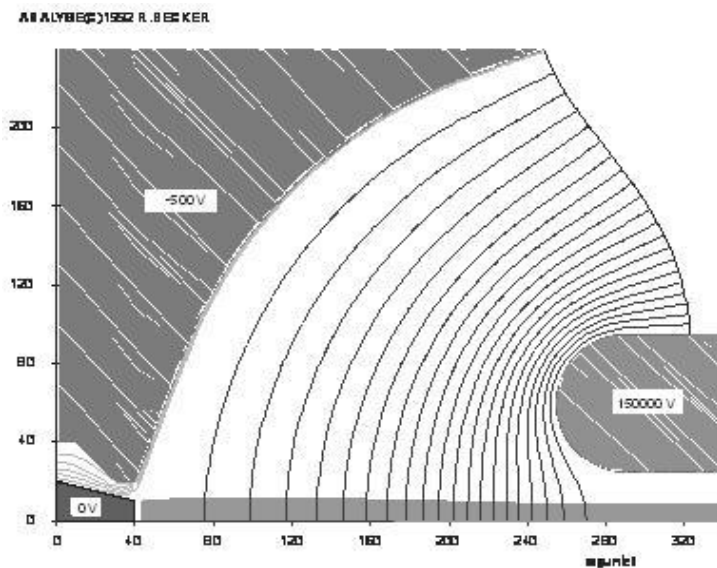


Figure 62. Calculation of electron emission in the gun part of Figure 61 with 10 times smaller mesh size, using the curved Neumann boundary, shown in Figure 61 to close the boundary

A more detail explanation of the problem follows. For computer simulations with a finite difference method (FDM) Poisson solver such a gun presents a substantial difficulty, because a cathode of typically 1 mm radius is situated in a anode housing with a radius of 80-100 mm.

A good simulation of the electron beam, however requires that the mesh size is much less than the cross-over radius of the beam, which is in the order of 1/10 mm. In response to this, about 10 000 meshes will be needed in radial direction. This is impossible, even for to days fastest PCs and only attainable on super computers, not everywhere available. The problem can be solved principally by a non-uniform mesh, best introduced by a logarithmic transformation of the radial coordinate [35]. This procedure, however, needs too many program modifications for well established programs, while developing a new program, which incorporates such a transformation, will require too much development to include all required features of well established programs. For the existing programs of the EGUN family it has been simplest to subdivide the problem by the calculation of a field line in an appropriate position – see Figure 61 – and to use this field line as a slanted and curved Neumann boundary for the calculation of the cathode part of the gun. The field line is written on a file with proper syntax for the direct inclusion into a input file. For the inner part of the whole gun the position and kind of curvature of this Neumann boundary represents all electrostatic influences from the much larger outer part.

From Figure 61 to Figure 62 the mesh size has been reduced by a factor of 10. Only by this the close vicinity of the strip cathode inside the bore of the Wehnelt electrode becomes visible. The trajectory end data from this calculation then can be used to calculate the beam through the lens and deflection system, which will not be performed here. Another point is more important for the optimization of such a gun. This is the reduction of surface fields on those surfaces where electrons could start and be accelerated to full power. A program provides a special tool for this, consisting of a plot of the geometry in connection with a plot of potentials and surface fields shown in Figure 63.

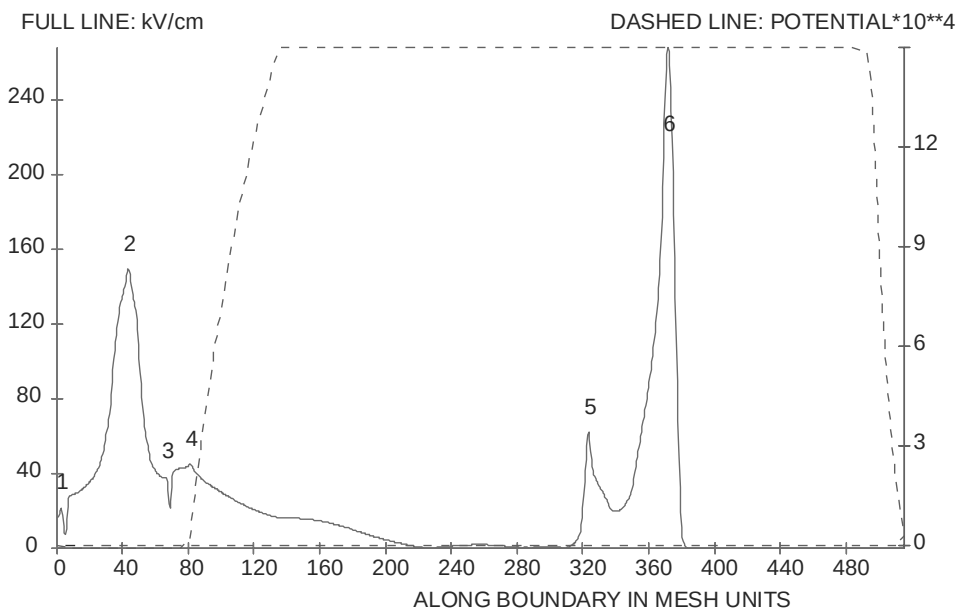


Figure 63. Electric field (full) and potential (dashed) along the boundary with numbers indicating maxima of surface field, synchronized with Figure 61, showing their locations. Dangerous for sparking is maximum No 2, because electrons from there will be accelerated to the anode

From an inspection of Figure 61 and Figure 63 it becomes clear at once, that the field maximum No. 2 on top of the Wehnelt electrode of about 160 kV/cm should be reduced by increasing the radius of the electrode curvature there. The field maxima No. 5 and 6 are located at the anode and do not need cure, because no electrons can be accelerated from there. By removing the anode disk and increasing the radius of curvature at the Wehnelt tip the surface field strength could be reduced to about 60 kV/cm. This improvement has been essential for the continuous welding of aluminum parts over more than 100 hours.

The problems of computer simulation of electron guns with point (hairpin) emitters are typical for analyzing electron beams but also in low power EBW and EB machining guns such emitters are utilized. As example space charge limited emission from the emitting tip of a shaped as “V” tungsten direct heated wire was simulated in [41]. There, after a suitable choice of the suitable calculation mesh the current density emitted from such thermionic emitter surface is iteratively established from the potential distribution near this surface.

Instead conclusion it can be seen that an inherent drawback of the trajectory analysis is its qualitative character. From the representation of the beam as a set of trajectories not a single quantitative characteristic of the beam structure which is of paramount importance in technological applications can be found. As individual trajectories carry different space charge it is difficult to evaluate their contribution to the beam formation as well as to study how the structure of the beam as a whole evolves along its axis.

2. APPROACH OF PHASE ANALYSIS OF THE BEAM FORMATION

The main features of the proposed by the author and collaborators approach [36,37] are as follows.

The thermal velocities distribution and formation of a potential barrier in front of the cathode are taken into account.

A new method (the so called phase-space method) for calculating the space charge density and its allocation is used.

The implementation of a phase space concept, i.e. phase analysis instead of the commonly used trajectory analysis of individual or “quasi”-individual particle tracks.

The physical model of our software package is as follows. The potential distribution is calculated again in the domain of gun electrode configuration with a boundary composed of gun axis (Neumann boundary condition), cathode surface, electrodes and suitable inter-electrodes segments with suitable distributed potentials (Dirichlet condition).

A beam of electrons in a static electromagnetic field including space charge can be described by a six-dimensional phase space density $f(x, y, z, p_x, p_y, p_z)$, where p_x , p_y and p_z are components of the momentum of an electron at a point (x, y, z) . The space-charge density at an arbitrary point of coordinates (x, y, z) is:

$$\rho(x, y, z) = e \int f(x, y, z, V_x, V_y, V_z) dV_x dV_y dV_z \quad .(114)$$

Phase space conservation (Liouville's theorem) yields:

$$dx dy dz dp_x dp_y dp_z = dx_0 dy_0 dz_0 dp_{x0} dp_{y0} dp_{z0} ,$$

and

$$f(x, y, z, p_x, p_y, p_z) = f(x_0, y_0, z_0, p_{x0}, p_{y0}, p_{z0}),$$

if the point $(x_0, y_0, z_0, p_{x0}, p_{y0}, p_{z0})$ in the phase space transforms into (x, y, z, p_x, p_y, p_z) by electron motion. Then (114) can be written:

$$\rho(r, z) = \frac{1}{J \cdot V_z} \int V_{z0} \cdot f_0(x, y, z, V_x V_y V_z) dV_{x0} dV_{y0} dV_{z0} , \quad (115)$$

where $r = (x^2 + y^2)^{1/2}$, J is Jakobian matrix of the transformation between x_0, y_0 and x, y .

Thermal electrons emitted from the cathode obey the Maxwell-Boltzmann's law, therefore the phase density on the initial plane (i.e. on the cathode) will be

$$f(x_0, y_0, z_0, V_{x0}, V_{y0}, V_{z0}, z=0) = \frac{K_1 \cdot j_s}{T^2} \exp\left[-\frac{K_2 (V_{x0}^2 + V_{y0}^2 + V_{z0}^2)}{T}\right] , \quad (116)$$

being j_s the saturation current density of the cathode, $K_1 = m^2 / 2\pi e k^2$, $K_2 = m / 2k$, e and m the electron charge and mass, respectively, k the Boltzmann's constant, T cathode temperature and V_{x0}, V_{y0}, V_{z0} , the components of the initial electron velocity.

Space charge density at a point (r, z) caused by electrons being emitted from an elementary cathode area $dx_0 dy_0$ with initial velocities in the range of $(V_{x0} - V_{x0} + \Delta V_{x0})$, $(V_{y0} - V_{y0} + \Delta V_{y0})$ and $(V_{z0} - V_{z0} + \Delta V_{z0})$ that are energetically to pass the potential minimum in front the cathode is

$$\begin{aligned} \Delta \rho(r, z) = \frac{j_s}{4J V_z} \{ \exp(-K_3^2 V_{z0}^2) - \exp[-K_3^2 (V_{z0} + \Delta V_{z0})^2] \} \cdot \\ \cdot \{ \text{erf}[K_3 (V_{x0} + \Delta V_{x0})] - \text{erf}(K_3 V_{x0}) \} \cdot \\ \{ \text{erf}[K_3 (V_{y0} + \Delta V_{y0})] - \text{erf}(K_3 V_{y0}) \} , \end{aligned} \quad (117)$$

where $K_3 = (m/kT)^{1/2}$ and J is determinant of the Jacobian matrix of the transformation between dx, dy and dx_0, dy_0 . The axial velocity V_z of electrons at a point (r, z) derived from energy conservation law is

$$V_z = [2 \frac{e}{m} U(r, z) - V_x^2 - V_y^2 + V_{x0}^2 + V_{y0}^2 + V_{z0}^2]^{1/2} \quad (118)$$

The motion of electrons is given by differential equations:

$$\dot{q}_i = \frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}; \quad \dot{p}_i = \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \quad (119)$$

where $i=1,2,3$.

In (119) p_i are the components of the momentum of an electron at a point with coordinates q_i (namely x, y, z). There are assumed (i) that the beam is non-relativistic and (ii) self-magnetic field of the beam is negligible. The Hamiltonian H for such a beam is given by:

$$H = \frac{(p_x^2 + p_y^2 + p_z^2)}{2m} = \frac{(p_{x0}^2 + p_{y0}^2 + p_{z0}^2)}{2m} + eU. \quad (120)$$

Here p_{x0}, p_{y0}, p_{z0} are the components of the initial momentum, and e is the charge of electron.

Therefore the equation for motion of non-relativistic electrons takes the form:

$$\frac{dV_i}{dT} = -\eta \frac{\partial U}{\partial q_i}, \quad V_i = \frac{dq_i}{dt}, \quad i=1,2,3 \quad (121)$$

where $\eta = e/m$ is the electron charge-to-mass ratio. The electron trajectories equation than is:

$$\frac{dV_i}{dz} = \frac{\eta}{V_z} \frac{\partial U}{\partial q_i}, \quad \frac{dq_i}{dz} = \frac{V_i}{V_z} \quad i=1,2. \quad (122)$$

Here the axial velocity of an electron ejected from a point (x_0, y_0) on the cathode, with initial thermal velocity components (V_{x0}, V_{y0}, V_{z0}) can be evaluated:

$$V_z = (2\eta U - V_x^2 - V_y^2 + V_{x0}^2 + V_{y0}^2 + V_{z0}^2)^{1/2} \quad (123)$$

$$\frac{dV_i}{dz} = \frac{\partial V_z}{\partial q_i}, \quad \frac{dq_i}{dz} = -\frac{\partial V_z}{\partial V_i}, \quad (124)$$

with initial conditions $q_i(z_c) = q_{i0}$ and $V_i(z_c) = V_{i0}$ determined at cathode plane $z = z_c$.

The potential distribution obtained through solution of Poisson's equation after approximation by series take the form:

$$U(x, y, z) = \sum_{k=0}^N g_{2k}(z) (x^2 + y^2)^k \cdot \sum_{i=1}^M b_{ik} [U(ih, jh) - U(0, ih)] \quad (125)$$

There $g_0(z) = U(0, 0, z)$; $g_{2k}(z) = h^{-2k}$, and h being the step of the grid used for calculation of the potential distribution; $U(ih, jh)$ is the potential in the grid point with coordinates $r = ih$ and $z = jh$ as the b_{ik} are constant coefficients. Than the axial velocity V_z can be re-written as:

$$V_z = \sum_{n=0}^2 \sum_{k=0}^{2-n} C_{nk} (x^2 + y^2)^n \cdot (V_x^2 + V_y^2 - V_{x0}^2 - V_{y0}^2)^k, \quad (126)$$

where the coefficients C_{nk} depend on the potential distribution and V_{z0} , $C_{00}=[g_0(z)+V_{z0}^2]^{1/2}$,

$$C_{01}=-\frac{1}{2}C_{00}; C_{02}=-\frac{1}{8}C_{00}^3; C_{10}=\frac{g_2}{2}C_{00}; C_{11}=\frac{g_2}{8}C_{00}^3; C_{20}=\frac{g_4}{2}C_{00}-\frac{g_2^2}{8}C_{00}^3.$$

Substitution of eq.(126) in eq.(124) yields :

$$\frac{dq_i}{dz} = -\sum_{n=0}^1 \sum_{k=1}^{2-n} 2kV_i C_{nk} (x^2 + y^2)^n (V_x^2 + V_y^2 - V_{x0}^2 - V_{y0}^2)^{k-1} \quad (127)$$

$$\frac{dV_i}{dz} = \sum_{n=1}^2 \sum_{k=0}^{2-n} 2nq_i C_{nk} (x^2 + y^2)^{n-1} (V_x^2 + V_y^2 - V_{x0}^2 - V_{y0}^2)^k \quad (128)$$

where $i=1,2$. The solution of Eqs.(127) and (128) can be found in the form (129), see [38]:

$$\begin{pmatrix} q_i \\ v_i \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} q_{i0} \\ v_{i0} \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} q_{i0} R_0 \\ q_{i0} V_0 \end{pmatrix} + \\ + \begin{pmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{pmatrix} \begin{pmatrix} q_{i0} W_0 \\ v_{i0} R_0 \end{pmatrix} + \begin{pmatrix} a_{15} & a_{16} \\ a_{25} & a_{26} \end{pmatrix} \begin{pmatrix} v_{i0} V_0 \\ v_{i0} W_0 \end{pmatrix}, \quad (129)$$

where $i=1,2$; $R_0=x_0^2+y_0^2$, $V_0=V_{x0}^2+V_{y0}^2$, $W_0=x_0V_{x0}+y_0V_{y0}$. Coefficients A obey the following set of differential equations:

$$A'_{1i} + 2C_{01}A_{2i} = 0, A'_{2i} - 2C_{10}A_{1i} = 0, i=1,2 \quad (130)$$

$$a'_{1l} + 2C_{01}a_{21} = \varphi_{1l}, a'_{2l} - 2C_{10}a_{1l} = \varphi_{2l}, \quad (131)$$

where $l=1,2,\dots,6$, $\alpha' = \frac{d\alpha}{dz}$ and $\varphi_{1l}, \varphi_{2l}$ depend on C_{nk} and A_{ij} .

The solution of Eqs. (129), (130) and (131)) as well eventually of Eqs. (127) and (128) allows the dependence of both the current coordinates and velocities on the initial their values to be obtained.

The initial velocities of electrons in the beam are determined by the temperature of the emitter and usually are in the range of energies less than 1eV. The space charge density is high and significantly affects the potential distribution. Often this space-charge cloud produces a potential minimum in vicinity of the emitter.

The distribution of the current emitted from the cathode is governed by the location and depth of the potential minimum in front of the emitter. Different sections of the cathode can function under three possible operating conditions: (i) initial currents' regime, (ii) space-charge limited flow and (iii) saturation or temperature limited mode. The emitted current is calculated by application of Langmuir's theory to the virtual parallel diodes in front of the cathode.

Having computed the Laplacian potential distribution the cathode surface is divided into $n_1 = R_c/h$ small annular regions, where R_c is the cathode radius and h is the mesh step. The emission current density as well as the location and depth of the potential minimum are then

calculated by considering each annular region as a small diode and applying the well known exact solution for planar geometry to each of them. The space charge of different energy groups and different cathode regions is computed. Allocating Δp to the grid nodes and summing the contributions of all energy groups and cathode regions the charge density ρ_{ij} in each node (ij) is obtained. In that way the number of initial conditions whose contribution to the electron beam formation is accounted.

Having computed the space charge density distribution the potentials are recomputed solving Poisson's equation and the whole process is repeated iteratively until the self consistent solution is obtained. The iteration technique applied for the determination of potential distribution by means of finite-difference method is the successive over-relaxation.

3. ILLUSTRATIVE NUMERICAL EXAMPLES OF ANALYSIS OF ELECTRON GUNS

Interpretation of the data obtained by the computer simulation of beam formation under implementation of the phase analysis is much more informative as compared to the trajectory analysis.

It was mentioned that in absence of coupling between motions in the xz and yz planes, the beam may be described by points distribution in two independent planes of phase space- x, x', and y, y'. For each phase plane, the area occupied by these points divided by π defines a two-dimensional emittance ε_x , (or ε_y) which is also a constant. In real beams the phase density distribution is neither uniform nor has a sharp edge. For this reason it is convenient to consider a set of concentric phase contours (on which the phase density attains a certain constant value) called "emittance diagram". Each contour encompasses a certain part of the beam and thus determines the emittance of the beam fraction considered.

In EBW the need of obtaining required power density distribution on the work-piece surface is of great importance. For example, formation of the cavity by the electron beam during electron beam welding is possible after reaching a critical power density). The critical power density p_{cr} depends on the thermo-physical properties of the solid material, mass and dimensions of treated details as well as on the type of the technological process (welding, drilling, melting etc). The region of electron beam in which power density exceeds the critical power density is called the electron beam active zone (EBAZ). As shown earlier both the configuration and dimensions of EBAZ strongly depend on beam parameters, namely total power, emittance; characteristic length of the beam and also on the critical power density determined by the specific features of the technological processes.

Figure 64 shows the emittance diagram of the beam in the initial cross-section i.e. in the plane of cathode in EOS with geometry given on Figure 58 and trajectory analysis shown on Figure 60a ($U_a=30\text{kV}$; $U_m=-1.5\text{kV}$). The modification of the phase contour enclosing 90% of the beam current is illustrated in Figure 64. This results from a linear transformation, one can see that the beam converges in a cross-section $z = 2.22$ and diverges in $z = 3.9$ cm. The cross-over of the beam is at $z = z_{cr} = 2.84$ cm. The emittance which corresponds to the most outward phase contour is $\varepsilon_{0.9} = 7.64 \cdot 10^{-6}$ m . rad. Integral invariants of the beam, namely normalized emittance $\varepsilon_{n,0.9}$ and normalized electron brightness B_u are $\varepsilon_{n,0.9} = 2,5 \cdot 10^{-6}$ m.rad und $B_U = 4.7 \cdot 10^4$ A.m-2 rad⁻².V⁻¹. Phase contours in cross-sections $z= 2.0$ and $z= 2.9$ cm of the beam formed in the same EOS but at modulating electrode potential $U_m = 0$ V are shown

in Figure 17. In this case the normalized emittance of the beam is $\varepsilon_{n,0.9} = 2.7 \cdot 10^{-6} \text{ m} \cdot \text{rad}$ and the calculated normalized electron brightness is $B_u = 8.36 \cdot 10^4 \text{ A} \cdot \text{m}^{-2} \cdot \text{rad}^{-2} \cdot \text{V}^{-1}$.

Another numerical experiments were performed to analyse an axially symmetrical electron gun for welding with a bolt-type tungsten cathode heated by means of bombardment with electrons emitted from a coil filament. Such cathodes yield higher currents and possess superior electron optical characteristics compared with directly heated filament cathodes. Additionally, the bolt-type cathode stands up well to ion bombardment and has a longer service life. Because of this, bolt-type cathodes are especially good for welding guns working in poor vacuum conditions.

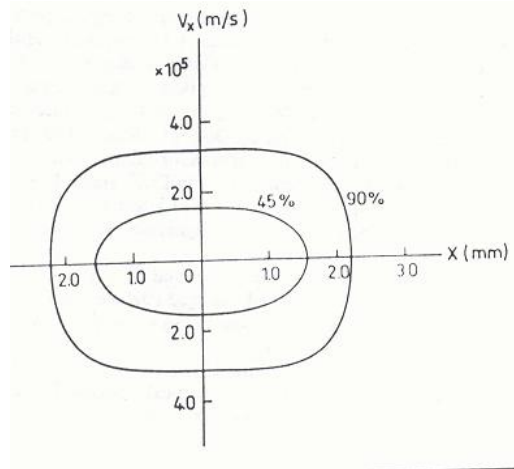


Figure 64. Emittance diagram of the beam in the cathode plane. The phase contours encompass 45% and 90% of the beam current respectively

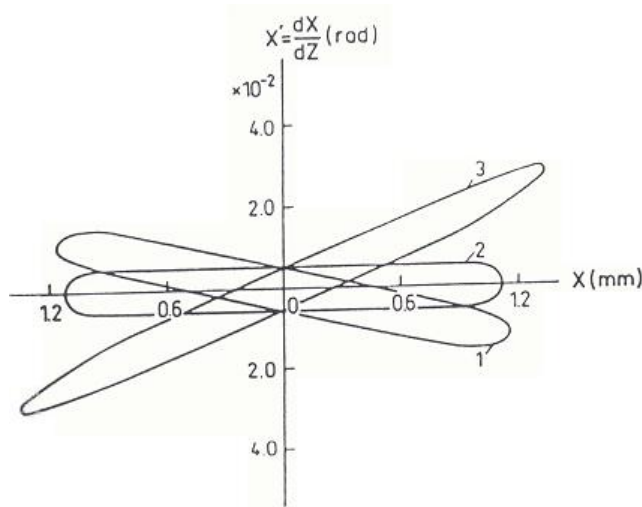


Figure 65. Transformation of the outward phase contour (90%) along the beam axis in three transverse cross-sections: 1) $z=2.22\text{cm}$; 2) $z=2.84\text{cm}$; 3) $z=3.9\text{cm}$

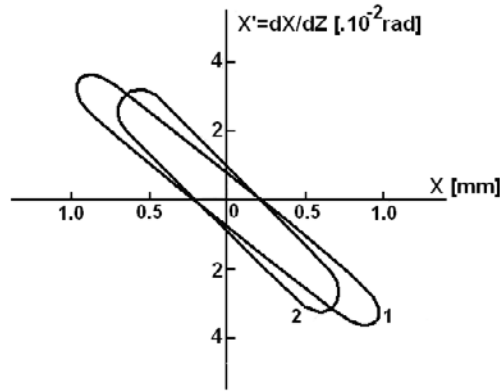


Figure 66. Outward phase contours of beam in the cross-sections:
1) $z=2.0\text{cm}$; 2) $z=2.9\text{cm}$; for $U_m=0\text{V}$; $U_a=30\text{ kV}$.

Here we present only the results from the analysis of the final (optimized) version of the gun, obtained for an accelerating voltage of 25 kV. The geometrical configuration of the gun is shown in Figure 67(a) together with the trajectories of the beam formed at a Wehnelt electrode potential of $U_w = -200\text{ V}$. It should be noted that the electron trajectories are computed after the self-consistent solution for beam space charge has been reached; that is, this computation is not performed on each iteration as in programs for which electron trajectories are used for the generation of a space charge map. The reason for the inclusion of trajectory output as one of the options of our program GUN-EBT is twofold. First, by doing this we pay tribute to the tradition and demonstrate that, although based on a novel phase approach, GUN-EBT is able to provide all the information available in numerical experiments carried out with packages implementing ray-tracing (trajectory analysis). Second, we would like to illustrate here the main difficulties in representing the beam in the configuration space as a set of trajectories. As it were mention yet, from Figure 67(a) one can see, that due both to the great number of overlapping trajectories and to the large difference between the longitudinal and transverse dimensions of the beam, the internal structure of the beam is effectively lost. By considering such plots one can gain only a general qualitative idea of the beam configuration. An attempt to show the internal structure of the flow produced at a Wehnelt potential of -400 V is presented in Figure 67(b). In this plot, the number of trajectories is reduced considerably and different radial and longitudinal scales are used. However, even in this representation, one of the inherent drawbacks of the trajectory plots still remains. It stems from the fact that different trajectories “carry” different space charges and therefore make different contributions to the beam formation; however, this is not directly visible from the plot. This being said, we proceed with the main portion of our analysis in an attempt to demonstrate that numerical experiments performed with GUN-EBT provide adequate information for the assessment both of the beam quality and of the electron-optical performance of the gun without considering explicitly the electron trajectories.

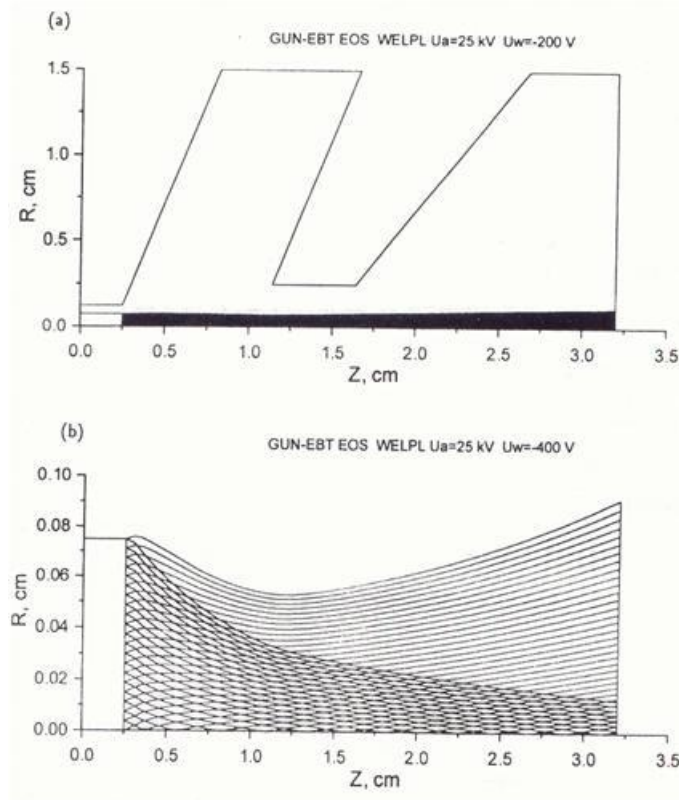


Figure 67. The geometry of analyzed electron gun and trajectories of electrons at accelerating voltage 25kV. a) $U_w=-200\text{V}$ (heating filament is not shown here) and b) $U_w=-400\text{V}$ (the electrode structure is not shown)

Intensity modulation of the beam is one of the most important processes in electron guns for electron beam welding. By varying the beam current density, one can control the beam power (and eventually the electron beam active zone) over a wide range from zero to maximum. The current density of the beam is controlled by the electric field in the near-cathode area in front of the emitter. For this purpose a Wehnelt electrode at a negative bias with respect to the cathode is used. Variations in the field shape and strength markedly affect the current extracted from the cathode. Different regions of the cathode can function under one of the following possible operating conditions: (i) the initial current regime, (ii) space-charge-limited flow and (iii) saturation (temperature-limited range of operation). It is well known that various grains of a tungsten crystal have slightly different work function. Since tungsten cathodes have a poly-crystalline structure composed of randomly oriented crystals the work function will vary in consequence in a random fashion across the emitting surface also. Another reason for variation of the saturated current density is the irregular heating of different cathode regions. As a result, the current density of a thermionic cathode working under saturation can be highly non-uniform and unstable. To avoid these problems, thermionic cathodes are usually operated in the space-charge-limited mode. This requirement is fulfilled in the analysed electron gun. As can be seen in figure 68, in front of the entire cathode surface there is a potential minimum which reflects a fraction of the electrons back to

the emitter. The location of the potential minimum for different potentials of the Wehnelt electrode is shown in Figure 69.

Accordingly, the retarding held region (from cathode surface to $z(U_{min})$) contains not only electrons traveling towards the anode, but also electrons falling back to the cathode. With a constant heating, a dynamic equilibrium sets in, so that the number of electrons reaching the anode of the elementary diode and falling back to the cathode is equal to the number of electrons emitted by the cathode.

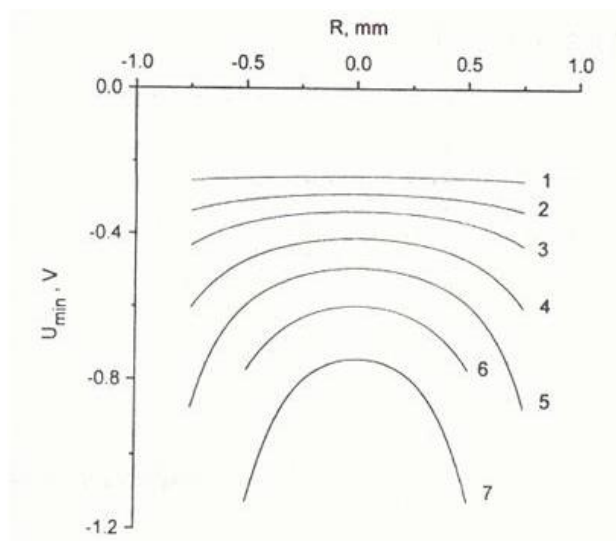


Figure 68. The potential minimum in front of the cathode. U_w : 1)-200V; 2)-400V; 3)-575V; 4)-800; 5)-1000V; 6)-1200V; and 7)-1400V

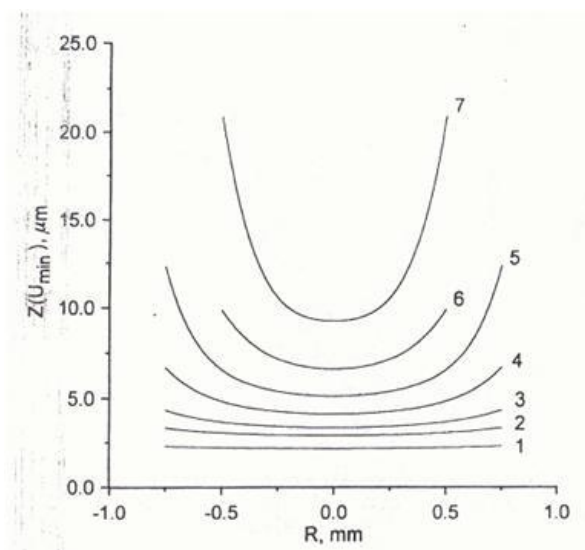


Figure 69. The distance: potential minimum-cathode. U_w : 1)-200V; 2)-400V; 3)-575V; ; 4)-800; 5)-1000V; 6)-1200V; and 7)- 1400V

Therefore, the anode current is smaller than the emission current. In this mode of operation the space charge in front of the cathode acts as a reservoir, or a source, which reduces the variations in current resulting from emission fluctuations. As is seen from Figure 68 and Figure 69, the Wehnelt electrode potential affects both the depth and the position of the potential minimum. When the Wehnelt electrode becomes negative, the potential barrier near the cathode increases, thus decreasing the extracted current. The cathode current undergoes an additional change due to the variations in the emitting area of the cathode. This is illustrated in Figure 70, in which current density distributions computed for different potentials of the modulating electrode are presented. It can be seen that, for potentials $U_w < -1.2$ kV the peripheral area of the cathode is facing a deep potential minimum and the cathode current is extracted only from the central regions of the emitter.

Measured and computed modulation characteristic of the gun are shown in Figure 71. It can be seen that measured and calculated values are in good agreement.

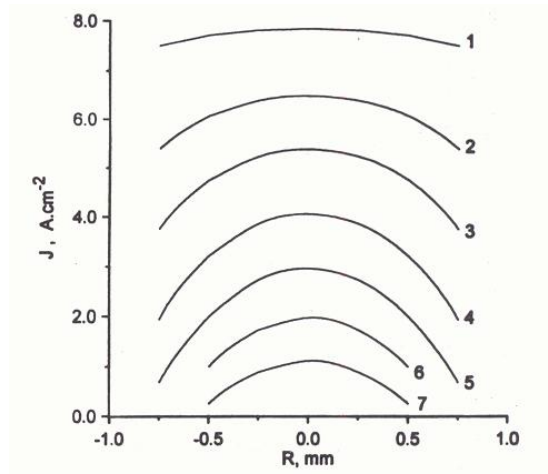


Figure 70. The current density distribution on the emitter plane. U_w : 1)-200V; 2)-400V; 3)-575V; 4)—800V; 5)-1000V; 6)-1200V; and 7)-1400V

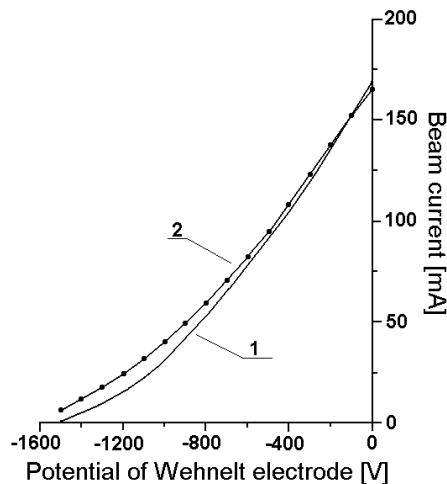


Figure 71. Modulation characteristics of the gun: 1) computed values; 2) measured experimental curve

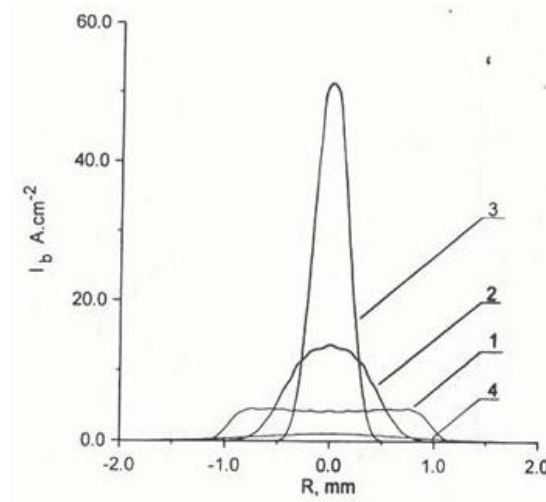


Figure 72. The current density distributions at the gun exit. U_w : 1)-200V; 2)-400V; 3)-575V; 4)-1000V

The Wehnelt electrode not only controls the current of the beam but also is an element of the immersion lens (made up of the cathode, Wehnelt electrode and anode) which focuses the electron beam. Generally speaking, in an arbitrary triode gun not all of the emitted electrons which have overcome the potential barrier near the cathode reach the target. Some electrons fall onto the anode. Owing to this, the beam current past the cathode may be a fraction of the cathode current. The electrons which fail to pass through the opening in the anode are lost from the beam and may even destroy the anode by overheating. Although the anode is water-cooled, such losses are undesirable. That is why the minimum-loss criterion is among the decisive requirements which one has to satisfy by choosing an appropriate geometrical configuration of the gun. The results of computer simulation predict loss free transport of the beams through the anode orifice in the analyzed electron gun. This conclusion is corroborated by measurements which demonstrated that the beam current practically equals the cathode current of the gun. The profiles of the current density distribution at the exit plane $z = 3.2$ cm for different potentials of the Wehnelt electrode are shown in Figure 72.

The quality of generated beams can be evaluated using the phase space analysis, based on the emittance concept. In an arbitrary transverse cross section $z = z_i$ of the beam each trajectory is characterized by its radial position r and slope $r' = dr/dz$ relative to the optical axis. Accordingly, the trajectory can be represented by a single point in a two-dimensional phase space (trace plane) with coordinates r and r' . The representative points of individual trajectories form a phase space portrait of the beam. Because the trajectories in phase space cannot intersect, a certain number of representing points lying on a contour that envelops a given region remain on the same contour regardless of the possible changes in its configuration.

One of the major advantages of such a description is related to the fact that it is much more convenient to trace the motion of a limited region of the phase space rather than to follow the individual particle trajectories. Knowing the behavior of the boundary enables one to draw a conclusion concerning the intervals within which the positions and moments (or the slopes) of all particles undergo changes. As a result, the general behavior of the electron beam can be considered instead of individual trajectories. Therefore, the concept of phase space

analysis provides a properly macroscopic description of a beam. It gives us a far deeper insight into the behavior of the gun than does the commonly used ray tracing.

Outward phase contours (emittance diagrams) encompassing the area occupied by the beam in the phase plane corresponding to the exit section of the analyzed electron gun for some values of the Wehnelt potential are shown in figure 24. The projections of phase contours on the Or and Or' axes indicate the maximum radial dimension and divergence angle of the beam. In our numerical experiments carried out with GUN-EBT the emittance diagrams can be obtained in different cross sections (including cross over) along the beam axis. During the acceleration the axial momentum of electrons increases, leading to a reduction in the beam emittance. In order to remove the effects of acceleration, the normalized emittance is used, (by multiplication on the ratio of the axial velocity to the speed of light). It is a useful invariant, which can be used to compare the quality of beams formed at different accelerating voltages. The emittance is measure of the beam non-laminarity and characterizes the disorder and the irreversible changes occurring in the beam. One common goal for the optimization of an electron gun design is the reduction of the beam emittance while producing a given amount of beam current. The production of low-emittance high-brightness beams is limited by several factors including electron momentum spread, space charge effects and aberrations.

In Figure 74 (curve 1), the normalized emittance as a function of the Wehnelt electrode potential is shown. It can be seen that ε_n decreases monotonically with the increase in negative modulation potential. This reduction is a result of the decrease both in the radius of the emitting area and in the maximum emission angle of electrons on the cathode $r'_{\max}$. The

latter corresponds to electrons emitted with maximum total energy (in the present model 1 eV, because the probability of emitting more energetic thermo-electrons is low and their contribution to the beam can be neglected) distributed so as to have maximum transverse and minimum axial velocity, namely

$$r'_{\max} = \arctan\left(\frac{V_{x0,\max}}{V_{z0,\min}}\right). \quad (132)$$

where for each elementary cathode region $V_{z0,\min} = [(2e/m)U_{\min}]^{1/2}$.

Although the normalized emittance is conserved along the beam axis, the aberrations can distort and wrap the shape of the phase contour, enlarging the effective area occupied by the beam in the phase plane. This situation is illustrated in Figure 73. It can be seen that, at $U_w = 0$ V, the emittance diagram is aberrated and surrounds regions of unoccupied phase space. In this case, the effective area of beam is larger than the actual area filled by the beam particles. This effective area divided by π gives the so-called effective emittance ε_{eff} . A commonly used method for the evaluation of the effective area and eventually ε_{eff} is to fit the two-dimensional phase space distribution with the minimum area ellipse that just encloses all particles. When the distribution is distorted, the ellipse must enclose a larger area containing empty regions of phase plane. An alternative approach for estimation of the effective emittance is to use the RMS emittance:

$$\varepsilon_{rms} = 4(\langle r^2 \rangle \langle r'^2 \rangle - \langle rr' \rangle^2)^{1/2} \quad (133)$$

where values in bracket $\langle \rangle$ are the mean squared values of r of all the trajectories. The RMS emittance is a figure of merit for the beam quality and provides a useful quantitative measure of the effective emittance. The smaller is the ε_{rms} , the better is the quality of the beam. This statement reflects the fact that a beam of smaller ε_{rms} (other parameters being equal) can be transported easily and can be focused to a smaller size on a target. It should be noted that, in the GUN-EBT code, the RMS emittance is calculated according to equation (41) using the phase space coordinates of all 'equivalent trajectories' whose contribution to the space charge map is taken into account.

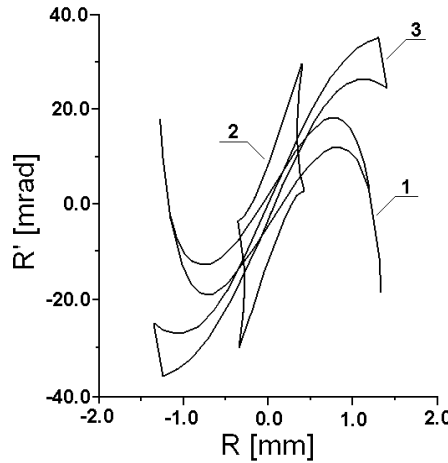


Figure 73. Phase contours of the beam at the gun exit.
Uw: 1)-100V; 2)-575V; 3)-120V

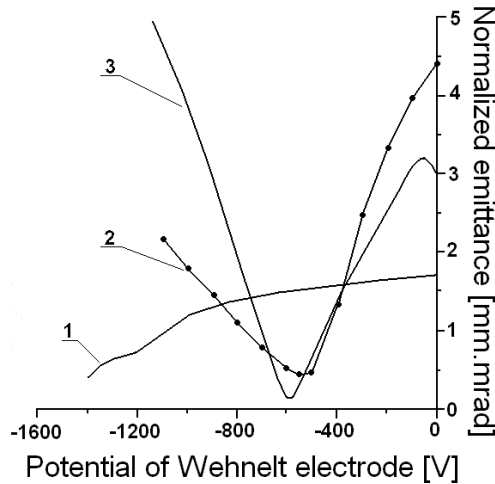


Figure 74. The emittance vs. Wehnelt voltage.
1) the normalized emittance (calculated),
2) the normalized RMS emittance, the measured normalized emittance

Figure 74 (curve 3) shows the dependence of the RMS normalized emittance on the Wehnelt bias. Initially, with increasing negative potential the RMS emittance decreases until a minimum is reached at $U_w = -575$ V. This (right) branch of the curve is a result of the formation of a more and more narrow and thus less and less aberrated beam.

Further increase in the Wehnelt potential, however, leads again to the formation of beam having greater divergence and consequently subject to greater aberrations.

The emittance was measured with a computer-controlled hole-slit analyzer. This method was chosen for two reasons. First, it is very suitable for studies concerning the aberrations of rotationally symmetrical charged particle beams, Second, the practical realization is easier compared with more sophisticated methods. In our experiments, a pin hole of 0.32 mm diameter was used to select the sample beam let by scanning across the beam along its radius. The angular distribution of the beamlet was analyzed with a slit of 0.14 mm width 'infinite' in direction perpendicular to the movement direction. The distance between the hole plane and the slit plane was 50 mm. Electrons of the beam passing through the system were collected by a Faraday cup. The distribution of the signal corresponds to the density of points obtained by taking a cross section, defined by the plane X-Z through the domain occupied in four-dimensional trace space, followed by a projection in the Y direction. On this basis, the iso-density contours known as 'section-projection' emittance diagrams were obtained for different potentials of the Wehnelt electrode. Measured values of section-projection normalized emittance $\epsilon_{sp,n}$ versus Wehnelt bias are shown in Figure 74, curve 2. It can be seen that there is qualitative correspondence between the computed rms emittance and the measured section-project! on emittance. It must, however, be emphasized that it would be inappropriate to seek more than qualitative agreement between ϵ_{rms} and ϵ_{sp} because these two values are different by definition.

On the basis of the analysis made, one can conclude that the electron-optical properties of the analyzed electron gun meet the requirements and are quite appropriate to the intended application. In order to test the performance of the gun a set of technological experiments was carried out. As an illustration, in Figure 75 are shown profiles of welds produced in stainless steel (type 304) at accelerating voltage $U_a = 25$ kV, welding speed $V_w = 5$ mm.s⁻¹ and different beam currents I_b . In each run the focusing of the beam was so adjusted, as to obtain the maximum attainable depth of the weld. The results are quite typical for the corresponding levels of beam power and they give a lot of confidence in the technological capacity of the gun.

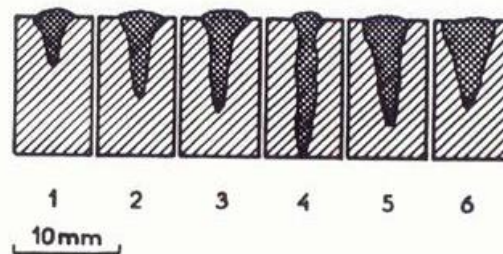


Figure 75. Profiles of welds ($U_a=25$ kV, $V_w=5$ mm/s) The beam current: 1) $I_b=60$ mA – $h=4.8$ mm; 2) $I_b=80$ mA – $h=7.2$ mm; 3) $I_b=95$ mA – $h=9.0$ mm; 4) $I_b=125$ mA – $h=13$ mm; 5) $I_b=145$ mA – $h=9.9$ mm; 6) $I_b=170$ mA – $h=8.4$ mm;

CONCLUSION

In this Chapter micro-characterization and integral macro-characterization of EB are described shortly. The monitoring of the beam current distribution across the EB is described. Modified pinhole approach and use of rotating wire, changing position of measuring device around the beam as well as the modified Faraday cup in combination with computer tomography reconstruction technique are most prospective methods for evaluation of beam profiles.

Measuring of the angular distribution of beam electron trajectories by application of a direct application of pinhole approach requiring about 10^6 sampling measurements, that is not practical for the practices of the welding workshops. Realistic way to map the brightness or emittance of the intensive EB is use of transverse beam profiles (2-3 measured transverse current distributions by two orthogonal slits in entrance refractory plate of a Faraday cup and assuming Gaussian distribution, or by 4-7 such distributions, measured by radial slits in entrance disk of the modified Faraday cup and a computer tomography code with minimal entropy and estimating the emittance and relative brightness per one volt as quality invariance of the technology intense electron beam

These data could be used at standardization of EBW machines, at transfer the concrete EBW technology from one machine to another and by periodic tests during welding of serial joints aiming an achieving the improved reproducibility and quality of EB welds at responsible applications.

There are discussed the requirements, the physical problems and many details of design of the high brightness electron guns for EBW. Accumulated knowledge during long term studies, design and use of powerful electron guns could be of use for many researchers with activity in physical problems and successful application of EBW.

Important place in discussion of EBW guns take the computer simulation and characterization of the produced intense beams. The authors discuss and apply for investigation of the generated intense beams the wide spread trajectory analysis. A important new approach – the phase analysis is apply for modeling the generation, control and directed transportation of suitable for EBW electron beams. More wide distribution of this new method for simulation of generated beams in EB guns for welding could be base of design of a new generation of perfect technology guns with high brightness and low emittance.

REFERENCES

- [1] International standard ISO14744 *Welding-Acceptance inspection of electron beam welding machines*-parts -2 and 3, 2000, (E)
- [2] International standard ISO/TR 11146 *Lasers and laser-related equipment - Test methods for laser beam widths, divergence angles and beam propagation ratios*-part, 3, 2003, (E)
- [3] Orlinov, V; Mladenov, G, *Electron and ion methods and equipments for treatment and analysis of materials*, (in Bulgarian), Techniques Publ.House, Sofia, 1982, 308.
- [4] Gabovich, MD; Kovalenko, VP; Metallov, OA. *J. Tech.Phys.*, (Russia), 1977, Vol.47, 1569-1571.
- [5] Boersh, H; *Z. Phys.*, 1954, Vol. 139, 115.
- [6] Sabchevsky, S; Mladenov, G; *J. Phys.D Appl.Phys.*, 1994, Vol. 27, 690-697.

- [7] Reiser, M. *Theory and Design of Charged Particle Beams*, Willey-VCH, 2008.
- [8] Lejeune, C; Aubert, J. Emittance and Brightness: Definitions and Measurements. In: *Applied Charged Particle Optics, Part A*, A. Septier, ed; Academic Press, New York, 1980, 159.
- [9] Lawson, JD. *The Physics of charged particles beams*, Oxford, Clarrendon, 1977.
- [10] Mladenov, G. *J. Scientific devices*. 1979, Moscow, (In Russian), 14-16.
- [11] Elmer, JW., et al. *US Patent*, 5 382895 , issued Jan.17, 1995.
- [12] Elmer, JW., et al. *US Patent*, 5 468 996, issued Nov.21, 1995.
- [13] Giedt, H., et al. *US Patent* , 5 483036 issued Jan.1996.
- [14] Wojcicki, S; Mladenov, G. *Vacuum*, 2000, Vol. 58, 523.
- [15] Dilhey, U; Masny, H. *Electronika & Electrotechnika*, Vol. 41, No 5-6, 2006, 61-65 (Publisher CEEC, Sofia, Bulgaria)
- [16] Elmer, JW. et al. *US Patent*, 6 300755 issued Oct. 9, 2001.
- [17] Elmer, JW. et al. *Science and Technology of Welding and Joining*, 1998, Vol.3, No 2, 51.
- [18] Mladenov, G; Koleva, E. *Vacuum*, 2005, Vol. 77, No4, 457.
- [19] Dilhey, U; Boehm, S; Dobner, M; Trager, G. In: *Proc. of 5-th Intern. Confer. on EBT*, Varna Bulgaria, 2-5 June 1997, 76-83.
- [20] Rykalin, N; Uglov, A; Zuev, I; Kokora, A. *Laser and Electron Beam Material Processing*, MIR Publishers, Moskow, 1988, 77.
- [21] Koleva, E; Vutova, K; Wojcicki, S; Mladenov, G. *Vacuum*, 2001, Vol. 62, 105.
- [22] Koleva, EG; Mladenov, GM. *Russian Physics Journal*, 2006, Vol. 11, 49-53.
- [23] Koleva, E; Menhard, Ch; Loewer, T; Mladenov, G. *Electronika & Electrotechnika*, 2006, Vol. 41, No 5-6, 51-60. (Publisher CEEC, Sofia, Bulgaria)
- [24] Koleva, E; Mladenov, G. *IEEE CPMT, Annual School Lectures*, 2005, Vol. 25, No1, 3-6.
- [25] Menhard Ch.G. *Proc. 8-th Intern. Conf. EBT*, 5-10 June, 2006, Varna, Bulgaria, 2006, Vol. 2, 11, Publisher IE BAS. Sofia.
- [26] Kasper, E. *Optik*, 1985, Vol.71, 129.
- [27] Kasper, E. *Nucl. Instr. and Methods A*, 1987, Vol. 268, 446.
- [28] Kumar, L; Kasper, E. *Optik*, 1985, Vol. 72, 23.
- [29] Weber, C. *Pilips Res.*, Reports, Suppl. 6, 1964, 1.
- [30] Ninomiya, K; Urano, T; Okoshi, T. *Trans. Inst. Electron. Commun. Eng.*, Jpn. 1971, Vol. 54B, 490.
- [31] Monro, E. *Nucl. Instr. and Methods A*, 1987, Vol. 258, 443.
- [32] van den Broek, MHL. *J. Appl. Phys.*, 1986, Vol. 60, 3835.
- [33] Mladenov, GM; Sabchevski, SP; Popowa, GS. *J. Tech. Phys.*, (Russia), 1986, Vol 56, No 4, 652-659.
- [34] Becker, R. *Electronika & Electrotechnika*, 2006, Vol41, No5-6, . 15-19. (Publisher CEEC, Sofia, Bulgaria)
- [35] Thomae, H; Becker, R. *Nucl. Instrum. Methods A*, 1990, Vol. 298, 407.
- [36] Sabchevsky, S; Mladenov, G. *Optik*, 1992, Vol. 90, 117.
- [37] Sabchevsky, S; Mladenov, G. *J. Phys. D: Appl. Phys.*, 1996, Vol. 29, 1446.
- [38] Ivanov, A; Titov, A. *Izvestia LETI*, 1975, Vol. 181, 60. (In Russian) Publ. St.Petersburg Electr.University, St.Petersburg, Russia.
- [39] Pelletier, J; Pomot, C., *Appl. Phys. Lett.*, 1979, 34, 249.
- [40] Yu, N; Tang, Ch; Zeng, Ch; Li, Q; Gong, K. *Proceed. of 2005 Particle Accelerator Conference*, 2005, Knoxville, Tennessee, USA, 4323-4325 (Publ.IEEE)
- [41] Jansky, P; Zlamal, J; Lencova, B; Zobac, M; Vlcek, I; Rdlicka, T. *Vacuum*, 2009, Vol. 84, No2, 357-362.

Chapter 2

PROCESS PARAMETER OPTIMIZATION AND QUALITY IMPROVEMENT AT ELECTRON BEAM WELDING

Elena Koleva and Georgi Mladenov*

Institute of Electronics at Bulgarian Academy of Sciences, Sofia, Bulgaria

ABSTRACT

The complexity of the processes occurring during electron beam welding (EBW) at intensive electron beam interaction with the material in the welding pool and the vaporized treated material hinders the development of physical or heat model for enough accurate prediction of the geometry of the weld cross-section and adequate electron beam welding process parameter selection. Concrete reason for the lack of adequate prognostication is the casual choice of the heat source intensity distribution, not taking into account the focus position toward the sample surface and the space and angle distribution of the electron beam power density. This approach, despite extending the application of solution of the heat transfer balance equations with the data of considerable number of experiments, results in prognostication of the weld depth and width only in order of magnitude. Such models are not suitable for the contemporary computer expert system, directed toward the aid for welding installation operator at the process parameter choice and are even less acceptable for automation EBW process control.

Various approaches for estimation of adequate models for the relation between the electron beam weld characteristics and the process parameters, the utilization of these models for process parameter choice and optimization are considered.

A statistical approach, based on experimental investigations, can be used for model estimation describing the dependence of the welding quality characteristics (weld depth, width, thermal efficiency) on the EBW process parameters - beam power, welding speed, the value of distance between the electron gun and both the focusing plane of the beam and the sample surface as parameters. Another approach is to estimate neural network-based models. The neural networks were trained using a set of experimental data for the

* Corresponding author: e-mail: eligeorg@abv.bg

prediction of the geometry characteristics of the welds and the thermal efficiency and the obtained models are validated.

In the EBW applications an important task is to obtain a definite geometry of the seam as well as to find the regimes where the results will repeat with less deviations from the desired values. In order to improve the quality of the process in production conditions an original model-based approach is developed.

Process parameter optimization according the requirements toward the weld characteristics is considered. For the quality improvement in production conditions, optimization includes finding regimes at which the corresponding weld characteristics are less sensitive (robust) to variations in the process parameters.

The described approaches represent the functional elements of the developed expert system.

INTRODUCTION

The use of electron beam welding (EBW) for joining applications has more than 45 years history. The technology recommended itself as reliable and universal tool, which is able to solve wide range of problems. EBW occurs to be the only solution for problems such as joining of reactive at high temperature metals or of heavy constructions. The main advantages of this technique are the deep and narrow welds and small thermal affected zone, as well as the high joining rate.

Power Beam Technology, often known as Concentrated Energy Flux (CEF) Technology, belongs to a class of novel manufacturing techniques. The primary attribute, which distinguishes the beams from conventional sources, is the power density, normally expressed as GW/m^2 . The power density characterizes the interaction of beams with materials and the relative importance of various thermal processes, as shown in Table 1. The highest power densities are available with electron beams and laser beams as they can be tightly focused. On the other hand, if one considers source strengths, the currently available plasma sources are at MW level, whereas the electron beams are at hundreds of kW/MW and lasers beams at a few kW.

With the advent of the beam technologies, it has become possible to localize heat transfer processes both spatially and temporally. The use of power beams in welding, melting, deposition of thin films, local evaporation of material for machining of holes or channels in irradiated sample, as for surface thermal modification is known for more than five decades. It should be noted that the total energy is equally important parameter in addition to beam power density in material processing. Electron beams up to 150 keV are employed for heat processing, whereas energy range of 100 keV to 10 MeV is most suited for radiation (non thermal) processing. Here some of the developmental efforts in the area of electron beams are presented along with a short discussion on the comparative performance of competing technologies.

Power Beams are characterized by high energy density at the impact point with excellent control of power and movement. The beams have to be distinguished in terms of their generation, transport and impact as illustrated in Table 2, where the corresponding auxiliary systems are also indicated. The critical parameters of power beams are beam size, divergence, location of beam waist, source stability and reproducibility. The beam diameter and the position of the minimum cross section of the beam relative to the work piece strongly

influence the beam processing of materials. Experimental techniques to evaluate them are based on the width of the melted zone or cut area, electrical/optical and thermal measurements.

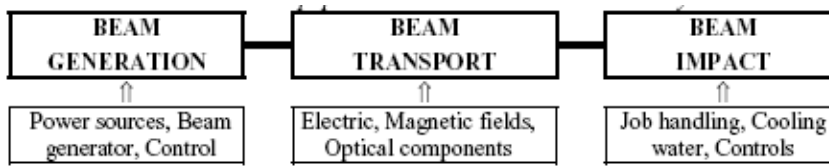
Beam generation is carried out in an electron gun, a laser cavity or in a plasma torch. The transfer of the beams from the generator to the work piece is achieved by the beam transport system. With electron beams and laser beams, the travel time is almost instantaneous and nearly at the speed of light, whereas plasma velocities are much slower. Unlike plasma beams, laser and electron beams are normally transported as large diameter beams and subsequently focused to a fine point on target. The electron beam needs vacuum for beam transport. The electron beams, being charged, can be focused by electromagnetic fields directly. A photon or a charged particle or a neutral atom impinging on a surface may be reflected, absorbed or re-emitted.

Electron beams (EB) due to space charge have limited power density in the focus spot of order of 10^8 - 10^{13} W/m². Laser beams can be focused up to higher power densities. In the same time due to higher efficiency of transformation of electrical power in energy of EB (near to 99%) if one do comparison with the laser beam (where the efficiency of this transformation is only few %) EB have no competitor in the area of powers higher than tens of kW. Due to possibility of transportation of the laser beam in air or gases with pressure of the order of 1atm. lasers are used for cutting by local melting and sequential melt flashing (often one say ablation), due to reactive force of evaporating material, or due to laser ablation mechanisms. Plasma cutting use also local melting, but melt is transported by gas flow (in the some laser cutting regimes that mechanism also takes place).

Despite of the wide use of EBW and of similarity to the laser welding, the knowledge of the physical processes governing EBW is still incomplete. The weld geometry characteristics and the weld defects depend on a large number of parameters, describing the material and the EBW device properties as well as itself technology process. The complicated interactions between the energy flows and treated material as the unknown fully drilling mechanism of the beam and complicated dynamics of the molten weld pool lead to uncompleted physical equations model controlling the beam penetration and the heat transfer. Instead of an exact description only rough approximations of many parameters are generally used. The real power distribution over the spot, where the beam hits the sample, is a complex function of co-ordinates and time. This is due to generation of a crater of variable shape in the molten metal, through which the electron beam with changing during interactions angular and radial energy distributions penetrates the treated sample. Phase transformations, mass transfer, as a change of material characteristics with the temperature take place too.

Table 1. Beam power density and related thermal processes

Beam Power Density in W/m ²										
10 ³			10 ⁶			10 ⁹			10 ¹²	10 ¹⁵
			Arc Processes		P l a s m a s			CO ₂ Lasers E B e a m s		
Negligible melting		Radial conduction dominated with melting			Vaporization, conduction and melting				Vaporization dominated	

Table 2. Power beam equipment - Common features & auxiliary facilities

To simplify the problem a quasi-steady-state model (involving a linear and uniformly distributed movable heat source (see Figure 3), sometimes modified as combination of a linear and an added point sources) is created. Using this heat model solely [1] as well as in a combination with experimental data [2] a prognosis of the windows of the possible weld geometry parameters was done. The model was used successfully for evaluation of the geometry of deep penetrating EB welds [3], welds at EBW of thin plates [4,5], as well as in the EB surface modification [6].

In the paper (follow closely [2]) using thermal model of EBW the expected ranges of the observable weld depths versus the EB power P , especially on the parameter P/H (where H is the weld depth) are given for the deep penetrating beams of power range (1 - 40) kW. The expected weld width range vs. welding speed V has been prognosticated too.

But the rough choice of an arbitrary and continuously steady beam power distribution in this model is reason for the loss of the influence of the focus position (relative to the sample surface) as well as the influence of the beam oscillation on the process results. An other type variations (uncontrolled by operator) of the beam energy distribution could be caused by adjustments of the gun as well as by the changed states of the electron optical system electrodes during the gun working time, or during the different runs of the same welding machine. Differences of the electron guns design of various machines are also neglected in the predictions of welding results. All approximations eliminate from such evaluations the behavior of the beam radial and angular energy distributions that are strongly machine dependent. Consequently, at unknown energy distribution, as well in the some time also at approximated values of the process parameters an exact calculation of the weld characteristics is unexpected. From computer simulation and experiments [7-12] is known that repeatable obtaining of the best welding results and transfer of the regime parameters for EBW of concrete details from one to other machine are possible only under knowledge of practical useable parameters (for example emittance in different meanings) and suitable measuring systems for the beam characterization.

The mentioned complications in the physical models and determination of the beam characteristics as in the control of the beam heat transfer require creation of an adequate statistical model of EBW. That model must be able to help the achievement to reliable choice and control of the process parameters, as well as to the estimation of the expected geometry characteristics of the EB welds at given regime and treated sample material. Our attempts to develop such approach are reviewed in the presented paper. In reference [13] EBW is studied as multi-response experiment, implemented at a set of working conditions. In this technique the surfaces for the depth and width for the factors - beam power, welding speed and the position of beam focus toward the sample surface, beam oscillation parameters had drawn. A suggestion for a sequential procedure of optimization as part of further improvement to the model had been also described [13]. Some useful data for values of heat efficiency of process

and further understanding of relationships between welding seam parameters and welding regimes were given in reference [14, 15, 16]. Model based approach for quality improvement of electron beam welding applications in mass production is given in [17].

PLACE AND APPLICATIONS OF ELECTRON BEAM WELDING

The basic advantage of power beam welding is the small heat input, which means minimal and easily controlled bead-width, heat-affected-zone and weld-distortion. In addition, the range of combination of the joining materials is wide including those with high melting points and widely different physical properties. When selecting a process for a specific joining application, a number of questions such as joint preparation, cleaning, inert gas or vacuum shielding, depth of penetration, weld joint accessibility, productivity, and cost must be answered. Comparison of various aspects associated with electron, laser and plasma beams are listed in Table 3.

It is impossible to state with conviction, which welding process should be used for maximum efficiency in a given application. Both electron beam and laser are good choices for critical, heat sensitive weld joints and widely dissimilar materials. Electron beam is the indisputable candidate for penetration beyond 6 mm without preparation of the weld joint. For not very high volume welding (of order of thousand or tens thousands of small component assemblies, laser offers the best approach. It should be mentioned that the ability of the lasers to be transported to inaccessible areas using optical fibbers makes it particularly useful in hazardous work. For maximum flexibility, immediate use, lower critical joint tolerances, and low capital investment, plasma arc and gas tungsten arc are the dominant choices.

EB welding have benefits in mass production (hundred thousands pieces). Vacuum as shielding environment is 35 times cheaper (if not include capital costs) than pure gas shielding of molten pool. At welding of lightweight metals EB not need anti-reflex coatings. High voltage EB (of order of 150 kV) can be brought out the vacuum chamber in air environment, but radiation protection of the operator is need. An intermediate evacuating by differential pumps space and Helium flow are used at such EB welding at atmospheric pressure.

Table 3. Comparison of Welding Processes

Parameter	E-BEAM	LASER	PLASMA
Penetration Thickness[mm]	0.5-200	0.5-50	0.1-10
Welding Speed	Fast	Fast	Medium to Fast
Distortion	V. Low	Low	Moderate
Power Density [W/m^2]	10^9 - 10^{12}	10^{11} - 10^{13}	10^8 - 10^{10}
Maximum Power [kW]	100	10	15
Equipment Size	V. Large	Small	Medium
Cost Comparison	5 -10	10	1
Operational Constraints	HV, X-Rays	Optical	Ultraviolet
Difficult Locations	V. Poor	V. Good	Fair

Figure 1 present a comparison between cross-sections of welds (at equal depth) obtained after 1) EB welding; 2) micro-plasma welding and 3) Ar arc welding. The heat input in the sample is proportional to area of the melt zone. So the distortions of welded sample and the need of position-fixing equipment for welded pieces is lowered or avoided.

In the last few decades, EB welding of the refractory metals and alloys, of heterogeneous metal junctions and of heavy engineering components were wide spread. The high joining rate, the deep and narrow weld (Figure 2 and Figure 3) and the minimal heat affected zone are basic advantages leading to the most often use of this process.

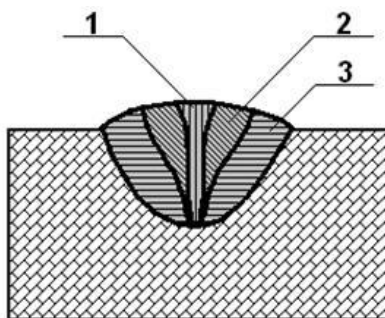


Figure 1. Cross-sections of various welds

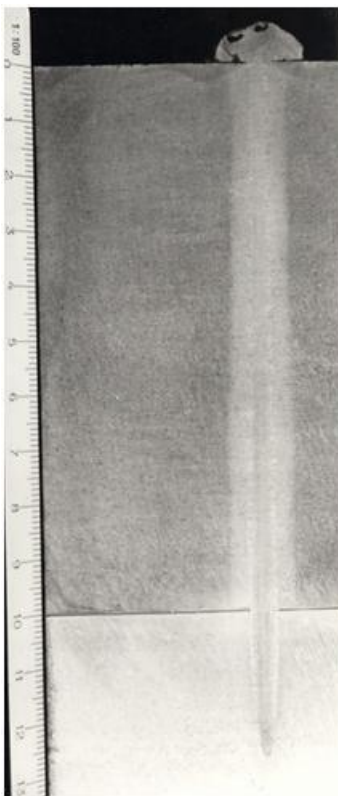


Figure 2. EB weld with deep penetrating beam with power density 1011 W/m^2 ; (165 kV, 320 mA , 3.5 mm/s)

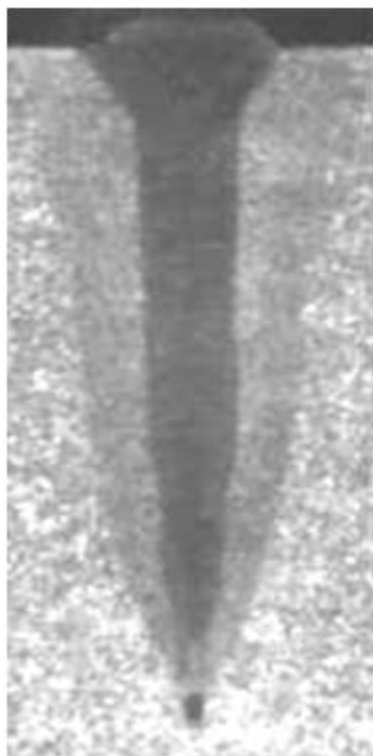


Figure 3. Metallographic photographs of the transverse cross-section of the EB welded junction of two plates with thickness of 78 mm. A deep and narrow molten zone and two heat affected zones are shown. The beam power is 15 kW, welding speed is 1 cm/s, the beam is focused 60 mm below the sample surface

The development of new high-intensity heat sources such as electron beams (EB) has facilitated welding of refractory metals and alloys, of heterogeneous metal junctions and of heavy engineering components. Electron beam welding (EBW) of materials has a number of decisive advantages over conventional techniques. The focused electron beam is one of the highest power density sources and that way high processing speed are possible, narrow welds with very narrow heat affected zone can be produced accurately. The weld cross-sections may have a "knife" shape. This is one the main advantages of the EBW method over the conventional methods of welding - the lower energy needed for the formation of a joint with equal width. The narrow heat affected zone allows the welding of materials and components near the weld zone that are not suitable for such processing. The crystal structure near the welded area is preserved unchanged, which on the other hand leads to preserving of the physical and mechanical properties of the welded materials. The thermal deformations are minimal, i.e. less are the cavities in the zone around the weld. The welded details may be thin or wide, and also can have different thermal conductivity. EBW is suitable for the welding of chemically active at high temperatures metals (Zr, Ta, Ti, Hf, Mo, W, Be, V etc.) and their alloys due to the fact, that the process is held in vacuum.

The Change of the weld and thermal affected zones at opening of the key-hole from the back side of work-piece is presented in Figure 4.

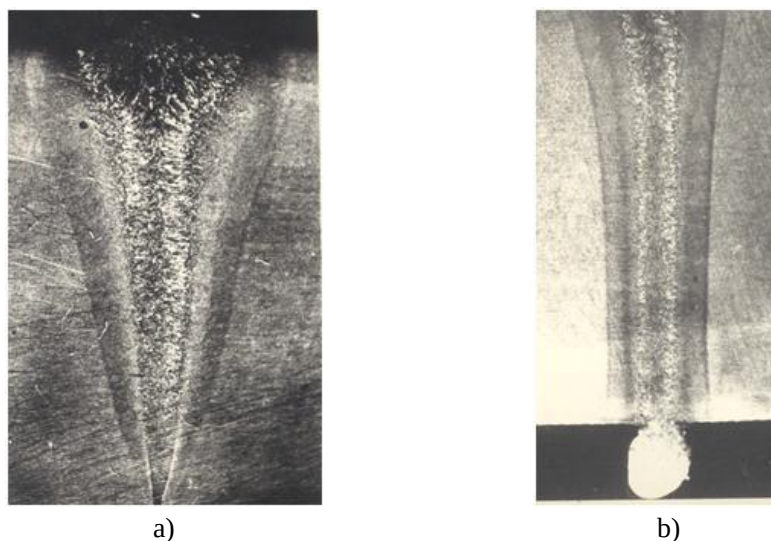


Figure 4. Change of the weld and thermal affected zones at opening of the key-hole from the back side of work-piece

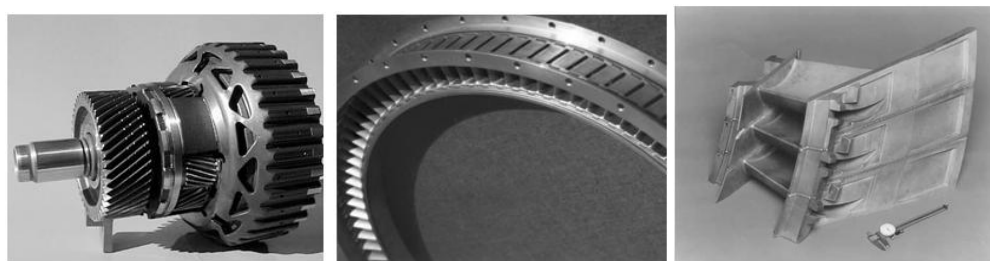


Figure 5. Some typical applications of EBW technology. a) Automatic CVT gear. Planetary and drive gear, welded with 4 EB-welds b) Aircraft - stator ring assembly with more than 300 EB-welds to join the vanes to the ring and the ring to flanges c) Industry - nozzle guide vanes for large turbines

Another characteristic of the welded seams is its hardness. In the area of the weld the hardness is usually higher than in the non-welded areas that can do it brittle. The crystal structure of not-melted metal is changed only in the narrow thermal affected zone.

Some applications of EBW technology are shown in Figure 5.

ELECTRON BEAM WELDING EQUIPMENT

Practically the electron beam welding is based on the use of the kinetic energy of a beam of accelerated electrons for a local heating of the welded material in the region of the joint up to temperatures higher than its melting temperature. The principal scheme of an electron beam welding installation is given on Figure 6, where there are: (1) – electron gun - the generation, acceleration and focusing of the electron beam are held there; (2) – vacuum chamber - the welded details and the electron gun are situated there; (3) – fixing system - for fixing and moving the details. There are scuttles for changing the samples and for the observation of the process – (4). The volume of the vacuum chamber depends on the welding

samples - from some dm^3 to hundreds of m^3 . The chamber walls provide the necessary mechanical hardness, vacuum density and the protection of the personnel from the x-ray radiation, appearing due to the interaction of the accelerated electrons with the welded material. The vacuum system (5) consists from diffusion pumps for high vacuum and mechanical pumps for low vacuum, as well as the needed vacuum faucets, pipelines and measuring devices. Except these installations designed for welding at high vacuum (10^{-2} – 10^{-3} Pa), there are others for medium vacuum welding (10^2 – 1 Pa) and welding at normal pressure. The high voltage generator – (6), includes a powerful source of voltage for accelerating the electrons, and a source for heating the cathode and control of the beam (when the triode electron gun is with thermal cathode). The installation includes low voltage sources for the electric supply and control of the focusing and averting system (7) of the electron gun and the manipulator – 8 (and the electron gun, if it is movable). The installation includes an optical or TV system (6) for observation of the process. To prevent the optical elements or the windows for the observation there are built-in appliances. The control deck (9) is used for the control of the process of welding and the supportive operations.

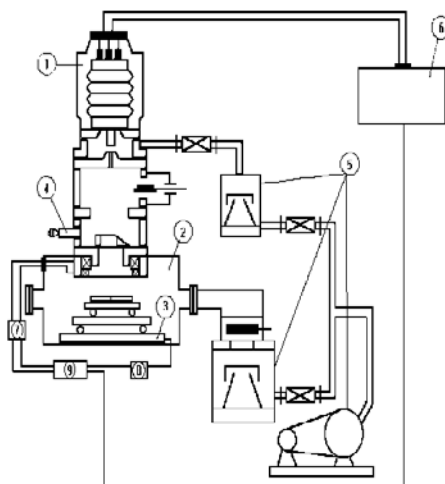


Figure 6. Block-scheme of EBW equipment



Figure 7. First EBW machine in Bulgaria-1974

The first EBW machines in the world were built by Stor (France) and Steigervald (Germany) – before 1956. In the next five years N. Olshanski (Russia), B. Paton and O. Nazarenko (Ukraine), W. Ditrich (Germany), design own EB welders. In Poland Dr. K. Friedel, J. Felba (Wroclaw) and W. Barwicz, S. Wojcicki (Warsaw) were the pioneers.

The first EBW machine built and operating in Bulgaria (and Institute of Electronics at Bulgarian Academy of Sciences – IE-BAS) was designed from a small team headed by Prof. G. Mladenov during 1973-1974 (Figure 7). The first developed in IE BAS EBW technology for the Bulgarian industry was EBW assembling of a sensor for the angular velocity. It can be seen in Figure 8. A method for calculation of regimes of EBW of thin films was created during experiments for mastering these devices. On Figure 9 an idea of an electron gun for welding is given. There the main parts are: (1) is isolator – protected against self-coating; oil & water cooling; (2) - triode electron beam generating and accelerating system – heated tungsten band cathode – beam power is up to 7.5 kW at maximum 60 kV accelerated voltage – the shape of focusing and anode electrodes is computer simulation optimized; (3) in anode water-cooled plate an adjustment of electrical and geometrical axes of the gun system is provided. (4) is partial vacuum pumping system – usually turbo-molecular pump - oil free evacuation of the residual gases in acceleration part of the gun provide longer cathode life; (5) isolation valve – an important element controlling often the output of the welding machine; (6) visual observation system give a possibility for beam's eye view of work-piece before, during and after processing; (7) and (8) electromagnetic focusing and deflection systems respectively.



Figure 8. Sensor assembled by EBW

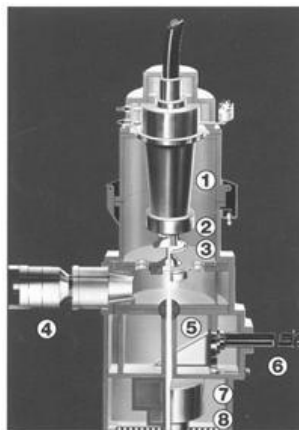


Figure 9. Electron gun

In 1981 in IE-BAS were produced electron beam installation "ELI"1300 ordered by Academy of Sciences of Belarus (Figure 10). During design of that plant were get (or requested by application) four patent- for universal x-y manipulator, for an active operating filter to decrease ripple component in D.C. output voltage [18] and circuits for lower intensity of discharges [19]; a new method for optimal focusing [20].

Figure 11 shows a Leybold-Heraeus EBW machine 7.5 kW, 60 kV. CN control unit is also seen there.



Figure 10. Front view of ELI 1300



Figure 11. EBW plant, produced in Germany

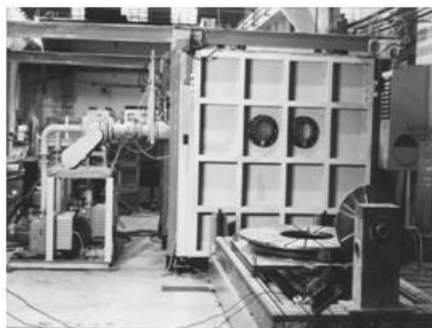


Figure 12. EBW plant for heavy industries

EB welding plant shown in Figure 12 is intended for joining of big machine parts (for heavy industry) weighting up to 5 tons; with a diameter up to 1500 mm, maximal length 2800 mm; the thickness of welded walls are (at use of 60 kW 120 kV gun) - up to 75 mm steel; 120 mm copper and up to 400 mm Al. Manipulators permit longitudinal straight line horizontal and vertical welds, rotational in horizontal or vertical planes welds. Added material is not provided. Welding speed is between 0.4 to 15 cm/s (0.24-9 m/min).

The vacuum chamber dimensions are 3×3×2.5 meters (volume is 22 m³). The pumping system provides achieving the working vacuum (10 Pa) for 16 min. Note that welding is realized under intermediate pressure of the residual atmosphere. Turbo-pump of the electron gun works directly in free volume of vacuum chamber - no special mechanical pumps for partial pumping of accelerating gun volume. For assembling of such plant are needed 180 m² area (the area for plant parts is 80 m³). The height of the working hall is 6m. The elevator must provide manipulation of 5 tons weight. Electrical power installed must be 160 kW; the needed water is 2.4 m³/h.

The work vacuum chamber has two sliding doors. For loading and unloading the work-piece table is moved out of the work chamber onto a run-out platform. The movable table (stage) accommodates a universal rotator with horizontal or vertical axis, and a back center. The precision of guidance of the electron gun is gained by special unloading drive and electron gun manipulator mechanics. That 3-axis manipulator have equal to high precision machine tools operation with tolerances in the hundredth-of-a-millimeter range.

The gun can be mounted in any spatial position; has an independent turbo-molecular system; has a possibility to be deflected by mechanical rotation. The cathode area of the gun is isolated by the common vacuum volume by vacuum valve to keep the hot parts of the gun in vacuum when the work chamber is vented.

TENSILE, HARDNESS AND MAGNETIC INVESTIGATIONS AT ELECTRON BEAM WELDING OF DISSIMILAR MATERIALS

Nevertheless that the copper could be used to braze steel, the joining of these dissimilar metals by fusion welding is difficult. The copper and steel are not very compatible components for mixing in a weld. Often explosive or friction welding was applied [21,22] for that joints, but use of these methods are highly dependent of configuration of the components. Electron beam welding (EBW) process has been found to be especially well suited in this area. In aerospace applications, nuclear and scientific devices design various joints of these metals, such as heat exchanger tubes [23], copper cavities and copper beam lines with Conflate stainless-steel flanges [24] are done by EBW. Selection of the appropriate welding conditions and parameters needs [25-27] thorough investigations. In this paper are given results of a study of welding conditions and obtained welds of these dissimilar metals.

The EBW is done using a conventional 60 kV electron beam welder. The vacuum chamber volume is about 300 l and the vacuum pressure during welding is 10⁻⁴ Torr. The gun cathode is from tungsten sharp with width 1 mm. The experiments are performed with plates, placed horizontally on the manipulator in the vacuum chamber of EBW machine. They are weld together using a vertical electron beam. Below the joint welded a thin copper plate is placed (with thickness of about 2 mm) in the case of 10 mm copper plate thickness or by

machining a 2 mm sub-plate under the joint was formed in the case of 12 mm copper plate. After the EBW the welded plates were cut to pieces with width 12 mm and then machined in order to make narrower central parallel length (with width 10 cm) of the specimen for the tensile test. The welding was performed for one pass without preheating. Welding results at beam position on the steel or on the copper predominately are investigated. Oscillations of the beam are not used during the experiments. The accelerating voltage is 60 kV and the distance from the electron gun to the sample surface is 36 cm. The variations in the experimental conditions are given in Table 4.

Three standard types of copper were used during the experiments (see Table 4). The data for the copper used and the chemical composition of specimens are given in Table 5. The chemical composition of the stainless steel (SST) according the Bulgarian Standard (BDS) is presented in Table 6. These types of SST correspond closely to SST used in other standards (German, American Iron and Steel Institute, Russia):

BDS	DIN 1.4541	AISI321	GOST
X18h9t	X10crniti189	Ae30321	12x18h10t
BDS	DIN 1.4501	AISI316	GOST
X18h10m21	X5crnimo1810	Sae30316	04x19h11m3

Table 4. Welding experimental conditions

№	I_b, mA	v, cm/s	I_f, mA	Type of Cu¹	Type of SST²	PBD³
1	70	0.5	501	b	A	SST
2	65	0.5	501	b	A	SST
3	75	0.7	495	b	A	SST
4	70	0.5	509	a	A	SST
5	80	0.7	501	b	A	Cu
6	85	0.7	501	b	A	Cu
7	82	0.5	478	c	B	Cu (65%)
8	90	0.5	485	c	B	Cu (90%)

¹ types of Cu: a – M1; b – M3 grade I; c – M3 grade II

² types of SST: A – X18H10M21; B- X18H9T

³PBD – predominant beam direction

Table 5. Analysis of the chemical composition (wt.%) of Cu plates observed by optical spectral method

Copper type	Pb	Sn	Ni	Fe	As	Sb	Bi	Zn
M1 (99.9%)	0.0006	0.0003	0.004	0.014	0.001	0.002	0.00004	0.0003
M3 grade I (99.5%)	0.027	0.029	0.006	0.030	0.001	0.001	0.001	0.049
M3 grade II (99.5%)	0.0063	0.0085	0.0024	0.0031	0.001	0.001	0.001	0.011

**Table 6. Standard chemical composition according
BDS of SST in weight % (max) or (from-to)**

Type SST	C	Si	Mn	P	S	Mo	Cr	Ni	Ti
X18H9T	0.12	0.8	2.0	0.035	0.025	0.3	17-19	8-10	5xC%-0.8
X18H10M21	0.15	1.5	2.0	0.04	0.04	2-2.5	17-19	9-11	-

The strength of the welds is tested using Instron 1195 Testing machine and Alfred J, Amsler & Co testing machine. Extensio-meter model G 51 12 M with length L=25 mm is used in the case of Instron machine, while the extensometer at A.A & Co machine is with length L=50 mm. The measurements of the strength are performed at room temperature.

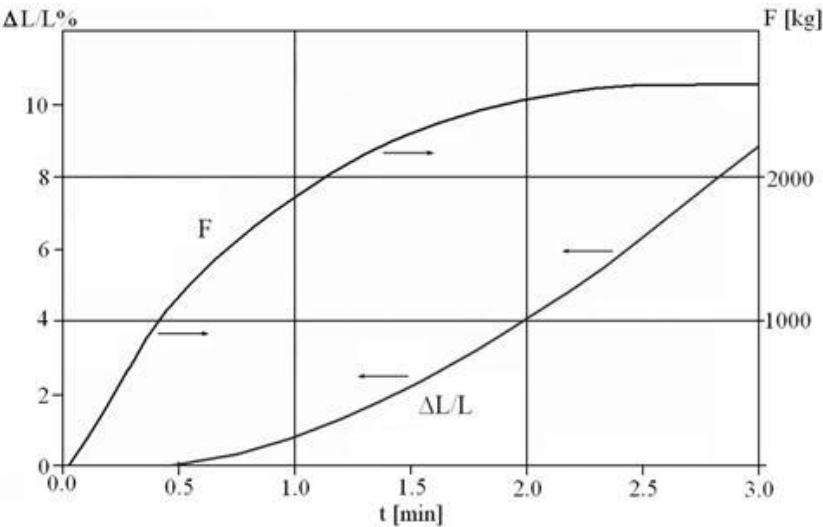


Figure 13. The relationship of the force and the relative extension vs. time for the weld 1 (see Table 4)

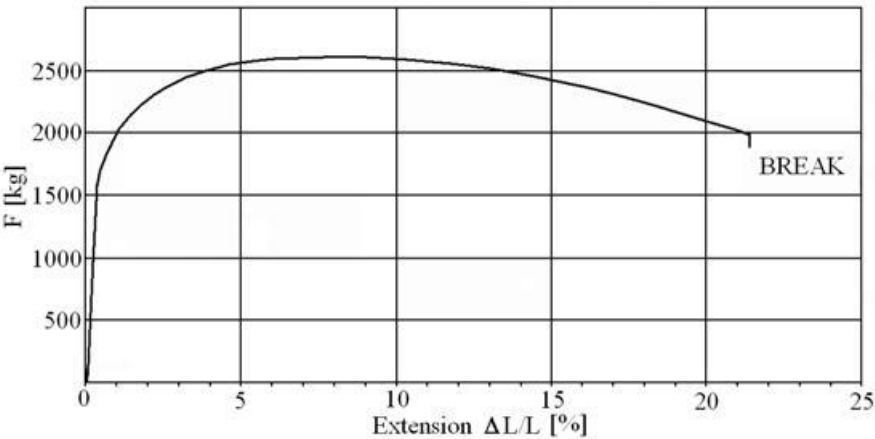


Figure 14. Tensile profile for weld 1 (see Table 4)

The placed in the jaws piece was stretched at a rate of 0.05" per min. During this time the force on the machine set of jaws increases. At the use of the Instron testing machine the relationship of the force and the relative extension vs. time as well as the force vs. relative extension or the force vs. displacement (extension measured in length units). The results, obtained for weld 1, when the beam was preliminary directed toward SST, are given on Figures 13 and 14 (see Table 4 for the welding parameters).

Figures 15 and 16 present results obtained for welds 5 and 8 (Table 4), when the beam was directed preliminary on Cu. The test in Figure 16 was stopped before reaching the breaking point.

The ultimate strength (UTS) and the proportional limit (PL) values are presented at Figure 17 for all the experimental conditions.

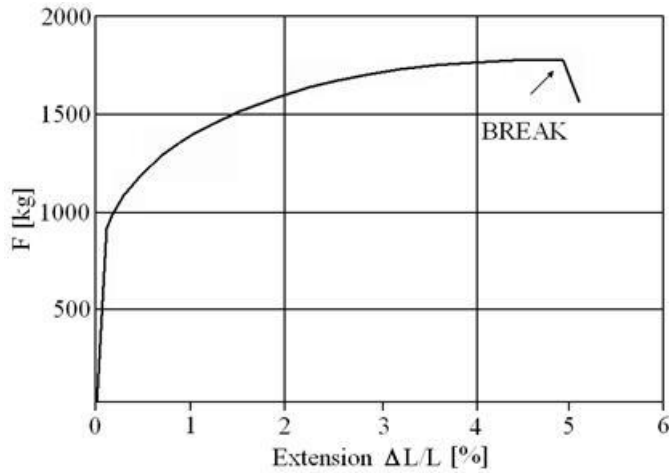


Figure 15. Tensile profile for weld 5 (see Table 4)

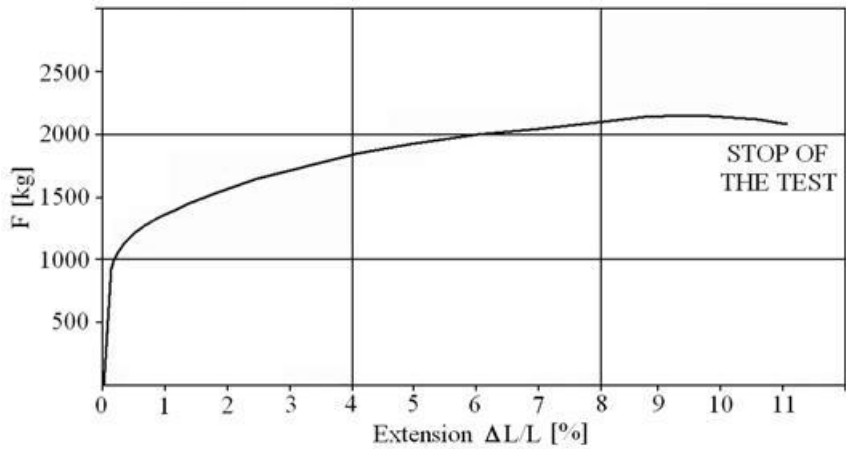


Figure 16. Tensile profile for weld 8 (see Table 4)

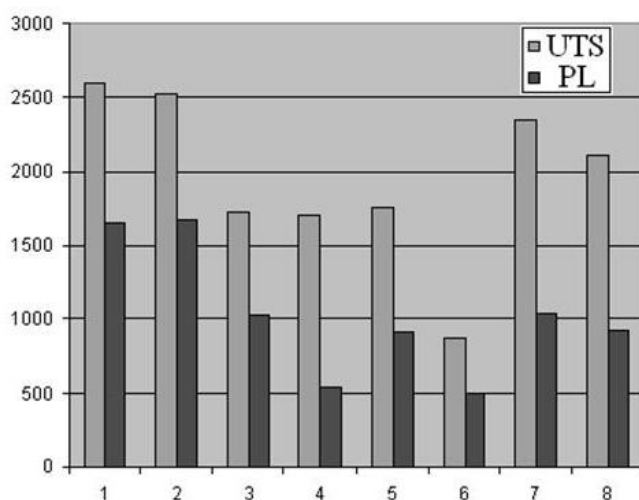


Figure 17. Ultimate strength (UTS) and the proportional limit (PL) for the welds, obtained at 8 experimental conditions (Table 4)

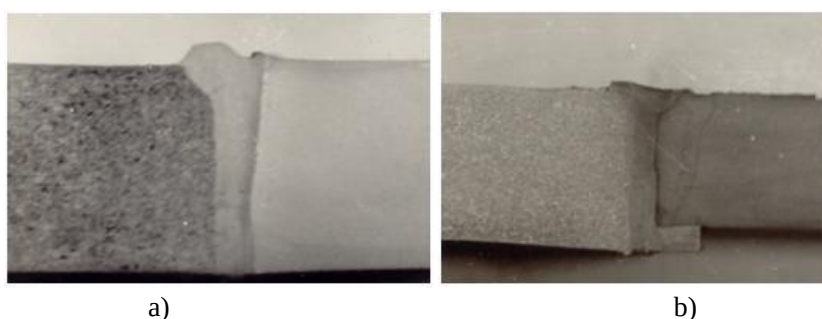


Figure 18. Micrographs of weld 1, weld 8 ($\times 5$)

On Figure 18 and 19 are shown the micrographs of the cross-sections of the welds performed under the conditions of the weld 1 and weld 8, using different level of enlargement. The first weld corresponds to preliminary beam direction during welding on SST (Figure 18a), while at the second weld – the preliminary beam direction is toward Cu (Figure 18b).

On Figure 19 can be seen the mixing of the welded materials in the interface zone.

Hardness distributions of these welds (1 and 8) are shown in Figures 20 and 21. They are measured in each cross-section using Vickers hardness tester with 10 kg load. The welds have satisfactory hardness (not very high). In the other cases of measurements of wider welds intermediate hardness is observed. When copper of type M3 grade I is used a decrease of the Vickers hardness in the thermally affected Cu zone during the process of welding is observed. The use of SST: X18H9T and Cu: M3 grade II is better then SST: X18H10M2 and Cu: M3 grade I from the hardness point of view.

A brief attempt for scanning electron microscope testing was done using SEM JEOL JSM 35 CF electron microscope analyzer (using TRACOR NORTNERN TN 2000 energy dispersion system).

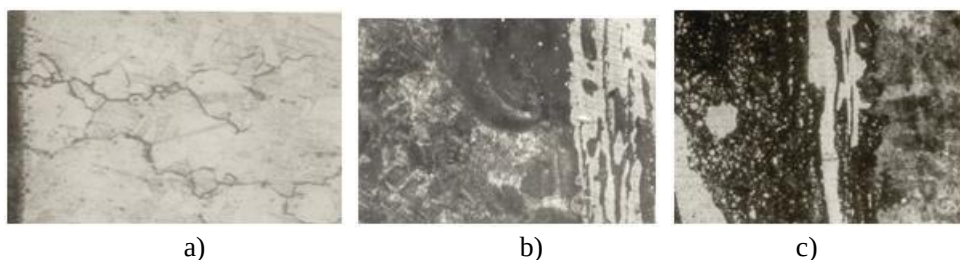


Figure 19. a) Microstructure of weld 1: X12H10M2 $\times 150$;
 b) microstructure of weld 1: interface ($\times 150$);
 c) microstructure of weld 8: interface weld-copper ($\times 125$)

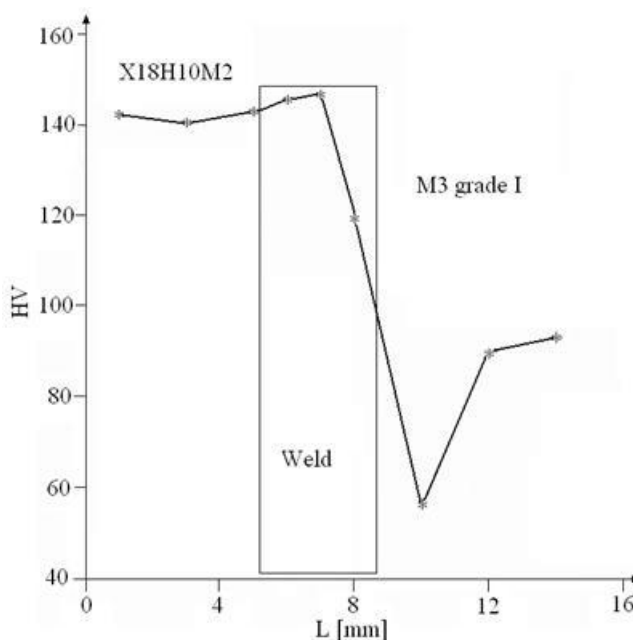


Figure 20. Hardness distribution of weld 1.

In Table 7 and Figure 22 are given the results of analysis along a line in the middle of the cross-section of two welds. It can be seen that in the used tensile tests copper content in the weld does not affect considerably the weld strength. According to the analysis of the welded metal it seems that there is a little vaporization of alloy components. In the given micrograph small SST and Cu drops in the metal can be seen. From our experience in investigating SST composition changes in such drops can be concluded that only Mn in SST drops has the ability to dissipate for a short time in the welding bath.

From some electron microscope examinations and from direct measuring with magnetometer (Ferritehaltmesser 1054, made by institute Dr. Forster, Reutlingen, Germany) small ferrite phase in the welds is observed. For the cross-section of the seam, produced at welding conditions of weld 1, this phase was at the top part of the weld. For the cross-section of weld performed at conditions of weld 11 (weld 10), the phase was at the root part.

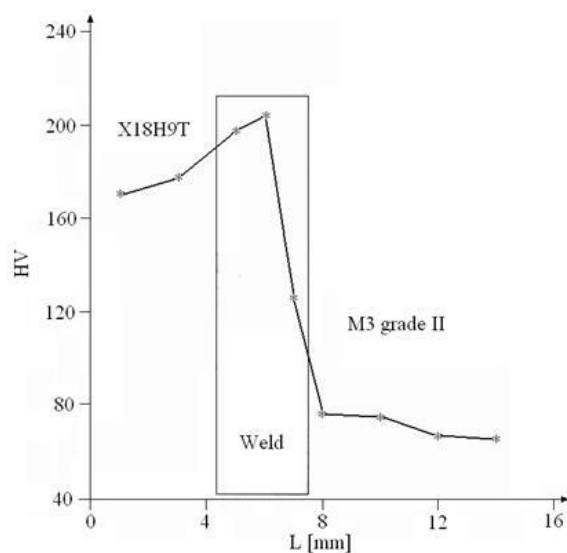
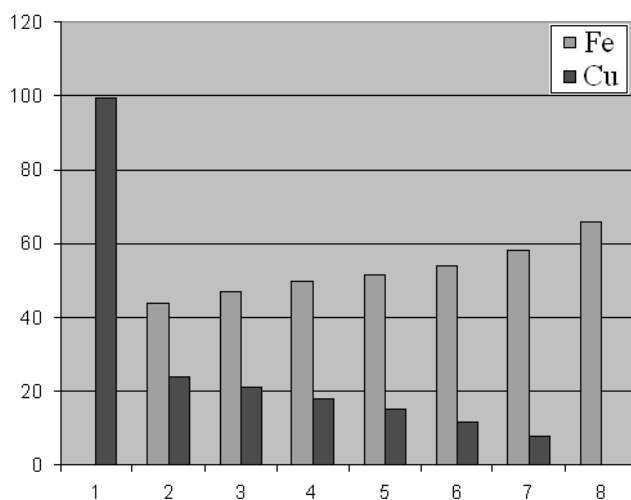


Figure 21. Hardness distribution of weld 8

Table 7. EDS analysis of weld 1 (wt%) averaged on the analyzing spot 0.8×0.8 mm²

Points	Components								
	Cu	Fe	Al	Si	Mo	Cr	Mn	Ni	Co
1	99.5								
2	60.08	23.30	0.88	0.92	0.86	6.62	0.87	4.46	
3	48.71	33.99	0.72	0.72	1.01	8.91	0.9	5.05	
4		66.36	0.27	0.45	2.3	16.86	2.31	10.68	0.72

Figure 22. EDS analysis of weld 2 (wt%) averaged on the analyzing spot 0.2×0.2 mm² (in 8 points)

This result is important for the special use of designed calorimeter working in magnetic field. The quantity of this phase is small: from 1% to 5% in the part of the weld.

The tensile tested specimens of electron beam welded joints of SST (X18H10M21 and X18H9T) and copper (M1 and M3) showed sound properties and were mostly fractured in the base metal. The absorbed energy of the weld metal in the case of beam predominately directed on Cu exceeds the assorted energy of the case of beam directed predominately on SST. The welding bath (and the surface) is more of the weld situated predominately in SST. The increase of copper contaminant concentration causes vaporization, boiling spattering and splashing of the welded material.

An important conclusion in the working conditions was that welding deep must be on full sample thickness to have guaranties that stress concentration shall be avoided.

The investigation made was directed mainly toward the increase of the knowledge about the process, rather than making technological instructions. Due to the difference in the weld depth in the case of Cu and SST and difficult control of the exact beam shift towards the position of the component contact before the welding in the real conditions, the choice of the welding with beam direction on Cu must be recommended.

PHYSICAL PROCESSES AND HEAT TRANSFER MODEL OF ELECTRON BEAM WELDING

The primary interaction of power beams with matter is manifested through the process of surface heating. Due to kinetic energy of accelerated electrons, converted in energy of electrons or at higher energies of primary electrons as energy of atom clusters of the target material (separately in the first time) and for a time of order of 10^{-10} s these two systems come to equilibrium state and a elevated temperature. As a result local melting and some evaporation can be observed. At EB welding melt pool in place of welded edges is produced. For control of the cross-section of the obtained welds an adequate physical and heat model of processes in the beam/sample interaction zone is needed.

Nevertheless of the wide use the knowledge of the physical processes that take place during the interaction of the electron beams with metals in the case of electron beam welding, are still incomplete. Accordingly there is an obvious lack of theoretical models describing adequately the appropriate operating mechanisms. The main reason for this is the complex nature of the deep penetration of electron beams during electron beam welding. The explanation of the deep penetration of the intense energy beams into the treated material is connected with the generating of a key-hole (crater or plasma cavity) within the liquid metal welding pool through which the energy beam entering in the heated samples.

The processing results at electron beam welding with a high power beam are strongly affected by the complexly interconnecting physical phenomena within both the plasma cavity and the welding pool, namely:

- i. energy dissipation and the phase changes of the materials in the interaction region;
- ii. neutral and ionized gas atoms are emitted from the heated sample. In the case of deep penetrating beam the metal vapor and outgazed molecules through the channel and from orifice of the plasma cavity of time-variable shapes and dimensions flows to the vacuum;

- iii. interaction of electron beam with vapor phase; change of the beam focusing parameters(respectively angular and radial distributions of the beam current)due to the electron scattering on the products of evaporation and the beam space charge neutralization or overcompensation;
- iv. heat transfer in the interaction region of the beam with the metal samples and near situated zones;
- v. liquid metal flows in the molten pool and surface tension changes on free liquid metal surfaces. In liquid metal waves and instabilities can be observed.

The processes of the generation of the cavity (filled with vapor and plasma) and the behavior of the liquid metal on cavity walls - i.e. the drilling mechanism of the deep penetrating electron beam are between the fundamental subjects of the investigations of the physical processes during EBW (Figure 23, Figure 24). The dynamics of the plasma cavity shape and the geometry of the welding pool has been studied experimentally [18-20, 28-31]. By X-ray observation it was shown [18, 28] that both the shape and the dimensions of the beam crater vary during the welding whereby the cavity entirely or partially but frequently is filled of the liquid metal by welding pool. The frequency of the filling is of order of a few Hz and can be observed on the weld surface as so called ripple weld surface as from spiking in the weld root on the metallographic longitudinal cross-sections of the EB seam. Alternatively, high speed camera [19, 20, 28-30] was employed for the study of welding dynamics.

In [31] CCD camera is used for high degree precision to follow the behavior of the weld pool and keyhole during electron beam welding ((Figure 25)). The shapes of the welding pool and of the keyhole are apparently asymmetrical. Front side of the welding pool has a few liquid metal compare with the back side of the pool (Figure 26). The dimensions of the cavity are more varying with time than the same characteristics of the welding pool. In other experiments [32] copper backing plate (1-2 mm thick) were used during welding at the full beam penetration of the sample plate (Figure 27 and Figure 28). After the welding metallographic investigations of the longitudinal weld cross-section was done (Figure 29).

The satisfactory description of all these processes is additionally hampered by the fact that the required general equations, as well as the corresponding initial and boundary conditions have not yet been fully formulated. The values of involved material characteristics are not exactly known too. That is why the physical and mathematical models proposed in the literature are very simplified and generally based on the assumption for quasi-stationary plasma cavity and the welding pool.

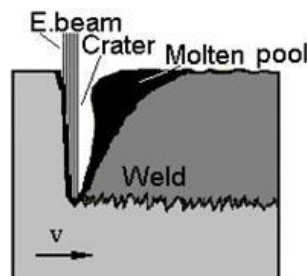


Figure 23. Schematic longitudinal cross-section of the EB welding process. The deep penetration of the weld is due to creation of crater (keyhole) in the welding bath

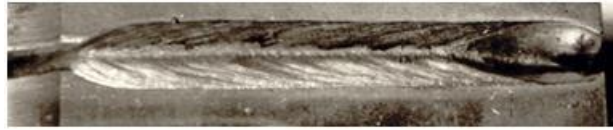


Figure 27. Weld face.

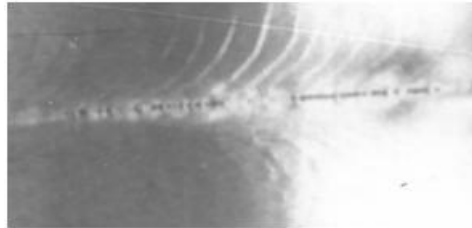


Figure 28. Back side of the weld, that penetrating through the whole height of the welded pieces.

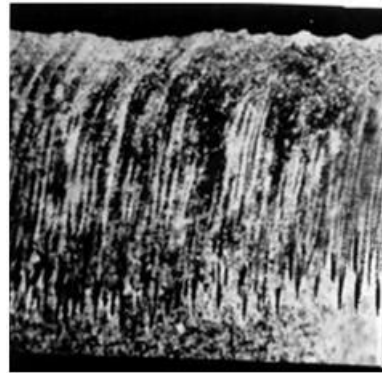
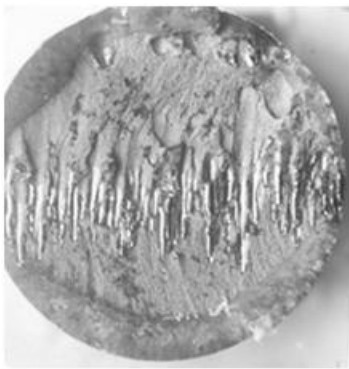


Figure 29. Quasi-periodic character of liquid metal transfer and spikes at weld root

A few of papers discussed the observed instabilities as result of interaction between the metal vapors, the electron beam and the cavity walls [33-35]. The principal results of these investigations are:

- (i) The geometry (width, depth, volume) of the molten welding pool formed in the work-piece by the continuously operating (CW) beam is not constant during the seam production. It depends on: the beam power density (or exactly on angular and radial energy distribution of the beam in the interaction zone), the welding conditions and the physical properties of the materials.
- (ii) The powerful beam penetrates deep in work-piece through a crater (keyhole) created in molten pool due the reactive force and pressure of generated vapor. The metal melts ahead of the hole and solidifies behind it after the beam has passed. The keyhole allows a more effective and directed beam energy transport and absorption.
- (iii) Less than 1% of the material of the welded sample is evaporated or blows off through scattered droplets. This quantity is less then keyhole volume. This means that bigger fraction of the molten metal is only shifted by the dynamic and static pressures of the vapors;

- (iv) The dimensions, shape of the welding pool and of the cavity undergoes quasi-periodical variations with time. Both the shape and the dimensions of the cavity are more varying with time than the same characteristics of the welding pool.
- (v) The material removal resulting as drilled cavity permit to gain the depth of a few millimeters from the target surface within a few milliseconds and the depth of several centimeters within tens or hundreds of milliseconds;
- (vi) The cavity is near to periodically filled with liquid metal. That partial or total filling of the hole by liquid metal is due to insufficient vapor pressure and vapor reactive force are insufficient to counteract hydrostatic and hydrodynamic forces and surface tension of the liquid metal. The frequency of partial or total filling of the crater by liquid metal depends on the metal thermal parameters as well as on the beam power and focusing of the beam (the position of the focal spot regarding the welded surface). As a result is determined the time of stay (holding) of the penetrating beam to any depth of the welding crater. At probability near to zero achieved total seam depth saturated.
- (vii) The filling of the crater is important also for the observed intensive weld metal mixing. In case of crossing the test rod position of a upper rod from other metal than base sample metal - for example cooper in steel welded sample - the root region of the weld is reached from the cooper after two-three strong pulses of the key-hole shape;
- (viii) The metal evaporated from the front side of the plasma cavity, as well as a portion of back scattered electrons, at reaching the rear side of the welding pool are elevate temperature and exerting the local pressures on the back liquid metal walls of the crater;
- (ix) The mass transport of the liquid metal from the melting front to the solidification phase boundary of the welding pool occurs around the plasma cavity walls through side wall of the crater (80%-90%) and in directions of the depth of the welding pool (20%-10%). The proportion of these portions is connected with the depth of the weld. At the deep levels of the seam more and more of liquid metal is going through the region of the root of the weld. In this region pulse character of the liquid metal fluxes is typical. The liquid metal fluxes have velocities of order of 2 - 10 cm/s. The molten metal velocity on the backside of the pool is in inward direction in case of a not fully penetrating beam through the sample thickness. This inward movement of the liquid metal gives the weld face height observed usually at the surface of such welds. The liquid fluxes in the welding pool are turbulent. The weld metal at distance 5 - 8 mm behind the crossed by beam test rod have a uniform and near to the base sample composition.
- (x) The front side of the welding pool has a few liquid metal compare with the backside of the quantity pool. Accordingly the variations of the positions of the welding pool walls are bigger for the rear side of the crater wall. The differences in the surface temperatures of the crater walls are resulting in differences of surface tension, which is responsible, together with the reactive pressure of evaporating atoms and hydrostatic forces for the liquid metal movement and surface oscillations. The roughness of the front side determine the angle of wall illumination by beam and in this way control the local power density distribution, the subsequent local rate of evaporation and the pressures (reactive and stationary), the initial velocities of the

removing liquid metal - accordingly also with the slope of the melting/solid phase boundary in this region.

- (xi) It should also be noted that the mass transport of liquid metal in the welding pool influences largely the process of metal crystallization, which determines both the shape of the seam's cross-sections and the presence of defects within its bulk. It has been shown (in [35]) that formation of such non-uniformity could be due to the generation of capillary waves in the welding pool;
- (xii) There is a scattering of the electrons at it's collisions with the vapor atoms and a subsequent focusing of the beam due to the compensation of the negative space charge in the electron beam by the generated positive ions and a magnetic pinch of the beam with neutralized space charge [36] as well as a gas focusing of the beam due to the overcompensation of the beam negative charge at higher densities of metal and gas molecules in the welding crater [37, 38]. In this direction of increasing or redistributing the local beam power density also can play role the reflection of beam electrons by crater walls and by plasma potential distribution near the wall of the welding cavity. Density of the energy in central part of the electron beam is controlled by the gas focusing, the electron scattering from both: the crater walls and the plasma potential drop around these walls. The focused portion of the beam in the weld root region produce intensive vaporization of the solid material in the bottom of the crater. The diameter of the drilled holes at the weld root are from several microns to some tens microns and the small heat affected zone around these holes speaks for the higher energy density and short working time of the beam there. At condition of the deep penetration of the beam the diameter of the crater in the weld root is smaller than in the upper part of the weld and due to this the spiking of the weld root occurs. The ring oscillation of the beam with a small amplitude or use of a beam with a minimum of the radial energy distribution around the beam axis (tube dispersion of the current and energy of the beam generating by a cathode with central hole) are increasing the welding root crater diameter and decrease the weld spiking;
- (xiii) The upper part of the crater is formed by molten metal removal (due to the mentioned reactive force of evaporated molecules and of their pressure), the lower portion of the crater near the bottom of welded plate (i.e. the weld root) is formed by vaporization removal of the sample material. The upward flow at the backside of the welding pool is superheated delayed the solidification and extended the welding pool in the upper part of the weld. The cross-section of the weld in that part is similar to the nail head. Opposite-due to gas focusing EB welds holds a tip-like root.
- (xiv) The interface between the melted metal/vacuum is a deformable free surface. The evaporation process and the temperature gradients surface tension affect to the dynamics of the shape of this surface and of the liquid metal fluxes;
- (xv) The shape and entrance of the cavity control as the vapor flows through them, so the beam power density and energy distribution on the cavity wall surface. The balance between the pressure and the recoil force of the vapor of one side and the surface tension on keyhole walls together with the gravitational and the dynamic forces in molten metal column from other side govern the melt movement and keyhole stability. The liquid metal fluxes are influencing crater shape and dimensions (and filling), the heat transfer as well as the energy input distribution through local irradiation of the keyhole wall unevenness or by beam shielding.

In such a way a complex interaction of many connected phenomena in work-piece vapor and of liquid pool dynamics govern the formation of the electron beam weld. The determination of the weld geometry is done by statistics of these variations.

The analysis of the energy dispersion processes is a way to evaluate the geometry characteristics of the weld. In the case of EB welding (heat treatment) of semi-infinite sample with electron beam, characterized with mean power density on work-piece surface less than a critical power density (of order of 10^5 - 10^6 W/cm²) a semi-spherical fusion zone can be obtained (Figure 30) due to near to point heating source. The same shape of weld cross-section is shown below for thin plate butt seam. In the general case of EB welding with a powerful deep penetrating beam (mean power density of which is more than mentioned critical value) the energy flux density that is absorbed on keyhole walls is a complex function of the coordinates (Figure 23).

On base of the solution of Rosenthal/Rykalin general theory of the heating of a infinite sample by movable source, using the electron beam characteristics, formulae and nomograms for the evaluation of the weld geometry parameters at electron beam welding of thin plates [39,40], as the depth of melted material at EB surface thermal treatment [37], were derived. The coincidence with experimental result is good.

In the case of disregarding the keyhole in the melting pool, the same model of heating can be assumed as an approximation [41-43, 44 – for laser welding]. Other possibilities for evaluation the weld geometry in case of deep penetrating powerful beam are models [42, 45-49] utilizing the ideas for heating by moving: the sum of linear and point heat sources as a cylindrical or conical steady, continuously operating heat sources.

From the heat dispersion calculations based on the heat balance assuming a quasi-stationary temperature distribution one are able to obtain approximately weld parameters and to explain many process features. The analysis of the energy dispersion processes is a way to evaluate the geometry characteristics of the weld. In the case of EB welding (treatment) of semi-infinite sample with electron beam, characterized with mean power density on work-piece surface less than a critical power density (of order of 10^6 W/cm²) a hemi-spherical fusion zone can be obtained due to near to point heating source. In the general case of EB welding with a powerful deep penetrating beam (mean power density of which is more than mentioned critical value) the energy flux density that is absorbed on keyhole walls is a complex function of the coordinates. In order to simplify the problem of non-known and non-steady distribution of the real heat source a quasi steady state heat source can be assumed.



Figure 30. Weld cross-section at EB power density less than 106 W/cm²

The power distribution over the spot into the depth at EBW is a complex function of the coordinates and time, due to generation of a crater of variable shape into the molten metal, through which the moving electron beam with a changing during interaction angular and radial energy distribution penetrate in the treated sample. In order to solve the problem of non-known and non-steady distribution of the real heat source a steady-state model involving a linear, uniformly distributed heat source in the moving with the beam respectively to sample coordinate system [42, 45-49] is used (Figure 24).

The solution of thermal balance equation at heating a sheet of thickness H from a linear moving thermal source of a constant intensity P , moving with speed V , assuming no phase changes in the sample during heat transfer, at known material physical parameters: thermal conductivity λ , thermal diffusivity a ($a=\lambda/(C\rho)$, where C is the specific heat and ρ is the sample density), will take the form [50]:

$$T(r,x)=P/(2\pi\lambda H) \cdot \exp(-Vx/2a) \cdot K_0(Vr/2a) + T_0. \quad (1)$$

There r is the radius-vector and x and y - coordinates in a moving together with the heat source co-ordinate system, and y is the distance from the EB movement axis x (coinciding with the V), $K_0(Vr/2a)$ is the modified Bessel function of second kind of order zero, P is also the EB absorbed energy input (beam energy P_b after correction for energy losses by back scattered and by secondary electrons). V is the welding speed and T_0 is the initial sample temperature. In order to minimize the effect of the temperature dependence of thermal constant the values of λ , C_p and a are taken at a intermediate temperature (between T_0 and T_m , where T_m is the melting point) and the heating process is assumed to be independent of the temperature. From the equality to zero of the first derivative of equi-thermal curve $T(x,y)=T_m$ at the maximal distance y_m , at which the temperature elevation is reaching that value the equation (1) gives a new equation (2), written in terms of the dimensionless maximal temperature θ_m and of Péclet numbers for the coordinates :

$$\theta_m = \frac{2\pi\lambda HT_m}{P} = K_0\left(\frac{r \cdot V}{2a}\right) \cdot \exp\left[\left(\frac{r \cdot V}{2a}\right)\right] \cdot K_0\left(\frac{r \cdot V}{2a}\right) / K_1\left(\frac{r \cdot V}{2a}\right), \quad (2)$$

where K_1 are the modified Bessel functions of second kind, of order one.

More practical is the function $\theta_m(yV/2a)$, which can be found by iterations from (2), and from $y = B/2 = r \cdot \sin(\varphi)$, where B is the weld width. This function is given in Figure 31 for a range of $Vy/2a$ appropriate for EB welding small distances y . Using values of known θ_m for given P and H the curve shown in the Figure 31 gives possibility to obtain the weld width for a concrete material. Opposite, at using choused width value one can obtain the weld depth value.

Another presentation $Y(X)$ of that relation is given also in Figure 31, where the tilted normalized coordinates are: $X=P/(H\lambda T)$ and $Y=VB/(2a)$.

An experimental test of these relations was fulfilled by EBW of thin plates of thickness in the region 0.4-4 mm from various metals, at welding speed from 0.5 cm/s to 2.5 cm/s. The obtained results confirm the assumed approximations. On that base had been developed equation (3) and nomograph (Figure 32) for evaluation of the beam current I for EBW of thin

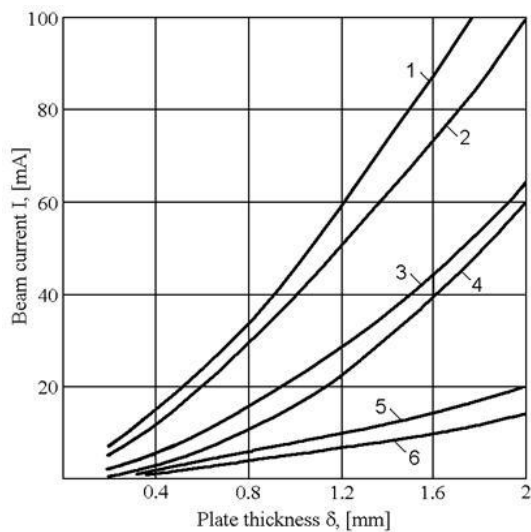


Figure 32. Dependence of current at butt welding of thin plates on plate thickness δ . 1- Mo, 2 – Cu, 3 – Ni, 4 – Covar, 5 – stainless steel, 6 – steel 08KP ($\lambda=0.51$ W/cm deg; $\sigma=1.24$ cm²/s; $C=0.52$ kJ/kg deg.) $U_a=30$ kV, $\xi=0$

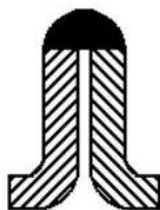


Figure 33. Butt welding

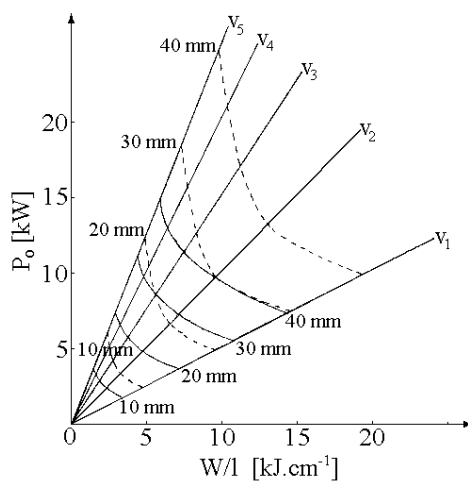


Figure 34. Relation h ($P, W/L$); $V_1=0.5$ cm/s; $V_2=1$ cm/s; $V_3=1.5$ cm/s; $V_4=2$ cm/s; $V_5=2.5$ cm/s; Continuous line (—) - for $H=1$ mm; dashed line (- -) - for $H=2$ mm

Other possibilities for evaluation the weld geometry in the case of deep penetrating powerful beam is to use models [50-55] utilizing the ideas for heating by moving: the sum of linear and point heat sources as a cylindrical or conical steady, continuously operating heat sources.

In all cases to evaluate the volume of the molten metal produced by energy beam per one unit time – 1 s (namely, the product of the desirable weld cross-section multiplied to the weld speed and the material characteristics) one needs the thermal efficiency of heating process.

The dimensionless thermal efficiency η_t defined as a ratio between the energy P_f absorbed and spent for heating of the metal of the volume of the weld up to melting temperature (including the fusion heat), and total beam energy converted in the thermal energy P .

$$\eta_t = P_f / P = V \cdot F_w \cdot S / P, \quad (4)$$

where v is the welding speed, F_w is the cross - section area of the melted zone, S is the heat content per unit volume of the material of work-piece during heating from room temperature up to fusion temperature T_m ($S = C_p \cdot T_m + H_f$, being C_p the mean specific heat for the temperature range between the room and fusion temperatures. H_f is the heat of fusion). The thermal efficiency value accounts for losses due to the following processes and mechanisms:

- (i) thermal conductivity towards cold sample regions,
- (ii) over-heating of the weld metal above melting temperature,
- (iii) heat transfer by vapor-gas flow leaving the welding crater,
- (iv) radiate dissipation of heat from weld surface.

It is easy to see that the thermal efficiency η_t can be evaluated as ratio Y/X . In this way Figure 31 presents a theoretical expectation for the character of the thermal efficiency changes at variation of the process parameters at used thermal model. It is evident, that theory give constant values of η_t at high powers and welding velocities (see the straight line $Y=0.484X$ observed in that part of the curve).

Figure 35 gives the transformed plot of $\theta_m (BV/2a)$ for stainless steel - namely $BV(P/H)$. On that figure three strait inclined lines presents 100% (upper inclined line), 50%(central inclined line) and 20% (below situated inclined line) of thermal efficiency of the EBW. The theoretical limit of that efficiency is 48.4% [50] seen as near to straight line part of the theoretical curve $BV(P/H)$. The experimental data of two wide studies of geometry characteristics of EB welds in stainless steel (there are partially presented more than 140 experiments) [2] denoted by points. The values of the experimental data cover the ranges of $P/H(1.33-10)$ kW/cm and of $BV(0.1-0.75)$ cm²/s.

The discrepancy of the experimental points and theoretical curve on Figure 35 are due to the assumptions:

- (i) Use a steady state instead a non-stationary heat source. Note that idea for a model using a non-stationary (variable or oscillating) intensity of the EB heat source in work-piece was discussed in [10, 56, 57].

- (ii) Presence of keyhole and due to dispersed heat source (in welding bath too) instead a concentrated heat source along the beam axis.

An experimental confirmation of the idea for the non-steady heat source, operating in the welding bath (i) was observed by direct measurements of the temperatures in the points placed at distances $y_1 = 0.01$ cm, $y_2 = 0.015$ cm, $y_3 = 0.02$ cm, $y_4 = 0.025$ cm and $y_5 = 0.03$ cm respectively from the line of the movement of heat source /beam/ using the W/W-Re thermocouples. The beam U_a was 60 kV, 50 mA and the welding speed was 1 cm.s^{-1} . The measured dependencies of the sample $T(t)$ are shown in Figure 36.

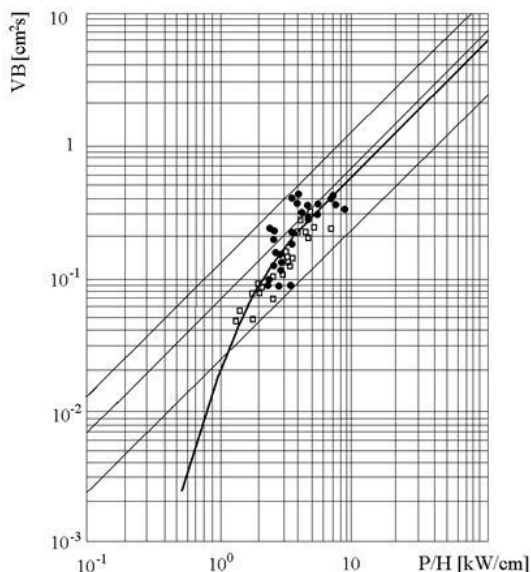


Figure 35. Comparison between the experimental and theoretical data for parameters VB and P/H as well as for the thermal efficiency

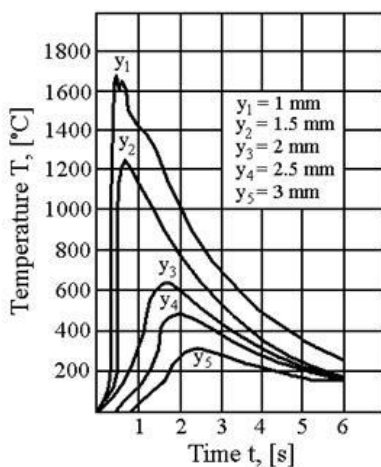


Figure 36. The measured dependencies of the sample temperature in time

Non-monotonous character of the temperature changes in the vicinity of the maximum of the first curve in the initial stages of the quenching of the material can be shown in Figure 36. The thermocouple is fixed just outside the weld but near the fast moved liquid - solid boundary. Analysis of great number of such curves showed that they all contain a component representing periodical temperature changes of a frequency of 3.5 - 5 Hz. Therefore, the real heat source operating within the welding bath has variable components too. The lowest frequency component is responsible for non-stationary changes of the temperature cycles. Avail from the welding pool (curves for bigger distances y in Figure 36) owing to metal capacity, the heat waves originating by the variable component of the non - stationary thermal source attenuate and observed thermal cycles are similar to the cycles, generating by a moving heat source of constant intensity.

The approximation (ii) of disregarding keyhole and of the real distributed volumetric heat source can be evident as follows.

In [56] paradoxically the reconstructed through calculations heat sources intensities by the experimental dimensions of the melted and heat affected zones prove to be different. That can be explained due to the lower distances between these zones and the cylindrical keyhole walls, where heat is absorbed. The phase transitions in solid state and turbulent flows as the reason for the variations in the shape of the crater in the melt together with the heat capacity of superheated liquid metal are additional reasons for that discrepancy as well as for differences between the theoretical curve BV (P/H) (see the solid curve on Figure 35) and the experimental weld geometry characteristics, presented by points. The same reasons lead to the shown increase of the thermal efficiency values of the some regimes of EBW that are higher than the theoretical limit 0.485 for the linear movable heat source there.

Figure 37 is a presentation of the ranges of the observable weld depths versus the EB power. The parameter window reflects the region of P/H observed in Figure 35. At powers bigger than 30-40 kW the maximal weld depth can be realized in the horizontal position of the beam and the welded sample is moved in vertical direction from the top to the bottom of the vacuum chamber and some additional care for preventing the out flow of the molten metal may be needed.

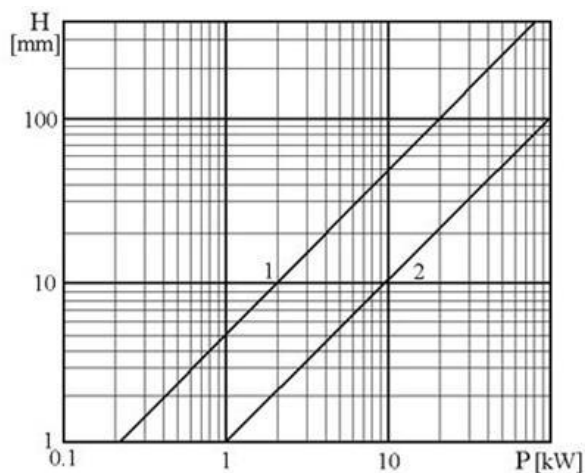


Figure 37. Welding depth ranges versus EB power at wide varieties of welding speed (0.2-15 cm/s). Curve1 :Hmax at P/H=2 ; curve2:Hmin=10

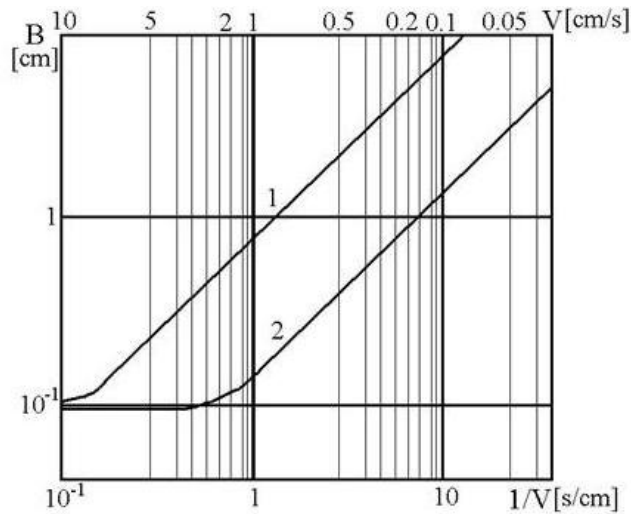


Figure 38. Welding width ranges versus welding speed V or $1/V$. Curve 1 presents $B_{\max}$ at $VB=0.75 \text{ cm}^2/\text{s}$ and curve 2: $B_{\min}$ at $VB=0.15 \text{ cm}^2/\text{s}$

Figure 38 shows the weld width range versus welding speed V (and $1/V$). The change of dependence character at high welding speeds to a limited value of the width is due to real minimum of the crater dimension.

EXPERIMENTAL INVESTIGATIONS

One of the experiments, considered in this chapter, is the electron beam welding of samples of austenitic stainless steel (SSt), type 1H18NT. The geometrical conditions of the experiments are shown in Figure 39. There z_p is the distance between the sample surface 3 and the main surface of the magnetic lens of the electron gun 1, z_o is the distance between the focusing plane 2 and plane 1 in the gun. The focusing parameter dz is the difference between these two distances. The values of z_o are determined from measurement of the focusing current of the electron beam. In the experiments, an inclined thick sample is treated along its length by an electron beam. The following operating parameters: weld velocity (v), focusing current of the beam – distance from the main surface of the magnetic lens to the beam focus (z_o), the distance to the sample surface (z_p) and beam power (P) are varied. In Table 8 are presented the regions of variation of the process parameters during performed experiments, as well as the performance characteristics of the welds: weld depth H , mean weld width B , the ratio of the electron beam power and the weld depth P/H , the product welding velocity v and the mean weld width vB , the thermal efficiency. The accelerating voltage is 70 kV.

Every sample is welded for values of the parameter dz over the chosen range using an electron beam with constant parameters, in particular, the angular and radial distributions of the beam current for a given P . Every welded sample is cut in at least three planes. These planes lie in the vertical direction coincident with the electron beam direction. This allows measurement of the weld depths and observation of the weld cross-section shapes. 81 experimental weld cross-sections were investigated [58].

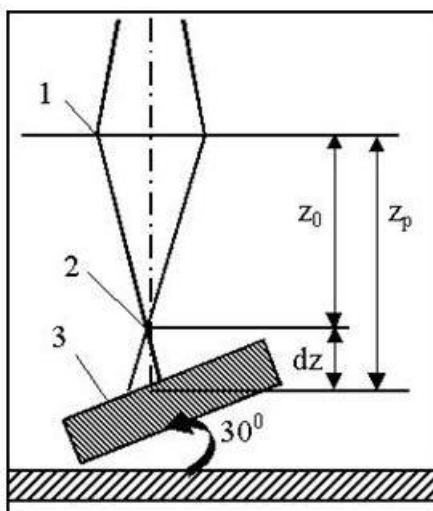


Figure 39. Geometrical conditions of EBW experiment: 1-magnetic lens; 2-electron beam; 3- welded sample

Table 8. Experimental conditions for EBW of Stainless Steel

Parameters	Dimensions	Min	Max
Beam power	P [kW]	4.2	8.4
Welding speed	v [cm/min]	20	80
Distance 1*	z_0 [mm]	176	276
Distance 2**	z_p [mm]	126	326
Focus parameter	dz [mm]	-78	62
Weld depth	H [mm]	4.9	43.8
Mean weld width	B [mm]	0.9	5.5
P/H	P/H [kW/cm]	1.333	8.571
vB	vB [cm ² /min]	3.2	25.6
Thermal efficiency	η_T	0.22	0.56

*Distance 1 - the distance between the EB gun and the beam focus;

**distance 2 - the distance between the EB gun and the sample surface

Electron beam welding (EBW) of steel 45 (St45) is performed at the experimental conditions shown on Figure 40. The welded samples are placed on 30° towards the horizontal plane and are moved by the manipulator in the vacuum technology chamber. The EBW of St45 is performed in a serial EBW installation "Leybold-Heraeus" ESW300/15-60, at acceleration voltage of 50 kV. The moving of the sample results in different distances between the magnetic lens of the gun and the sample surface (Z_s). The distance between the focus of the beam and the main surface of the magnetic lens of the electron gun (Z_o) is held constant and equal to 300 mm (the focusing current is 478 mA) for St45. In such way the moving of the sample results in different distances between the magnetic lens of the gun and the sample surface (Z_s) being in the region (from 228 to 362 mm) and the start of the weld is near to position 1 inserted on welded sample in Figure 40. The acceleration voltage is 50 kV.

The beam current (I_b) is changed on four levels: 30, 66, 100 and 133 mA. The welding speeds (v) are 0.5, 1 and 1.5 cm/s. The welded samples were rods of rectangular cross-section (20 mm $\times$ 34 mm and 25 mm $\times$ 34 mm) and length about 335 mm.

After the processing a blind weld the sample rods are cut to pieces through the inclined planes (as beam penetrate in sample material – signed on Figure 40 by 1, 2 and 3) and processed afterwards. The HAZ geometry parameters for St45 are easily distinguished (Figure 41) due to the hardness variation on the mechanically polished cut surface. For determination of the weld geometry parameters in this case metallographic images, like the one shown on Figure 42 for $P=5$ kW, $v=1$ cm/s, $dz=-7$ mm, are obtained. Due to suitable chemical etching of the polished weld cross-section on this photography two zones are clearly seen. They are: i) the surface area of weld fused zone (the inner part), and ii) the HAZ (presented by changing color areas around fused zone, situated up to beginning of the black structure elements).

The range of the values of these process parameters during the performed experiments are presented in Table. 9. The negative values of the focusing parameter correspond to a position of the focus below the sample surface.

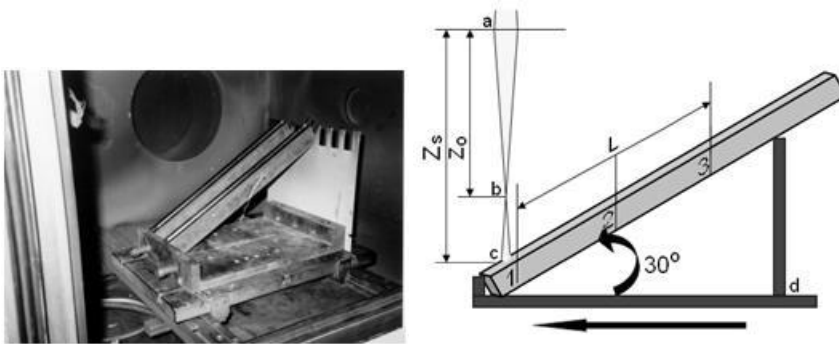


Figure 40. Experimental conditions: a) main surface of the magnetic lens of the electron gun; b) beam focus (or beam waist); c) surface of the sample; d) manipulator and EBW vacuum chamber



Figure 41. The heat affected zone geometry at beam current 100 mA, welding speed 0.5 cm/s and the distance to the sample surface: a) 238 mm in surface 3 (Figure 40); b) 352 mm at surface 1a (Figure 40)]

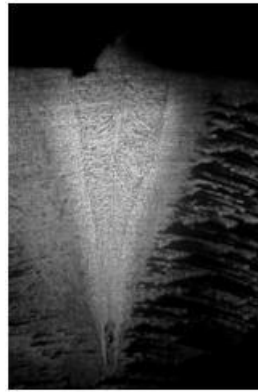


Figure 42. Metallographic etch of the cross-section of a weld (St45), where the weld geometry and the heat affected zone are clearly distinguished

Table 9. Experimental process parameter ranges

Process Parameter	Dimension	Coded	Tolerance limits	Stainless Steel		Steel 45			
				Min	Max	Weld		HAZ	
						Min	Max	Min	Max
P	kW	x_1	$P \pm 2\%$	4.2	8.4	3.3	6.65	1.5	6.65
v	cm/min	x_2	$v \pm 3\%$	20	80	30	90	30	90
dZ	mm	x_3	$dZ \pm 2$	-78	62	-72	62	-72	62
z_o	mm	x_4	$z_o \pm 1$	176	276				
z_p	mm	x_5	$z_p \pm 1$	126	326				

Table 10. Chemical composition of steel 45 in %

C	Si	Mn	P	S	Cr	Ni
0.42-0.50	0.17-0.37	0.50-0.80	≤ 0.040	≤ 0.040	0.25	0.25

1)

Table 11. Physical properties of steel 45

T [K]		300	400	600	800	1000	1200
λ [W/mK]	$57.321 - 0.026959 T$	48	47	41	37	32	23
C_p [kJ/kgK]	$0.2612 + 0.0007754 T - 0.00000042 T^2$	0.469	0.506	0.521	0.660	0.616	0.577
ρ [kg/m ³]	$7799.33 - 0.037778 T$	7788	7784	7777	7769	7762	7754

The chemical composition of steel 45 is given in Table 10. The content of As $\leq 0.08\%$ and residual copper content Cu ≤ 0.25 are acceptable. Steel 45 is tempered at temperature 850-900 °C. The melting temperature is 1403 °C (solidus). In Table 11 are presented the thermally dependent steel 45 material characteristics: thermal conductivity λ , specific heat capacity C_p and metal density ρ .

STATISTICAL APPROACH

A review of multi-response surface methodology is given in [59, 60]. For achievement of choice of operating conditions for obtaining concrete parameters of the EB seam or for assuring the optimal parameter of the welds the multi-response optimization methods: graphical optimization and desirability function approach [61] are used. On the stage of approbation and testing a lot of efforts are usually spent on the search for the optimal parameters of welding. The commonly used method for it remains still so called 'parameter welding' containing great number of model welding experiments, the main purpose of which is to determine the boundary of applicability of new methods and the best regimes for some particular application. In order to improve the quality of the welded product in mass production (to decrease the deviation from the target value of the performance characteristic) a model approach is applied. The variability of the welded features as a result of the errors in the process parameters, defined through the tolerance intervals, is considered. Models [62] describing the (i) mean value and (ii) the variance of the weld depth and width in mass production are estimated.

In order to apply methods for optimal process parameter choice, models describing the influence of the process parameters on the performance characteristics of the welds obtained at EBW are needed. Statistical approach is applied for the estimation of regression models describing the relationships between geometry parameters of the obtained welds (for SSt and St45) and the heat-affected zone (HAZ) for St45, as well as of the thermal efficiency η_T (for SSt) and the process parameters: electron beam power (P), welding velocity (v) and the focusing parameter ($dZ=Z_S-Z_O$), presenting the distance between the sample surface and the focus of the beam are estimated. The influence of the two distances: the distance from the main surface of the magnetic lens to the beam focus (z_o) and the distance to the sample surface (z_p) on the weld geometry are considered separately for stainless steel.

The obtained models are presented in Table 12 for coded in the region $[-1 \div 1]$ process parameter values. The relation between the coded (x_i) and the natural values (z_i) is given by:

$$x_i = (2z_i - z_{i,\max} - z_{i,\min}) / (z_{i,\max} - z_{i,\min}), \quad (5)$$

where $z_{i,\min}/z_{i,\max}$ are the corresponding values of the minimum and the maximum of the process parameters during the experiment (Table 9).

Some examples of the results are given in Figures. 43 and Figure 44. The figures present contour plots with lines of weld depths H and mean widths B depending of two of variables: power P [kW], velocity v [cm/min]. The focus position is at the surface of the sample ($z_o = z_p = 226$ mm). From Figure 43, where contour plot $H(P,v)$ is given one can see how, with the increase of beam power P together with the decrease of welding speed v , arise in weld depth H occurs. A decrease of the sensitivity of H to the P can be seen at higher values of P and lower v and also a decrease of the sensitivity of H to v at lower values of P and higher v . In Figure 44 the function $B(P,v)$ is presented also as contour plot. An unexpected optimal region of P and v for obtaining narrow welds exists. From the function $H(v,z_o)$ in Figure 45 it can be seen that, at higher welding velocities v the focus position is a decisive factor for increasing the weld depth H at constant beam power. From Figure 46 respectively weld width B is more

stable at high welding speed v . At the lower v the variations of the focusing position and v are more sensitive.

Table 12. Regression models

Case	Param.	Regression equation
STAINLESS STEEL (3 process parameters)	H	$22.8335+4.1065x_1-6.8632x_2-8.4127x_3-2.5658x_1x_2-2.0462x_1^2+$ $+3.7934x_2^2-6.88x_3^2+5.371x_1^2x_3+6.152x_1x_3^2$
	B	$1.7106+0.2986x_1-0.63263x_2+1.2335x_3-0.20335x_2x_3+0.4055x_2^2+$ $+1.125x_3^2-0.608x_1^2x_3+0.2983x_1x_2^2-0.9285x_1x_3^2$
	S	$34.61+12.358x_1-31.8850x_2-15.442x_1x_2+3.575x_1x_3-5.617x_2x_3+$ $+24.703x_2^2+13.383x_1x_2^2+6.442x_2^2x_3-4.196x_1x_2x_3$
STAINLESS STEEL (4 process parameters)	H	$20.82+5.975x_1-7.098x_2+3.742x_4-10.117x_5-1.202x_1^2+3.733x_2^2-$ $1.155x_4^2-14.534x_5^2-2.963x_1x_2-1.693x_1x_5+11.511x_4x_5$
	B	$2.166+0.195x_1-0.609x_2-0.785x_4+1.624x_5+0.427x_2^2+1.762x_5^2+$ $0.181x_1x_4-1.638x_4x_5$
STEEL 45	H _W	$14.0531+1.4160x_1-4.3478x_2-0.8375x_3-9.6644x_1x_3+3.2559x_1^2-$ $4.1089x_2^2$
	B _W	$4.89556+0.80192x_1+1.06983x_2+1.10761x_3+0.52610x_1x_2+2.02079x$ $_1x_3^2+$ $+0.83635x_2^2-0.48380x_2x_3^2$
	S _W	$29.1743+18.1348x_1+6.4578x_3+11.0342x_1x_2+3.3481x_2x_3-$ $1.4387x_2x_3^2$
	H _{HAZ}	$12.729+6.3641x_1+3.1611x_1x_2-4.4447x_3^2+3.219x_1^2x_2-2.1136x_2x_3^2-$ $-1.7884x_1x_2x_3$
	B _{HAZ}	$5.7224+1.5350x_1-1.2153x_2+1.6868x_1x_2+1.0011x_1x_3+1.1931x_3^2+$ $+1.2211x_1^2x_2-1.3502x_1x_2x_3$
	S _{HAZ}	$38.466+28.501x_1-13.371x_2+25.885x_1x_2+38.33x_1^2x_2$

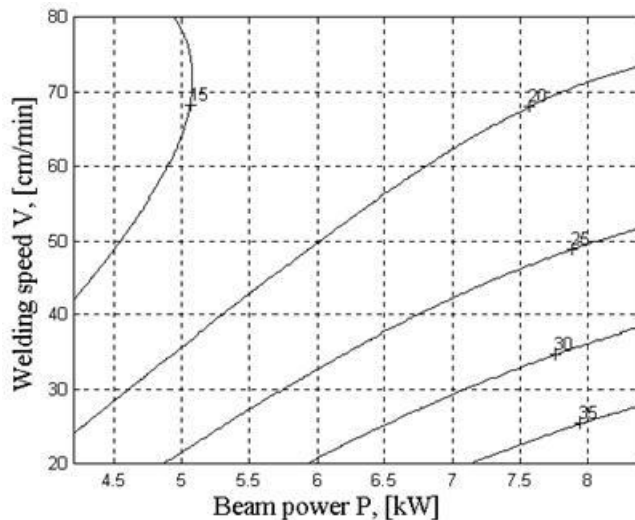


Figure 43. Weld depth $H(P,v)$ for $z_0=226$ mm, $z_p=226$ mm (SSt)

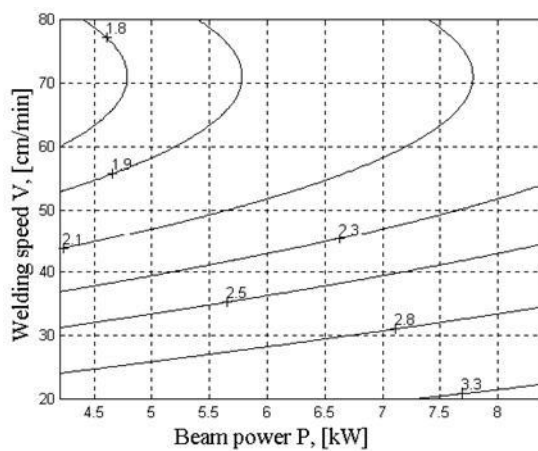


Figure 44. Weld width $B(P, v)$ for $z_0=226$ mm, $z_p=226$ mm (SSt)

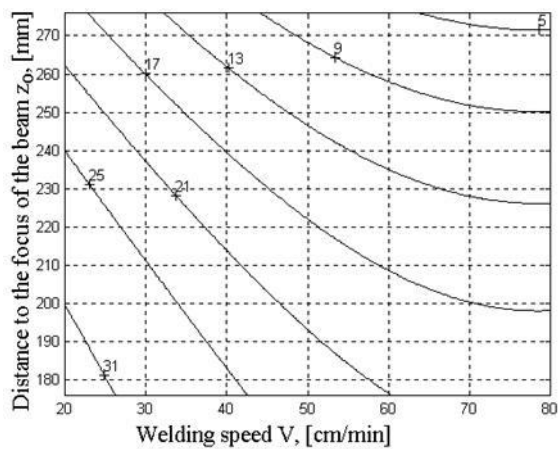


Figure 45. Weld depth $H(v, z_0)$ for $P=6.3$ kW, $z_p=126$ mm (SSt)

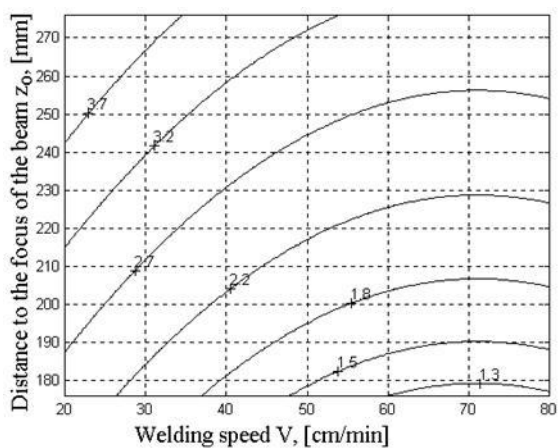


Figure 46. Weld width $B(v, z_0)$ for $P=6.3$ kW, $z_p=126$ mm (SSt)

Contour plots, presenting the dependence of the weld and HAZ depth H_{HAZ} and width at the top B_{HAZ} from the process parameters at EBW of St45, are presented on Figure 47 and Figure 48. They show that the HAZ is narrower and deeper for focus positions some millimeters below the sample surface (at $v = 1.5$ cm/s) and that the deepest and narrowest fusion zones are obtained for smaller welding velocities ($v=0.7$ cm/s) with a focus position deeply below the sample surface at a chosen beam power.

On Figure 49 the contour plots of the weld depth H and width at EBW of SSt. It can be seen that at chosen welding velocity ($v=50$ cm/min) and beam power the focus position toward the sample surface can be used as a tuning parameter for obtaining deep and narrow welds: the focus position must be moved from about 15 mm below the sample surface deeper up to 70 mm with the increase of the beam power.

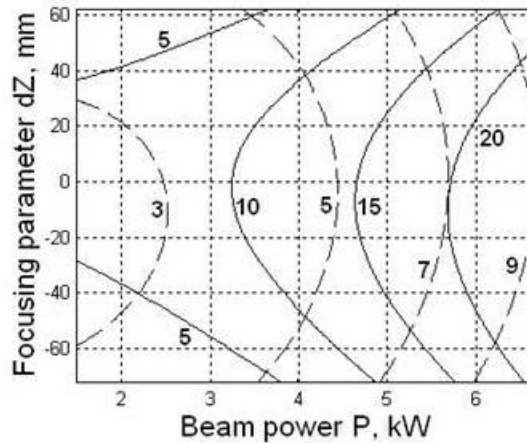


Figure 47. Contour plots of the HAZ depth $H_{HAZ}(P,dZ)$ (solid lines) and width at the top $B_{HAZ}(P,dZ)$ (dashed) for St45, $v = 1.5$ cm/s

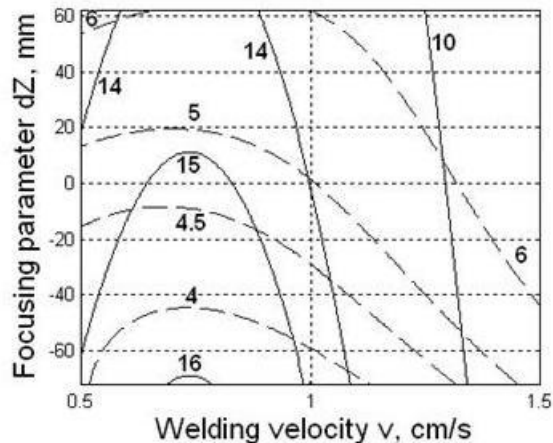


Figure 48. Contour plots of the weld depth $H(v,dZ)$ (solid lines) and width at the top $B(v,dZ)$ (dashed) for St45, $P = 4.975$ kW

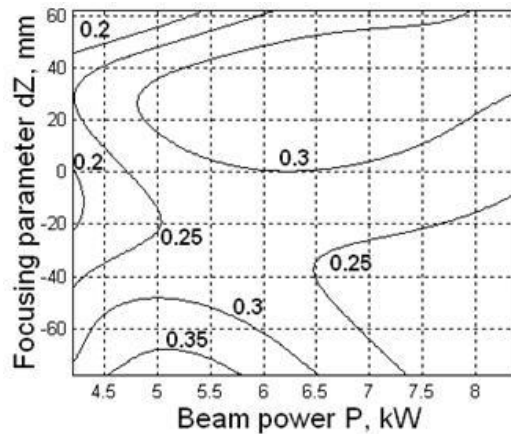


Figure 50. Contour plot of the thermal efficiency $\eta_T(P, dz)$ for SSt and $v = 50$ cm/min

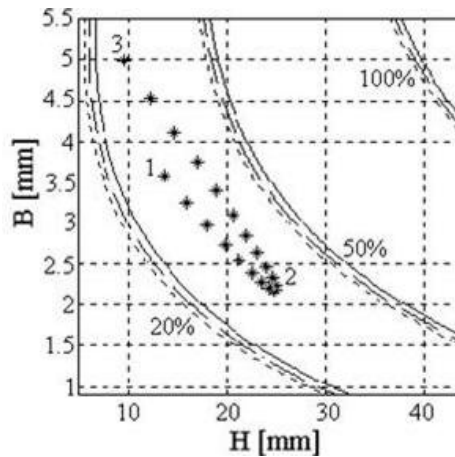


Figure 51. Contour lines for η_T of 20, 50 and 100% levels. Points * present the simulated H and B values for 21 dz in the region (-78÷68mm) at $P=4.2$ kW, $v=20$ cm/min. Point 2 is at $dz=-8$ mm and $\eta_T=0.34$

In Figure 52 and Figure 53 the relationships between the weld depth H and the beam power P depending on the changes of the welding speed and the focusing parameter dz are presented. It can be observed, that the increase of the welding velocity v leads to the decrease of the weld depth at keeping all the other conditions equal, and also that the increase of P leads to a considerable increase of H only at lower values of speed V. For positions of the focus above the sample surface at $dz=62$ mm there is a smooth increase of H with the increase of beam power, while at positions below the surface of welded samples this increase is more intensive up to certain level. At $dz=-8$ mm one can observe stable comparatively depths. The result depends weakly on P in direction of rise of H. The experimental results (signed with "*") and the limiting lines $P/H=1.33$ kW/cm and $P/H=10$ kW/cm (Figure 35), determined through the experimental data distribution. At small power (up to 5-6 kW) if one aim maximum H, the focus must be below the sample surface (see $dz=-8$ mm), while at 6-9 kW it is desirable to increase the depth of focus position. At upper studied powers (8-9 kW) deep welds can be obtained also at focus above the welded surface.

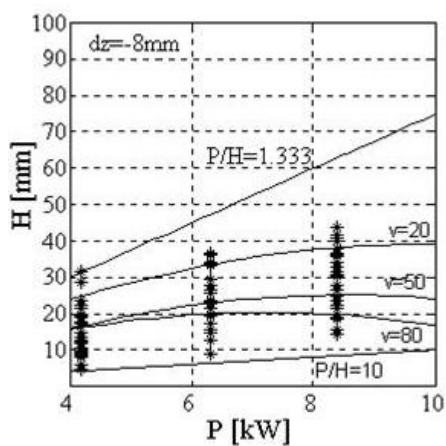


Figure 52. Weld depth $H(P)$ for $dz = -8$ mm at three level of welding speed: 20, 50, 80 cm/min

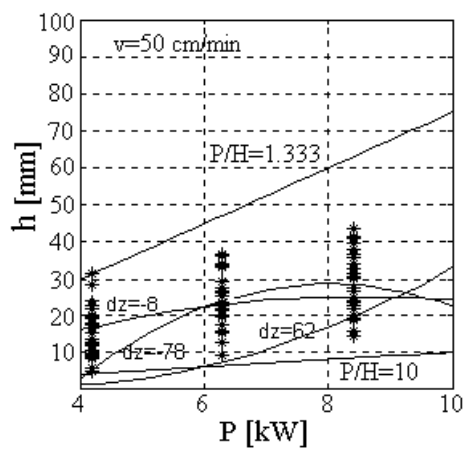


Figure 53. Weld depth $H(P)$ for $V = 50$ cm/min at three level of the focusing parameter dz : -78, -8, 62 mm

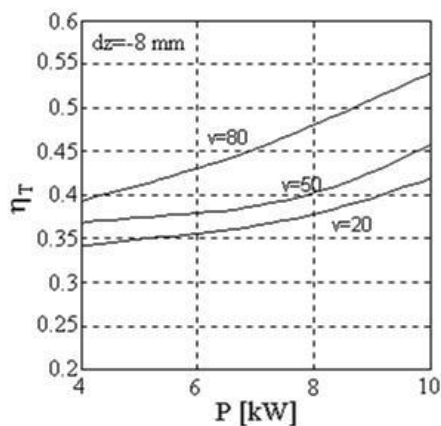


Figure 54. Thermal efficiency η_T versus beam power P at various speeds V [cm/min]

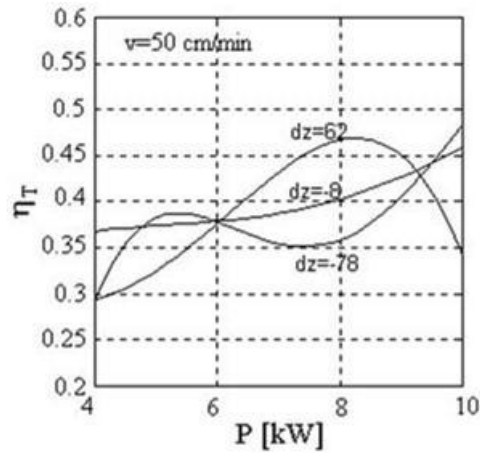


Figure 55. Thermal efficiency $\eta_t(P)$ for $V=50$ cm/min at various dz [mm]

In the Figure 54 and Figure 55 are shown the corresponding relationships of η_t for the generated cases. It is seen that the change of η_t has certain parity with the way H changes at the considered conditions. Maximum of η_t is reached at high speed and powers (instead of small velocities and powers where have a maximal H). At power 6 kW and 9 kW the value of η_t is not influenced by the focus position.

In the Figure 56 and 57 are shown the relationships between the width B and the welding speed v for different levels of the beam power and the distance dz . It can be seen that an elevation of P leads to an increase of B and the value of product vB (Figure 57). The position of the focus above the surface of the welded material increases B and the product vB (Figure 57). The limiting lines are at $vB=0.75$ cm/s and $vB=0.0533$ cm/s as they are obtained by the experimental data (see comments for Figure 35). On the figures are shown the experimental data too.

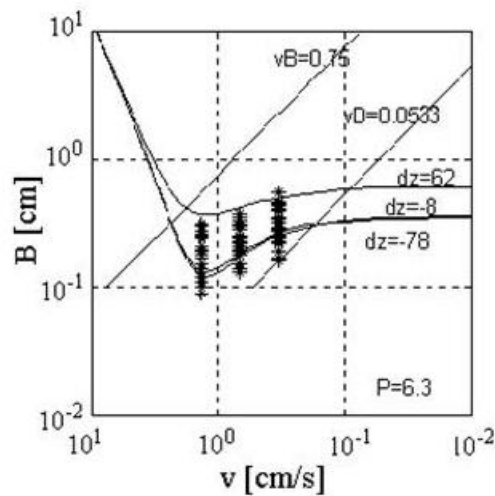


Figure 56. Weld width $B(V)$ for $P=6.3$ kW

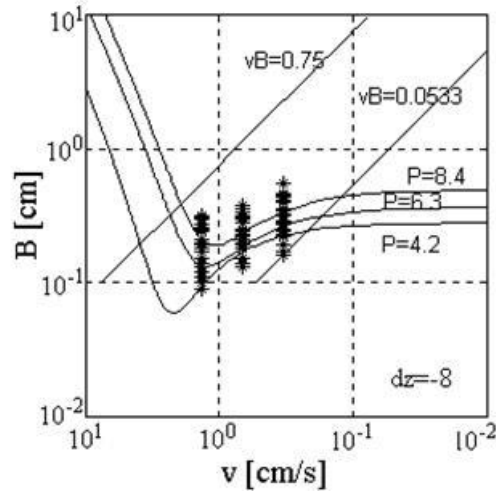


Figure 57. Weld width $B(V)$ for various P at $dz=-8$ mm

OTHER MODELLING ASPECTS

In order to examine the influence of the process parameters: beam current I_b , welding speed V and the distance between the main surface of the magnetic lens of the electron gun and the sample Z_s on the HAZ geometry parameters a statistical approach is applied. The data, obtained from nine available metallographic etch cross-sections of welds (both molten zone and HAZ) at EBW of St45, are presented in Table 13.

In the case of EB welding of semi-infinite sample with an electron beam, characterized with mean power density on the work piece surface less then the critical power density of 10^5 - 10^6 W/cm², a shallow or near to semi-spherical fusion zone is obtained due to the sample surface heating by beam near to point heat source. If the mean power density is higher, then a deep penetrating beam through a key-hole, generated in the molten pool [42, 45-49], as well as a quasi-steady state linear heat movable source can be assumed.

The form of the HAZ at EBW of St45 can be used also as a measure for a rough estimation of the transition from point to linear heat source. As a limiting value of the ratio of the depth to the width at the top H_{HAZ}/B_{THAZ} is accepted the value of 1.2. The beam spots evaluated correspondingly with the mentioned region of critical power density and the data of Figure 58 are of diameters $d \approx 1.43$ mm - 0.44 mm.

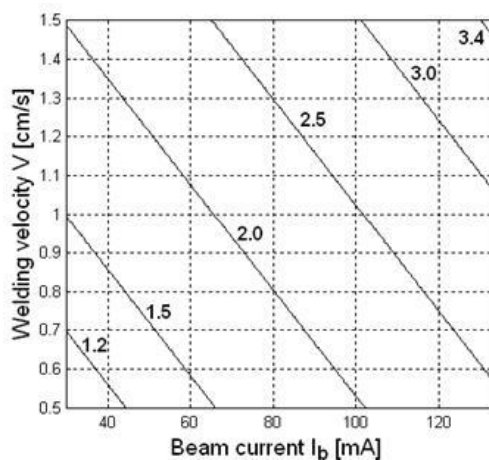
A statistical model for the ratio H/B (shown in [32, 34] as main characteristics of the EBW), measured as a function of the process parameters is estimated (see Figure 4):

$$H_{HAZ}/B_{THAZ} = 2.22 + 0.713x_1 + 0.507x_2 - 1.09x_6^2 - 0.650 x_2x_6^2, \quad (1)$$

where: x_1 , x_2 and x_6 are correspondingly the coded in the region $[-1 \div 1]$ values of the process parameters: beam current I_b , welding velocity V and the distance between the main surface of the magnetic lens of the electron gun and the sample Z_s (see Table 14), using the formula (5).

Table 13. Weld and HAZ geometry parameters and the intensity of the beam: (P/H)_L – calculated by θ_m linear heat source, (P/H)_E – heat source evaluated experimentally

N ^o Exper.	V	B _M	Y	R*	θ_m	(P/H) _L	(P/H) _{HAZ} /(P/H) _W	P _{exp}	H _{exp}	(P/H) _E
	mm/s	mm				W/cm		W	cm	W/cm
1 _W	15	5.5	5.0084	26.0430	0.1503	7014	1.4256	5000	0.66	7575.8
1 _{HAZ}	15	8.7	5.4182	30.3218	0.1391	9999.3		5000	1.13	4439.1
2 _W	10	5.3	3.2175	11.2629	0.2311	4561.7	1.6184	5000	1.33	3759.4
2 _{HAZ}	10	9.5	3.9750	16.7380	0.1884	7382.7		5000	1.40	3580.1
3 _W	10	4.3	2.6104	7.6891	0.2819	3739.6	1.8880	5000	1.4	3571.4
3 _{HAZ}	10	9.1	3.7966	15.3467	0.1970	7060.4		5000	1.43	3504.7
4 _W	5	5.0	1.5177	3.0417	0.4621	2281.3	1.9579	5000	1.36	3676.5
4 _{HAZ}	5	11.3	2.3474	6.3633	0.3114	4466.6		5000	1.40	3580.1
5 _W	5	4.3	1.3052	2.3945	0.5262	2003.4	1.6382	5000	1.37	3649.6
5 _{HAZ}	5	8.0	1.6741	3.5698	0.4238	3282		5000	1.49	3363.6
6 _W	5	3.4	1.0411	1.6998	0.6345	1661.5	1.4503	5000	1.57	3184.7
6 _{HAZ}	5	5.6	1.1688	2.0209	0.5772	2409.7		5000	1.71	2921.4
7 _W	5	2.7	0.8196	1.2064	0.7653	1377.5	1.9433	3300	1.60	2062.5
7 _{HAZ}	5	6.4	1.3248	2.4509	0.5196	2676.9		3300	1.62	2035
8 _W	5	5.5	1.6695	3.5536	0.4248	2481.6	1.8652	3300	0.77	4285.7
8 _{HAZ}	5	11.7	2.4386	6.8079	0.3005	4628.6		3300	0.81	4070
9 _W	15	4.7	4.2799	19.2629	0.1753	6013.7	2.1656	6650	2.03	3275.9
9 _{HAZ}	15	11.3	7.0856	51.1841	0.1068	13023		6650	2.07	3209

Figure 58. Contour plot of the ratio HHAZ/BTHAZ depending on the process parameters I_b and V , for the optimal value of $Z_s = 295$ mm**Table 14. Process parameters**

Parameter z_i	Dimension	Coded x_i	Min	Max
I_b	mA	x_1	30	133
v	cm/s	x_2	0.5	1.5
z_s	mm	x_6	228	362

The maximum value of H_{HAZ}/B_{THAZ} is 3.44, obtained for $I_b = 133$ mA, welding velocity $V=1.5$ cm/s and $Z_s = 295$ mm, which corresponds to a position of the beam focus 5 mm below the sample surface, where the most deep welds are expected.

Using the limiting criterion value $H_{HAZ}/B_{THAZ} = 1.2$, the experiments are divided into two sub-groups. First is (i) the experiments with $H_{HAZ}/B_{THAZ} < 1.2$, where a semi-spherical weld shape and respectively a movable point heat source can be assumed; the second group is (ii) the experiments with $H_{HAZ}/B_{THAZ} > 1.2$, where a deep penetration of beam and movable linear heat source can be assumed. Two series of regression models for the two-subgroups for the HAZ cross-section surface S_{HAZ} , depth H_{HAZ} , width at the top B_{THAZ} and mean width B_{MHAZ} can be estimated (Table 15).

From the obtained models it can be concluded that for the deep welds the distance to the sample surface in the investigated region does not affect significantly the cross-section surface of the HAZ, but it is a significant factor for the its shape.

The obtained models could be used as help of operator choice of regime parameters to obtain a desirable weld (namely HAZ of the seam) as well as for automatic control of EBW machine at welding Steel 45 pieces.

In Figure 59 as an illustrative example are given contour plots of the mean width B_M (dashed lines) and depth (solid lines) of HAZ, for $H_{HAZ}/B_{THAZ} > 1.2$ (at beam focusing 300 mm, beam surface $Z_s=295$ mm and accelerating voltage 50 kV). On the horizontal axes are given beam current values and on vertical axes of these plots are given the velocities values. It can be seen, that deeper and narrower HAZ could be obtained at $I_b=133$ mA and welding velocities 1.5 cm/s. The colored area is roughly the area where the equations for $H_{HAZ}/B_{THAZ} < 1.2$ from Table 14 should be used.

Metallographic etches of materials with two isotherms on the weld cross-section (HAZ and molten zone) allow the estimation of the role of the deviations from the ideal model – heating with moving linear heat source of a semi-infinite hard body. These deviations are due to the presence in the molten bath of a key-hole, and then the mass transfer is realized through the liquid pool by the moving heated liquid metal, the phase transitions presence in the heat transfer process. An interesting question arises: using the two zones contours is it possible to investigate the linear moving heat source intensity distribution that is acts during the deep penetration of the beam?

Table 15. Regression models for the HAZ geometry parameters

Parameter	$H_{HAZ}/B_{THAZ} < 1.2$	$H_{HAZ}/B_{THAZ} > 1.2$
S_{HAZ}	$32.9 + 50.4x_1 - 22.4x_2 - 3.57x_3 - 21.2x_1x_2 + 1.24x_2x_3 + 33.0x_1^2 + 15.6x_2^2 + 24.0x_1x_2^2 + 2.25x_3x_2^2$	$35.1 + 56.1x_1 - 12.5x_2 + 36.3x_1x_2 + 42.1x_1^2 - 39.2x_1x_2^2$
H_{HAZ}	$6.47 + 5.12x_1 - 3.26x_2 - 1.22x_3 - 3.44x_1x_2 - 0.782x_1x_3 + 2.40x_1^2 + 1.58x_2^2 + 3.14x_1x_2^2 + 0.581x_3x_2^2$	$12.7 + 9.82x_1 + 9.80x_1x_2 + 3.69x_1^2 - 7.83x_1x_2^2 - 6.13x_1x_3^2 - 4.02x_2x_3^2$
B_{THAZ}	$6.41 + 2.14x_1 - 2.71x_2 - 0.518x_3 - 2.45x_1x_2 + 1.08x_1^2 + 1.95x_2^2 + 3.72x_1x_2^2 + 0.507x_3x_2^2$	$5.85 + 1.42x_1 - 0.833x_2 + 1.43x_1x_2 + 3.11x_1x_3^2$
B_{MHAZ}	$9.18 + 4.84x_1 - 2.68x_2 - 2.37x_1x_2 + 3.62x_1^2 + 2.99x_2^2 + 5.59x_1x_2^2$	$5.73 + 2.55x_1 + 1.90x_3 + 2.93x_1^2 + 2.64x_1x_3^2 - 2.07x_2x_1^2 - 1.62x_3x_2^2$

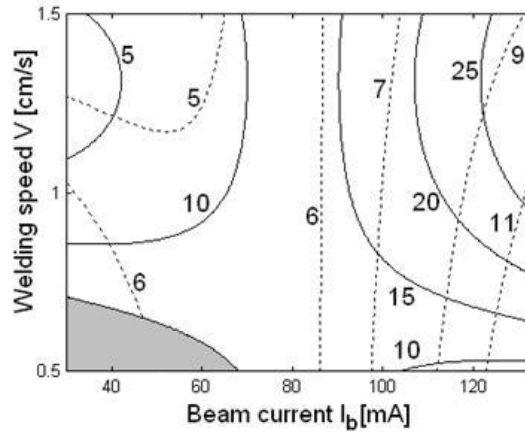


Figure 59. Contour plot for HHAZ (solid lines) and BMHAZ for HAZ, $\text{HHAZ/BTHAZ} > 1.2$, $Z_s=295$ mm

In order to answer this question the calculation of the temperature field at heating the samples with a moving linear heat source, which is the base of the EBW thermal model, are considered.

Form the metallographic etches of weld cross-sections regression equations for the molten zone parameters (cross-section surface S_w , depth H_w , width at the top B_{TW} and the mean width B_{MW}) of the welds available are estimated and given in Table 16. The region of the beam current considered is [66-133 mA]. The process parameters x_1 , x_2 and x_3 have coded values. Contour plots for H_w (solid lines) and B_{MW} (dashed lines) of the weld zone ($Z_s=228$ mm), for a position of the focus 72 mm below the sample surface are presented on Figure 60.

If the material physical parameters: thermal conductivity λ , thermal diffusivity a ($a=\lambda/C_p \cdot \rho$, where C_p is the specific heat and ρ is the sample density) are known, the solution of thermal balance equation at heating a sheet of thickness H from a linear moving thermal source (of a constant distributed intensity P/H) moving with speed V , assuming no phase changes in the sample during heat transfer can be found from eq. (1).

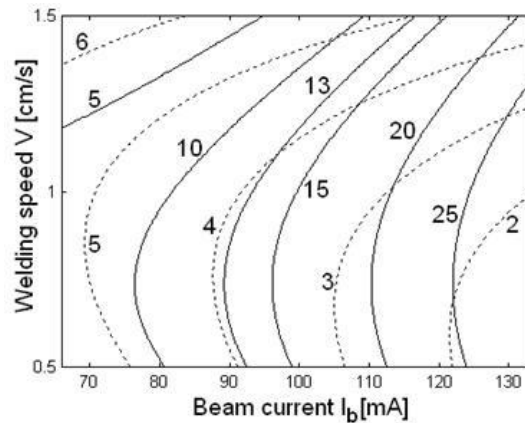


Figure 60. Contour plots for H_w (solid lines) and B_{MW} of the weld zone ($Z_s=228$ mm), position of the focus is 72 mm below the sample surface

Table 16. Regression models for the molten zone weld parameters

Param.	Regression models
S_W	$30.1 + 18.3x_1 - 4.73x_2 + 3.30x_6 + 11.6x_1x_2 - 6.13x_2^2 + 1.96x_6^2$
H_W	$14.1 + 1.42x_1 - 4.35x_2 - 0.837x_6 - 9.66x_1x_6 + 3.26x_1^2 - 4.11x_2^2$
B_{TW}	$4.88 + 0.892x_1 + 0.191x_2 + 0.709x_6 + 0.590x_1x_2 + 2.06x_1x_6 - 0.407x_2x_6 + 0.448x_6^2$
B_{MW}	$4.41 + 2.16x_1 + 1.03x_2 + 0.779x_6 + 0.579x_1x_2 + 3.76x_1x_6 + 0.859x_2^2$

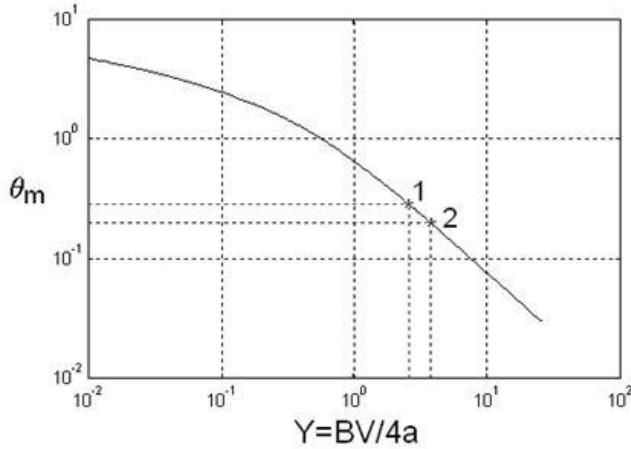


Figure 61. The dependence of the maximum dimensionless temperature on the dimensionless distances $Y=BV/4a$ as an example for experiment No3: 1-for the weld; 2-for the heat affected zone

The dependence of the dimensionless temperature θ_m as a function of $Y=yV/2a$ is presented in Figure 61 (eq.(2)). Using values of known θ_m for given P and H , the curve shown in the Figure 61 gives the possibility to obtain the weld width for a concrete material. Conversely, at using the chosen width value one can obtain the weld depth value.

In the Table 13 are evaluated intensities of uniformly distributed on the weld depth intensities of the linear heat source P/H , evaluated by two ways. Using the experimental depths of welds and HAZ and the beam power are estimated $(P/H)_E$. From data in Table 3 it is possible to calculate the dimensionless distance from the beam axis $Y=yV/2a = BV/4a$ and the dimensionless maximum temperature θ_m . The theoretical value of $(P/H)_L$ (Table 13) is evaluated on base of the equation:

$$\left(\frac{P}{H}\right)_L = \frac{2\pi\lambda(T_m - T_o)}{\theta_m}.$$

At analysis of the obtained data for $(P/H)_E$ and for $(P/H)_L$ in Table 13 could be observed discrepancies. The difference between the $(P/H)_E$ are within 10% and are due to errors in estimating H_W and H_{HAZ} . The calculated using the B_W (or B_{HAZ}) and θ_m the values of the $(P/H)_L$ are more variable and inexact. This is due to the deviation of the heat source from a linear one and due to the uncertainty of the measured B_m values. The more deep welds become the less this deviation will be. It can be noted also, that the estimated $(P/H)_E$ for the weld, systematically is higher with few % than $(P/H)_E$ for HAZ, due to the presence of a key

hole. But still the accuracy of estimation is considerably higher. It was also found out that the ratio of the obtained intensity of the linear heat source $(P/H)_L$ for the weld and $(P/H)_L$ for HAZ depends on the position of the focus towards the sample surface, which points to the importance of this parameter for the real heat source intensity distribution. Uncertainty of measurements of B did not give a possibility to create acceptable methodology to prognosticate the weld width B_{MW} from the data of B_{MHAZ} through the thermal model of EBW, described early. It is not easy by that data to study precisely the intensity distribution of heat linear source that act in the weld key-hole. The evaluation of HAZ width by different methods (hardness distribution, metallographic etching, corrosion experiments etc.) have to be compared and conclusions for their exactness are still needed.

In order to approximate the form of the cross-section of the welds and the heat-affected zones an approximation is made:

$$\hat{H}(x) = -\frac{S}{(B/k)\sqrt{2\pi}} \exp\left(-\frac{x^2}{2(B/k)^2}\right), \quad (6)$$

where S is the weld/HAZ cross-section surface, B is the weld/HAZ width and the coefficient $k=2.5$ for SSt and $k=3$, when the width at the top of the weld or the HAZ (St45) is used. The coefficients are estimated by ordinary least squares method.

Figure 62 presents superimposed the experimentally observed form of the weld cross-section and the approximation made by eq. (6) for SSt weld for $P=4.2$ kW, $v=80$ cm/min and $dZ=-60$ mm.

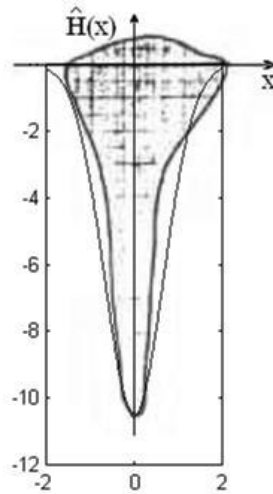


Figure 62. Approximation of the form of the cross-section of the weld for stainless steel

NEURAL NETWORK MODELING OF EBW PROCESS

One of the most promising fields of the Artificial Intelligence is related to the Neural Networks [63] that has the ability to learn and approximate any functional relationship. The

NN integration in intelligent control system [64] is based on such characteristics of connectionist systems as: availability of learning, generalization, classification; stability in relation to partial faults in the network and the noise; improving performance with increased experience; associative memory. The advantages of NN are demonstrated when the mathematical description of the plant is very complex or the computational task is not completely defined. In relation to control systems NN are attractive tools for solving problems in which classical analytic methods are difficult to be applied. It is appropriate to use neural network for process modeling and control, pattern recognition, fault diagnosis. Moreover, despite the possibility of equally comparable solutions to a given problem, several additional aspects of a neural network solution are appealing, including parallel implementations that allow fast processing; less hardware which allows faster response time, lower cost, and quicker design cycles; and on-line adaptation that allows the networks to change constantly according to the needs of the environment.

A number of process engineering problems have been studied and solved using the neural networks approach that exploits symbolic processing and knowledge representation [65÷69]. The majority of the neural networks utilized in the applications are the multilayered feed-forward networks. First and still widely used method of training the neural networks is the so-called back propagation method (BPM) [70]. It requires a preliminary generated (usually experimentally obtained) set of training data containing sets of input-output data for the neural network.

An example of a model structure in the form of a Neural Network is shown in Figure 63.

Further a procedure of creating neural network-based models and their application to the prediction of the electron beam welding (EBW) performance characteristics and to the parameter optimization are presented.

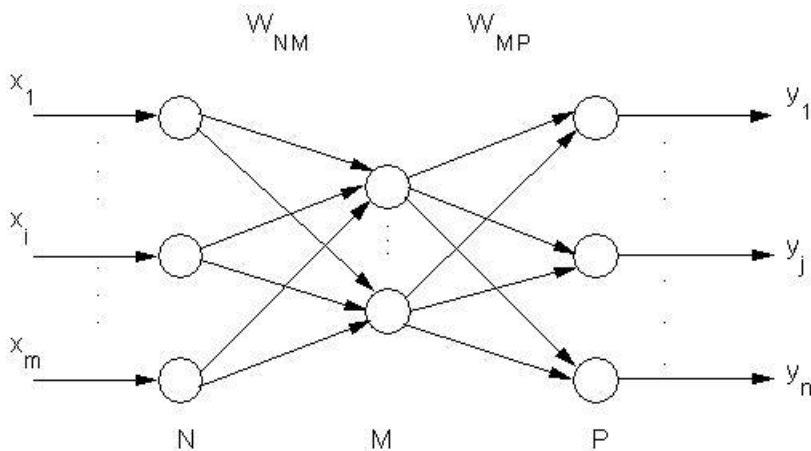


Figure 63. Neural network structure

The proposed methodology for developing NN-based models for EBW performance characteristics consists of the following general steps:

1. Construction of the neural network model structure.

2. Training of the created neural network by using the back propagation method [70] and experimentally obtained (and/or numerically simulated) set of training data to a satisfactory accuracy.
3. Recall of the trained neural network for prediction and parameter optimization.

The modelled EBW process parameters define the input-output structure of the neural network-based model used, i.e. the neural network should consist of 4 input neurons and 1 output neuron. NN models for each output (weld depth H and mean weld width B) are considered (illustrated in Figure 64).

The best results for Neural network models for the weld depth H and mean width B were obtained with 5 hidden units and different number of iterations for training (above 10000 iterations). For the purpose of validation the data were split into two parts: training datasets containing 73 observations and the testing datasets limited to 8 observations each (for H and for B). For each performance characteristic randomly were chosen 10 datasets (73 training and 8 test observations) and for each dataset the best network model was obtained and verified. For comparison of the models the absolute value of the error calculated as the difference between the predicted and the measured values of the weld geometry characteristics, as well as root mean squared error (*RMSE*) and the non-dimensional error index (*NDEI*) are used. The last two are calculated by:

$$RMSE = \sqrt{\frac{(\hat{y} - y)^2}{n}}; NDEI = \frac{RMSE}{\sigma},$$

where $\hat{y}$ and y is the predicted and the experimental values, n is the number of data and σ is the standard deviation of the data points. These error measures are defined on the basis of the training error (average loss over the training sample) and the generalization error (expected prediction error on an independent sample). Their values are minimized during the neural network training.

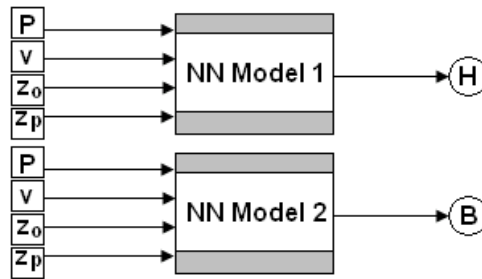


Figure 64. Neural networks input-output parameters for the weld depth H and mean width B

The experimental results (marked with points) and the predicted results (connected with the straight lines) using the estimated best model for the weld depth H using the training dataset (73 observations) are presented in Figure 65.

The absolute value of the errors, presented as the difference between the predicted and the measured values of the weld depths, are calculated and graphically presented in Figure 66,

connected with lines. Generally, the error values are situated in the region $(-2 \div 2 \text{ mm})$ with the exception of only 5 errors. The model precision is estimated quantitatively by RMSE and NDEI and the results are presented in Table 17.

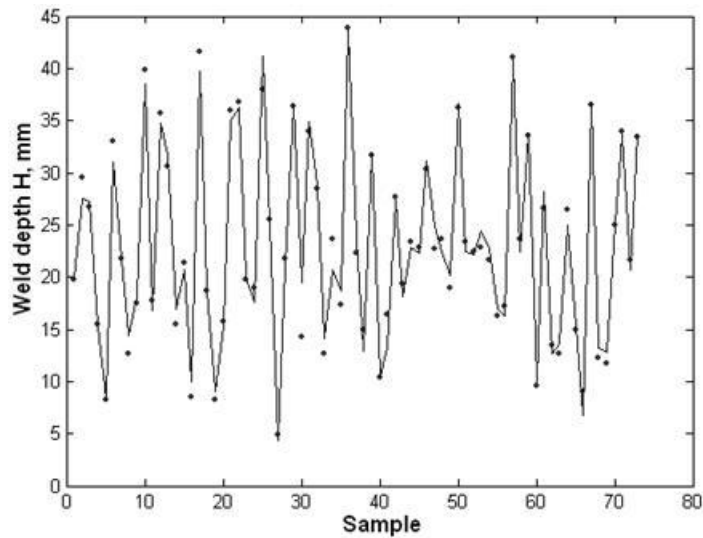


Figure 65. Predicted end experimental values for the weld depth H – training

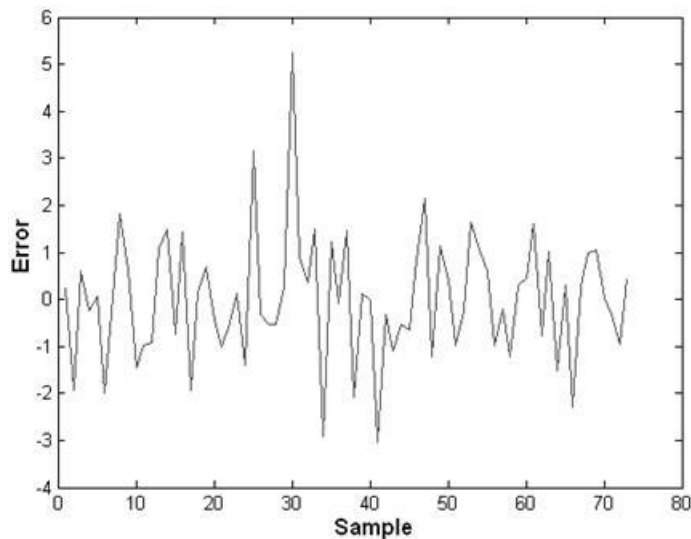


Figure 66. Absolute error values (the differences between the experimental and the predicted weld depths H) – training

In Figure 67 and Figure 68 are presented the results from the training of the best neural model for the weld mean width B.

A comparison between trained neural networks (Figure 69), describing the relationship of the thermal efficiency and different combination of factors: a) depth H and mean width B of the welds (2 factors); b) electron beam power P, welding velocity v, the distances between the

main surface of the magnetic lens of the gun to the beam focus z_0 and to the surface of the sample (4 factors) and c) all considered factors (6 factors). The results from the training and the cross-validation are presented in Table 17. It can be seen, that the trained neural network models with 4 factors give very good results. Visualization of the experimental and the predicted results for the thermal efficiency in this case are presented on Figure 70 and Figure 71.

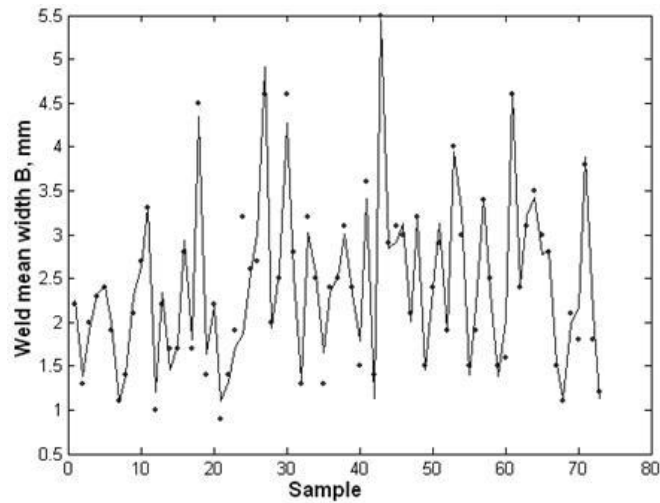


Figure 67. Predicted end experimental values for the weld mean width B – training

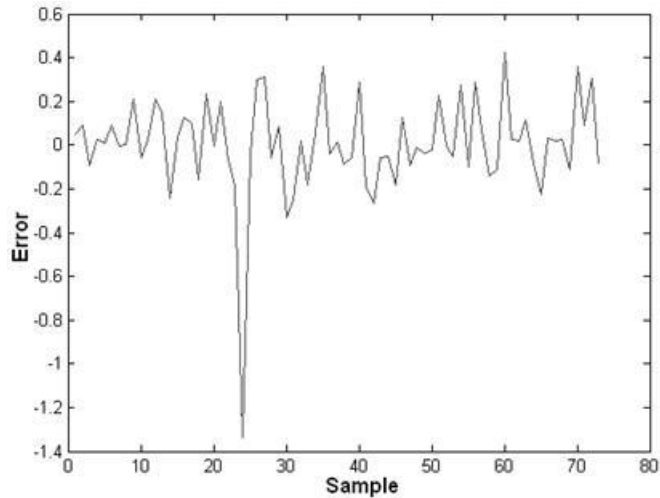


Figure 68. Absolute error values (the differences between the experimental and the predicted weld mean widths B) – training

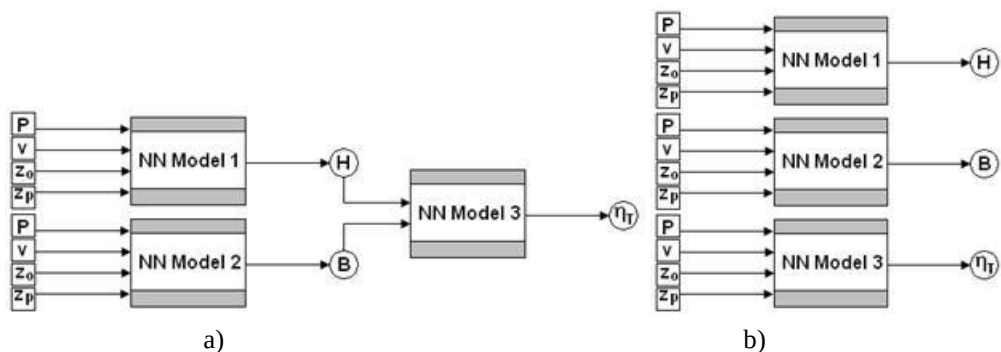
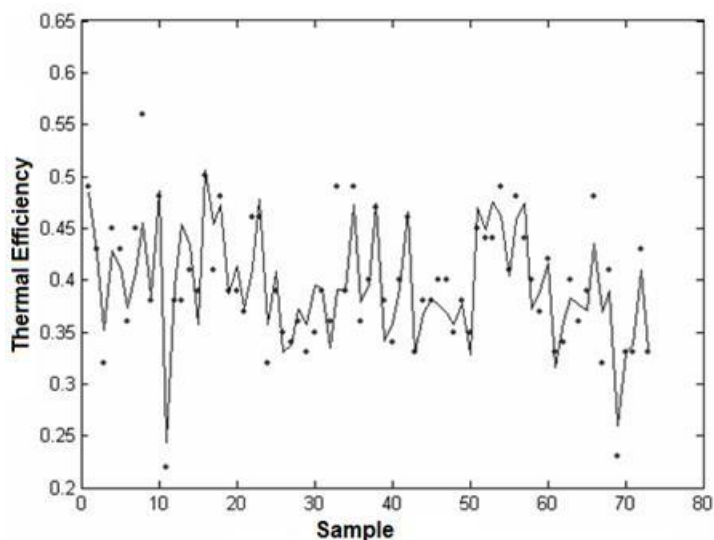


Figure 69. Neural networks input-output parameters for the thermal efficiency – inputs and outputs

Table 17. RMSE and NDEI error measures

Process Parameters		Training (73 experiments)	Testing (validation) (8 experiments)	Performance Characteristic
4	RMSE	1.33382	1.52107	H
	NDEI	0.141456	0.162708	
4	RMSE	0.226097	0.131611	B
	NDEI	0.231885	0.116459	
2	RMSE	0.0531979	0.0814766	η_T
	NDEI	0.908591	0.875771	
4	RMSE	0.0290363	0.0273294	η_T
	NDEI	0.47571	0.3782120	
6	RMSE	0.0253802	0.0222612	η_T
	NDEI	0.397551	0.557775	

Figure 70. Predicted end experimental values for the thermal efficiency η_T – training

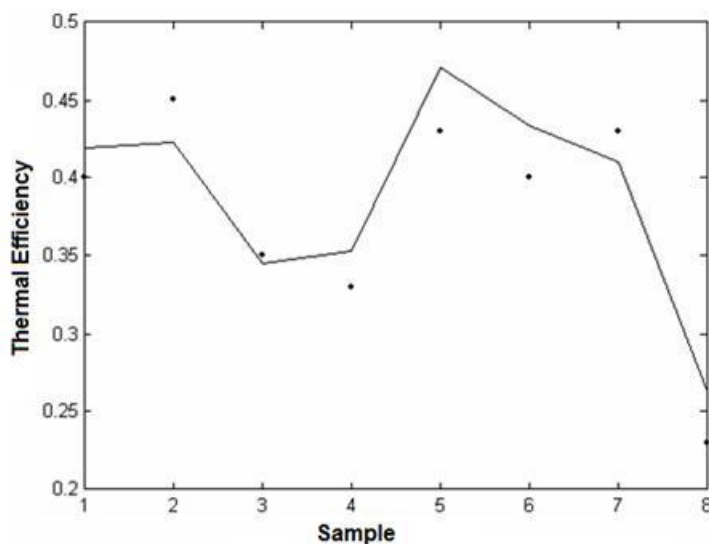


Figure 71. Absolute error values (the differences between the experimental and the predicted the thermal efficiency η_T) – training

QUALITY IMPROVEMENT IN PRODUCTION CONDITIONS

In production conditions (compared to laboratory installations) the prediction of the geometrical characteristics of EBW is an even more complex task due to the presence of errors, coming from the tolerances in the controlled EBW parameters, or, from other uncontrolled parameters [62]. The variations caused by these variables make it difficult to repeat weld geometry exactly under the same conditions. The quality improvement considered here is connected with finding regimes where the variation in the weld depth and width will be less sensitive to such variables [71].

In production conditions usually variations of the process conditions are usually observed. They result in increasing the variations of the performance characteristics of the produced welds. The robust engineering approach can be applied for the quality improvement related to the decrease of the variations of the obtained welds and its repeatability. The estimated regression models are used for the estimation of two new models for the performance of each quality characteristic in production conditions: a model of the mean and a model of the variance [62, 71]. These two models can be used for choosing process parameters, which satisfy both the characteristic being close to its target value and minimization of its variance.

A new method for estimation of regression coefficients takes into account both the correlation and the heteroscedasticity (the case when there are errors in the factors levels in the production stage resulting in variation of performance characteristics, which depends on the process parameters) of the performed experiments in order to improve the accuracy of the estimated regression models, as well as the models for the means and variances of the multiple responses, is proposed in [72]. This combined approach can be implemented for the sequential generation of industrial experimental designs.

The application of the proposed approach gives the possibility to use for the quality improvement using the robust engineering approach raw industrial experimental data, instead of the necessary very precise regression model estimations without errors in the factor levels, done usually in laboratory conditions.

The mean and the variance models for the two responses are estimated, applying the original new combined method. On Figure 72 contour plots of the weld depth H mean and variance at EBW of SSSt in production conditions are presented. For the estimation of the models the tolerance limits given in Table 9 are used.

Figure 72 and Figure 73 present the equipotential contour lines of the mean value (solid) and the variance (dotted) for both - the weld depth and the weld width depending on the beam power and the welding velocity at focusing parameter $dz=-40$ mm (Figure 72) $dz=-78$ mm (Figure 73).

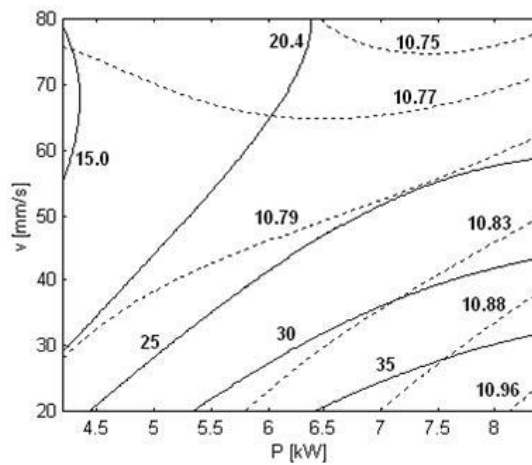


Figure 72. Contour plots of the mean $\tilde{y}_H(x)$ (solid lines) and the variance of the weld depth H (dotted lines) depending on P and v at $z_0=276$ mm and $z_p=236$ mm

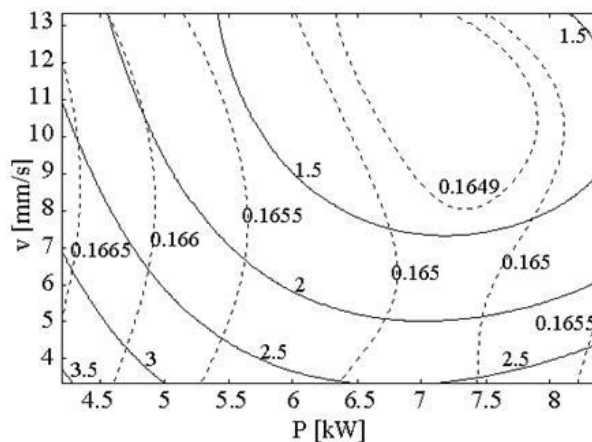


Figure 73. The estimated contour plots for the mean value $\tilde{y}_B$ (solid) and the variance $\hat{s}_B^2$ (dotted) of the weld width B for a focusing parameter $dz = -78$ mm

OPTIMIZATION

Using the multi-response surface methodology [57, 58], polynomial regression models or neural network models for the estimation of the behavior of the weld depth H and the mean weld width B (as well as the thermal efficiency or other performance characteristics) at EB welding with deep penetrating beam versus welding and material characteristics parameter optimization can be performed. A model is developed that includes the values of beam power and welding speed as well as the distances between the electron gun and both the focusing plane of the beam and the sample surface as process parameters. Computer procedures for the choice of operating conditions under some criteria for obtaining special parameters of the seam and for acquiring optimal weld parameters can be different, depending on the concrete requirements for the characteristics of the produced welds. As criteria for such optimization can be used desirability function for a property - values of the weld depth, the width or the thermal efficiency. In order to improve the quality of process (to decrease the deviation from the target value of the performance characteristics) in production conditions a model approach is applied. Two models: one describing the mean value (using the mentioned polynomial regression or other modeling method) and second calculating the variance for the weld depth and the weld width in the mass production are estimated. Utilizing these models quality improvement can be defined [62, 71] as an optimization problem of variance minimization while keeping the mean value of weld depth or/and width on the target values.

Additionally to the requirements for the geometry of the obtained welds and the process thermal efficiency, requirements for the defect-free welds are typical.

For the experimentally obtained weld cross-sections by EBW of stainless steel, the number of defects is counted. Several approaches (response surface methodology, discriminant analysis etc.) are applied for the prediction of the process parameter regions, where the probability for appearance of defects is smaller. The experimental welds are separated into two groups (classes): 1 – with defects and 2 – without defects. The type of the defects is not taken into account.

The analysis for concrete conditions shows that the most influential process parameters, which should be considered, in order to avoid the defect appearance, are electron beam power and the distance to the surface of the sample.

In the case of applying the regression analysis 94% of the observations are predicted correctly (95% - for the group 1 of observable defects in welds and 89.5% for the group 2). The regression model for the defects is estimated as follows:

$$\begin{aligned}
 D = & -0.177 - 0.341x_1 - 0.113x_2 + 0.562x_4 - 1.188x_5 + 0.495x_1^2 + 0.260x_2^2 + 0.314x_4^2 + \\
 & + 1.097x_1x_5 - 0.368x_2^2x_4 - 0.383x_1^2x_4 + 0.553x_1^2x_5 - 0.271x_1^2x_2x_5 + 0.379x_1x_4^2 - \\
 & - 1.867x_1x_4x_5 + 1.803x_1x_5^2 + 0.677x_2^2x_5 + 0.320x_1^2x_2x_4 - 0.586x_1^2x_4^2 + 2.037x_1^2x_4x_5 - \\
 & - 2.083x_1^2x_5^2 - 0.232x_1x_2^2x_4 - 0.742x_1x_4x_5^2 - 1.890x_2^2x_4x_5 + 1.886x_2^2x_5^2 - 0.310x_2x_4x_5^2.
 \end{aligned}$$

The value of $D=0.5$ is accepted as a conditional limit between the regions with ($D>0.5$) and without ($D<0.5$) defects.

The estimated regression models can be used for EBW process parameter optimization fulfilling the specific performance characteristic requirements for finding individual optimum and compromise solutions.

On Figure 74 is presented the result from maximization of the weld depth H . The maximum value obtained is $H=43.65$ mm at $P=8.4$ kW, $v=20$ cm/min, $z_o=176$ mm and $z_p=146.5$ mm (focus position at 29.5 mm below the sample surface). A requirement is added for lack of defects ($D<0.5$). The coloured zone contains all the regimes at which defects are not expected. Figure 75 shows the results from the parameter optimization for the thermal efficiency under the following constraints: $H>25$ mm, $B<3$ mm and no defects ($D<0.5$). The focus position in this case is on the sample surface ($z_o=226$ mm and $z_p=226$ mm). The maximum thermal efficiency is 0.43, obtained for maximum beam power and welding velocity of 26 cm/min.

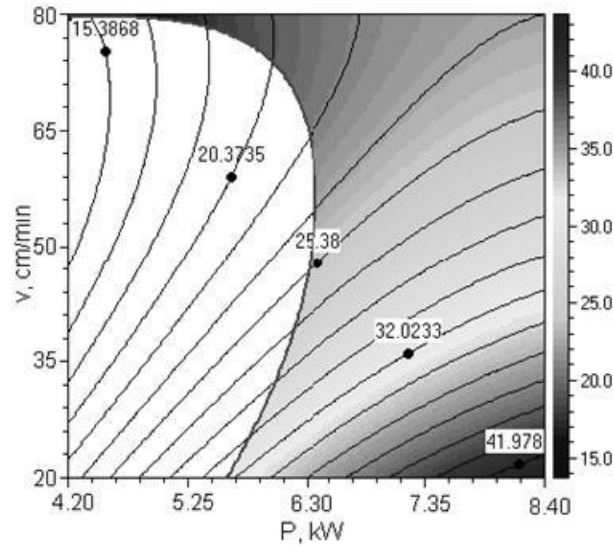


Figure 74. Contour plot $H(P,v)$, at $z_o=176$ mm and $z_p=146.5$ mm (SSt)

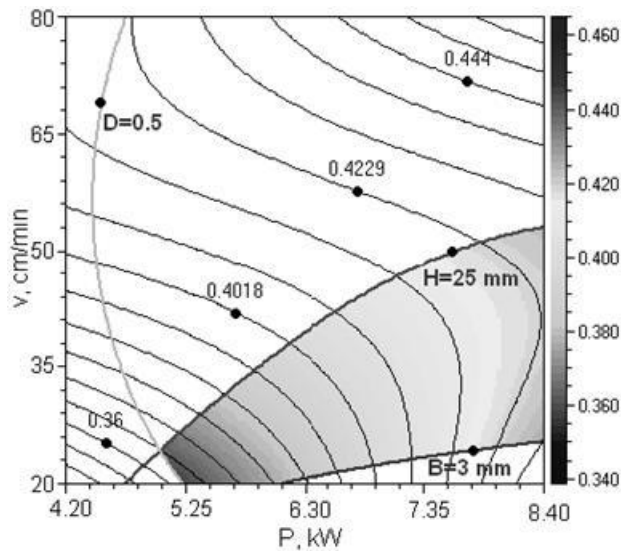


Figure 75. Contour plot $\eta_T(P,v)$, at $z_o=226$ mm and $z_p=226$ mm (SSt)

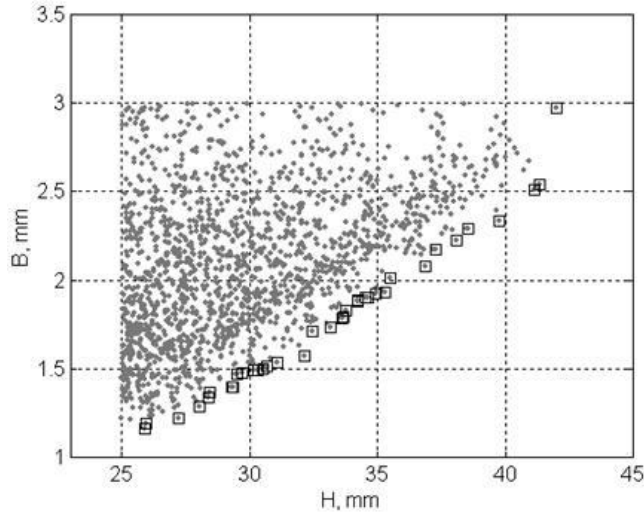


Figure 76. Pareto-optimal solutions ('□') and constraints: $H > 25$ mm and $B = [1 \div 3]$ mm] (SSt)

When optimum of more than one function at the same time is required, compromise solutions should be found, since the individual optima usually are reached at different regime conditions. Pareto-optimal solutions form a group of optimal solutions in the sense that moving away from a given Pareto-optimal point will worsen at least one of the considered performance characteristics. The choice among the Pareto optimal solutions should be made according other criteria. Figure 76 represents a set of points (calculated from 10000 randomly selected regimes within the experimental region), which fulfill the constraints: $H > 25$ mm and $B = [1 \div 3]$ mm] and Pareto-optimal solutions (signed with '□'), which maximize the depth H and minimize the width B within the acceptable region at the same time. In Table 18 are presented a few of these solutions such solutions (first three points). Each of these points is closer to one of the optimums: maximum H or minimum B .

Another approach of a compromise solution choice is the analytical technique for the optimization of a several functions, using the utility or the desirability of a property given by a certain performance characteristic function (in our case weld depth H and width B). One can specify certain desired values of the weld geometry characteristics d_i and they will be two-side constrained $y_{i*} \leq y_i(\mathbf{x}) \leq y_i^*$ (there y_{i*} and y_i^* are acceptable values of the lower and upper deviations from the desired values). Then the individual desirability for each function is evaluated by the function:

$$g_i = \begin{cases} [(y_i - y_{i*}) / (d_i - y_{i*})]^s, & \text{for } y_{i*} \leq y_i \leq d_i \\ [(y_i - y_i^*) / (d_i - y_i^*)]^t, & \text{for } d_i \leq y_i \leq y_i^* \\ 0, & \text{for } y_i \leq y_{i*} \text{ or } y_i \geq y_i^* \end{cases},$$

where the values of s and t are chosen within the domain $[0.1; 10]$ - the larger values of s and t are the desirability function is larger only for weld depths and widths that are closer to d_i . If all the values in the region $y_{i*} \leq y_i(\mathbf{x}) \leq y_i^*$ are almost equally acceptable, s and t are given

smaller values. A single function D is formed from all the individual desirability functions, which gives the overall assessment of the desirability of the combined responses, namely the geometric mean of the values of g_i . The overall desirability function for H and B is:

$$D = (g_H \cdot g_B)^{1/2}.$$

In Table 18 are presented three of the solutions (№4-6), having the highest overall desirability value G at desirable values $H=30$ mm and $B=2$ mm ($s=t=1$) and acceptable regions $H=[28-32$ mm] and $B=[1.5-2.5$ mm]. In Figure 77 is shown contour plot of overall desirability function D (and maximal desirability value G) for the optimal solution №4.

The quality improvement based on process parameter optimization is the cheapest way to utilize the available equipment and materials. The estimated models applying the statistical approach can be utilized for fulfilling that task. In Table 19 the optimal process parameters for obtaining maximum (minimum) of the performance characteristics at EBW of SSt are determined.

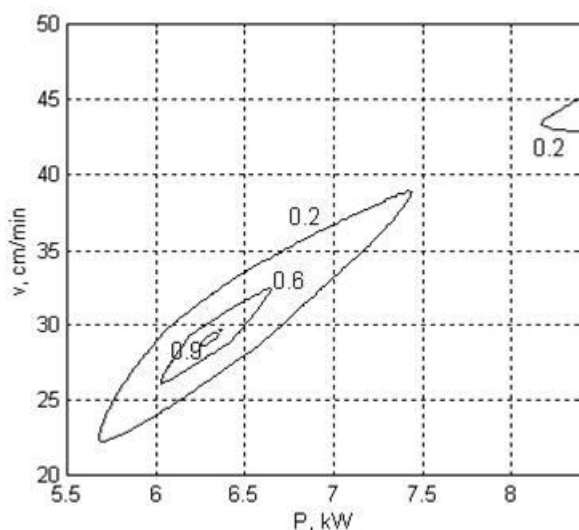


Figure 77. The overall desirability function ($z_o=226$ mm, $z_p=176$ mm) (SSt)

Table 18. Optimal solutions – Pareto-optimal and desirability function

	P, kW	v, cm/min	z_o , mm	z_p , mm	H	B	G
1	8.21	52.86	253.88	255.90	25.92	1.19	- (Pareto)
2	7.36	73.91	201.04	234.55	42.01	2.97	- (Pareto)
3	6.35	57.03	270.30	253.03	35.28	1.93	- (Pareto)
4	6.30	29.00	226.00	176.00	30.00	2.00	0.9912
5	8.40	35.00	216.00	126.00	29.98	2.00	0.9848
6	5.88	26.00	186.00	146.00	29.97	2.00	0.9767

Table 19. Optimal process parameters for maximum/minimum of the performance characteristics at EBW of SS

		P, kW	V, cm/min	dZ, mm	H, mm	B, mm	S, mm ²	η_T
Stainless steel	H _{max}	8.106	20.0	-78.0	40.5659	2.6014	107.8733	0.3371
	B _{min}	4.200	80.0	-15.0	16.3507	0.8649	16.9844	0.3658
	S _{max}	8.400	20.0	62.0	34.3465	4.3710	152.2110	0.4420
	η_{max}	5.775	31.1	24.9	21.4486	3.0877	59.7869	0.5287

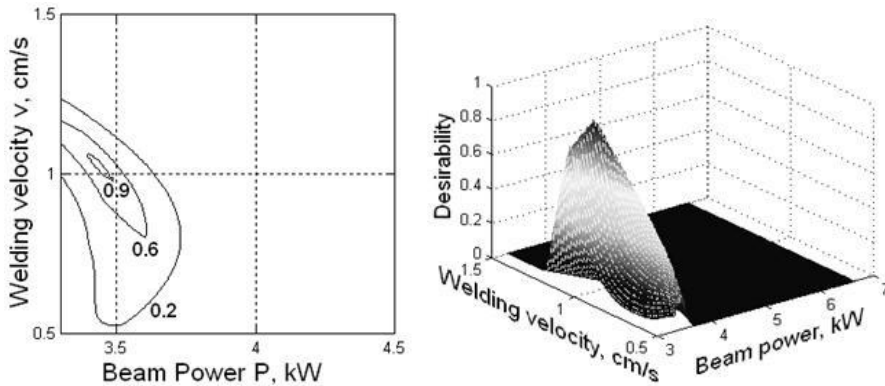


Figure 78. Desirability function G (2D- and 3D-view) at EBW of steel 45, dZ = 55.3 mm

On Figure 78 the desirability function is calculated for the fusion zone depth and width at EBW of St45. The required values for the weld geometry parameters are: $H_W=22.5$ mm, $B_W=3.5$ mm with tolerances: H in the region $[20\div 25]$ mm, B – $[2.5\div 4.5]$ mm]. The maximum value of $G_{max}=0.9442$, for $P=3.4675$ kW, $v=1.0000$ cm/s and $dz=55.3$ mm.

The trained neural networks can also be implemented for prediction of the considered performance characteristics over the experimental region and their individual optimization (for the H and η_T – maximum and for B - minimum) at EBW of stainless steels. In Table 16 are presented the optimal results, the corresponding optimal process parameter values and the values of the rest two performance characteristics predicted at the same EBW process conditions. It can be seen that the most deep welds do not coincide with the regimes with maximum thermal efficiency, the minimum width of the welds is obtained for weld depths about 25 mm, the maximum thermal efficiency is reached at regime conditions at which the focus position is 150 mm above the sample surface and the welds are comparatively wide and shallow.

In Figure 79 is presented a contour plot of the thermal efficiency, depending on the distances to the beam focus and to the sample surface (z_0 and z_p), at optimal values of the beam power $P = 7.14$ kW and the welding velocity $v = 20$ cm/min, at which the maximum thermal efficiency is reached (Table 20). It can be seen that values above 0.5 (50%) are reached at focus positions considerably below the sample surface. Figure 80 shows the corresponding (the same process parameters P and v) contour plots of the weld depth and mean width. At these conditions the most deep and narrow welds are obtained for small distances to the sample surface and focus positions a below its surface. Since the optimal

solutions for each performance characteristic are different, a compromise solution must be found, fulfilling the requirements for all the characteristics at the same time.

Table 20. Optimal regimes and weld quality performance characteristics (SSt)

	P, kW	v, cm/min	z_o , mm	z_p , mm	H, mm	B, mm	η_T
$H_{\max}$	8.40	20	196	126	45.69	2.60	0.356
$B_{\min}$	8.40	74	266	126	24.69	1.00	0.266
$\eta_{T, \max}$	7.14	20	176	326	12.38	5.27	0.687

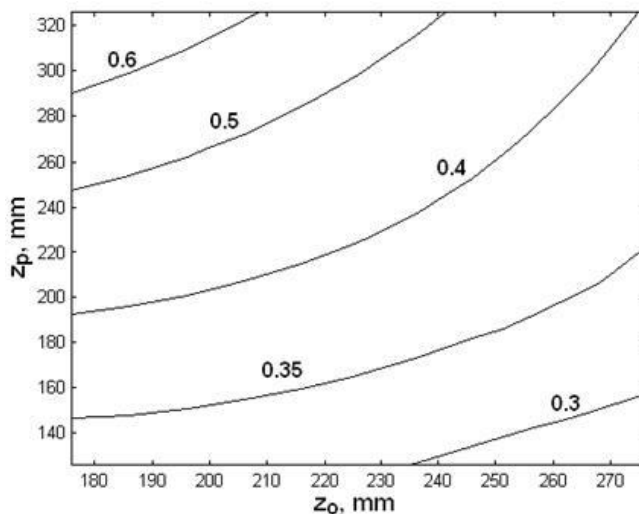


Figure 79. Contour plot of the thermal efficiency, depending on the distances z_o and z_p , at values of $P = 7.14$ kW and $v = 20$ cm/min (SSt)

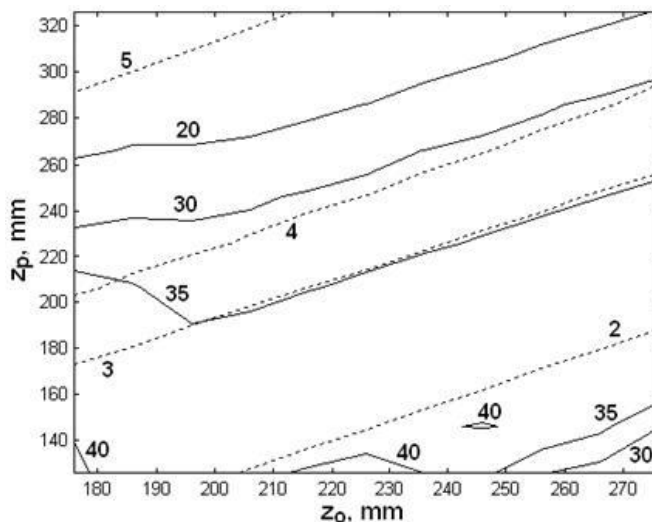


Figure 80. Contour plot of the weld depth (solid lines) and the weld mean width (dashed lines), depending on the distances z_o and z_p , at $P = 7.14$ kW and $v = 20$ cm/min (SSt)

The optimization task in the case of quality improvement at production conditions based on robust engineering approach is defined as variance minimization, while the weld geometry parameters are kept on the required values. If we want to obtain a weld depth of $H=20$ mm (with 2% tolerance), the parameter regime with the lowest variance $\hat{s}_{H\min}^2=0.1649$ in production is: $P=7.77$ kW, $V=13.333$ mm/s, $dz=-8$ mm. The estimated value of the mean for the depth is $\tilde{y}_H=20.2251$ mm. If the target value for the width is $B=2.5$ mm (with 5% tolerance), the regime with the minimum variance $\hat{s}_{B\min}^2=0.1649$ is obtained for: $P=6.93$ kW, $V=3.333$ mm/s, $dz=-78$ mm. The calculated value for the width is then $\tilde{y}_B=2.4887$ mm. A simultaneous optimization of the weld width and depth is done for the same target values for $H=20$ mm and for $B=2.5$ mm and the regime with a minimum variance at which these values are obtained is: $P=7.35$ kW, $v=8.333$ mm/s, $dz=27$ mm. This is a compromise solution in favor of both the weld depth and weld width. The values of the compromise variances and the corresponding estimated values of B and H are: $\hat{s}_{B\min C}^2=0.16519$, $\tilde{y}_B=2.509$ mm, $\hat{s}_{H\min C}^2=7.0979$, $\tilde{y}_H=19.621$ mm.

CONCLUSION

The results of calculations using steady state models (namely moving linear heat source) can be used for rough (initial) technology parameter choice. One can apply this model at admission of the known value of the width or the depth of the weld as well as at prognosis of the both values: the width and the depth of the weld as a pair at calculating its values on the basis of known welding and material characteristics. But such estimation has a big disadvantage due to not taking in the account the position of the beam focus relatively to the sample surface (or the beam focusing current changes and the variations of the distance gun-sample). The beam physical parameters (radial and angular distributions or the beam emittance) are not included too.

The proposed statistical approach gives more deep knowledge of the process characteristics influence on the weld geometry parameters. The region of application of created models is limited to studied material and EBW machine due to nature of the quantitative information obtained. It is appropriate for computer expert systems for EBW operator or technologist advice as well as for CNC systems and for computer optimization of results of EBW applications in the laboratories, at workshop services and mass production in the industry.

The functional elements of the developed expert system for electron beam weld characterization and parameter optimization, which gives the possibility for fulfilling various modeling and optimization tasks, are reviewed. This tool can be upgraded with new experimental data and now incorporates the accumulated knowledge for EBW of stainless steel and steel 45.

The tool integrates several options for:

Design of experiment for obtaining objective information on the influence of material and process parameters on EBW with minimum number of experiments.

Estimation of models. This permits to find acceptable regions for the EBW process; to estimate the significance parameters and to understand the interactions between the factors.

Process parameter choice at various requirements and conditions (defects, desirability function, robust engineering at industrial production processes etc.)

Multi-criteria parameter optimization - compromise Pareto-optimal regimes (for example maximum H and minimum B).

REFERENCES

- [1] Vutova, K; Mladenov, G. Evaluation of the dimensions of weld and thermal affected zones during EBW, Proc. Fourth Int. Conf. EBT'94, Varna, 1994, 6-11 June, 101-107.
- [2] Koleva, E; Mladenov, G; Vutova, K. Calculation of weld parameters and thermal efficiency in electron beam welding, *Vacuum*, 1999, 53, 67-70.
- [3] Swift-Hook, DT; Gick, AE. Penetration welding with laser, *Marcwood Engineering Lab*, R/M/N, 637, June 1972, 1-10, see also *Weld. J.* v.5, 1973, 492-499.
- [4] Dvorkin, II; Ledovskoy, VP; Mladenov, GM. *Electronnaia tehnika*, ser.16, No 4(8), (1970), (in Russian).
- [5] Mladenov, G; Petrov, P. Ermittlung der Prozessparameter zum Elektronenstrahlschweißen durch Computer. In *Schweißen und Schneiden*, 1993, 45-N3, 145-147.
- [6] Petrov, P; Mladenov, G. Theoretical analysis of heat flow and structural changes during electron beam irradiation of steel, *Vacuum*, 1991, v.42, No 1/2, 29-32.
- [7] Sabchevski, S; Mladenov, G; Wojcicki, S; Dabek, J. An analyse of electron gun for welding, *J. Phys. D: Appl. Phys.*, 1996, 29, 1446-1453.
- [8] Dilthey, U; Bohm, St; Dobner, M; Trager, G. Comparability and replication of the DIABEAM measurement device, Proc.of 5-th Int. Conference on Electron beam technologies, 1997, 2-5 June Varna, Bulgaria, 76-83.
- [9] Dilthey, U; Böhm, S; Welters, T; Ilyin, S; Turichin, G. EBSIM - eine Simulationssoftware für das Elektronenstrahlschweißen, *Große Schweißtechnische Tagung*, 1997, 10.-12.9.1997, Essen.
- [10] Friedel, K; Felba, J. Quantitative study of experimental emittance diagrams, Proc.of 4-th Intern.Conf. on Electron beam technologies, 5 - 11 June 1994. Varna, Bulgaria, 55 - 62.
- [11] Wojcicki, S; Mladenov, G. A new experimental investigation of high power electron beam, *Vacuum*, 2000, v 58, 523-530.
- [12] Koleva, E; Vutova, K; Wojcicki, S; Mladenov, G. Use of radial distribution of the beam current density for evaluation of the beam emittance and brightness, *Vacuum*, 2001 v 62, N2-3, 105-111.
- [13] Koleva, E. –Statistical modeling and computer programme for optimization of the electron beam welding of stainless steel, *Vacuum*, 2001, v62, N2-3, 151-157.
- [14] Koleva, E. EB weld parameters and thermal efficiency improvement, *Proc.7-th Intern. Conf. EBT, Varna*, 2003, 1-6 June 210-220.
- [15] Koleva, E. *Proceed.of Symp. Elektronika'2000*, Botevgrad, 2000, 5-6 Oct. 117-124 (In Bulgarian).
- [16] Koleva, E; Mladenov, G. Analysis of the Thermal Processes and the Shapes of Melted zones at Electron Beam Welding and Electron Beam Melting. *Bulg. J. Physics*, 2000, 27, No4, 83-96.

- [17] Koleva, E; Vuchkov, I. Model based approach for quality improvement of EBW applications in mass production, *Proc.7-th Intern. Conf. EBT*, Varna 1-6 June 2003, 221-229.
- [18] Arata, Y; Matsuda, F; Murukami, T. *Trans. of JWRI*, 1973, Vol.2, No.2, 23.
- [19] Irie, H; Hashimoto, T; Inagaky, M. *Trans. of Nat. Res. Inst. for Metals*, 1981, Vol. 23, No.2, 22.
- [20] Petrov, P; Dyakov, T; Mladenov, G. *Univ. Annual Report Technical Physics*, Sofia, 1987, Vol. 24, No.1, 171.
- [21] Lucas, WJ. *Inst. of Metals*, 1971, 99(2659), 335-340.
- [22] Bell, RA; Lippold, JC; Adolphson, DR. *Welding Journal*, 1984, 63(11), 325-332.
- [23] Bertinelly, F. et al, *Proceedings of EPAS 2004, Lucerne, Switzerland*, 1837-1839.
- [24] Geng, RL; Barnes, P; et al, *Proceedings of Particle Accelerator Conference*, May 16-20, 2005, Knoxville, TN.USA.
- [25] Nagawa El-Shahat, M.Sc.*Thesis*, Cairo University, 10, 1996.
- [26] Wei, PS; Kuo, YK; Ku, JS. *J.of Heat Transfer*, 2000, 122(3), 626-631.
- [27] Umino, T; Suzuki, M; Shida, T. *US Patent*, 3935417.
- [28] Tong, H; Gied, W. *Rev.Sci.Instr*, 1969, Vol. 40, No.10, 1283.
- [29] Dyakov, T; Petrov, P; Mladenov, G. *Proc.3 Int. Conf .Electron Beam Technologies*, 1991, May 30-June 4 Varna, 367-372.
- [30] Petrov, P; Georgiev, Th; Ivanov, R. *Int. J. for the of Joining of Materials*, 1996, Vol. 8(4), 152-157.
- [31] Petrov, P; Georgiev, Ch; Petrov, G. *Vacuum*, 1998, Vol. 51, n. 3, 339-343.
- [32] Bashenko, V; Petrov, G. *Automatic Welding*, 1997, No.9, 23, (in Russian)
- [33] Ledovskoy, V; Mladenov, G. *J. Technical Physics*, 1970, Vol. 40, 2260, (in Russian)
- [34] Mladenov, G; Ledovskoy, V; Krivkov, B. *J. Phys. and Chem. of Treatment of Materials*, 1974, No. 4, 134, (in Russian)
- [35] Stefanov, B; Petrov, P; Pirgov, P. *Vacuum*, 1988, Vol. 38, No.11, 1029.
- [36] Gabovich, M; Kovalenko, V; Metallov, O. et al., *J. Technical Phys.*, 1977, Vo. 147, No.7 , 1569. (in Russian)
- [37] Petrov, P; Mladenov, G. *Vacuum*, 1991, Vol. 42, No.1/2, 29.
- [38] Mladenov, G; Petrov, P; Sabchevski, S. *4th Int. Colloq. on welding and melting by electron and laser beams, Cannes*, 1988, 139-147.
- [39] Mladenov, G. *Welding*, 1977, No.4, 6, (in Bulgarian)
- [40] Dvorkin, I; Ledovskoy, V; Mladenov, G. *Electronnaia Technica, Ser.4-Vacuum and gas discharge tubes*, 1972, 3, 54 (In Russian)
- [41] Vutova, K; Mladenov, G. *Proc. Fourth Int. Conf. EBT'94*, Varna, 6-11 June, 1994, 101-107.
- [42] Mladenov, G; Petrov, P. *Schweiben and Schneiden*, 1993, Vol. 45, No. 3, 145.
- [43] Koleva, E; Mladenov, G; Vutova, K. *Vacuum*, 1999, 53, 67-70.
- [44] Rykalin, N.; Uglov, A; Zuev, I; Kokora, A. *Lazer and EB Material Proocessing*, Mir Publishers, Moscow, 1988, 412.
- [45] Hashimoto, T; Matsuda, J. *Trans. Nat. Res. Inst. for Metals*, 1967, Vol.9, No.1.
- [46] Tong, H; Giedt, W. *Pap. Amer. Soc. Mech. Eng. No. WA/HT-2*, 1.
- [47] Petrov, P; Mladenov, G; Michailov, V. *Proc. Int Conf Electron Beam Technologies*, Varna, May 26-June 2, 1985, 183-189 (in Russian)
- [48] Michailov, V; Petrov, P. *Automatic Welding*, 1988, No.5, 13, " (in russian)
- [49] Petrov, P. *Int J for the Joining of Materials*, 1992, Vol. 4, No.4, 110.
- [50] Rikalin, N. *Calculation of Welding Thermal Processes*, Mashgiz Publ. House, Moscow, 1951, 291. (in Russian).
- [51] Hashimoto, T; Matsuda, J. *Trans. Nat. Res. Inst. for Metals*, 1967, Vol. 9, No.1.

- [52] Tong, H; Giedt, W. Pap. *Amer. Soc. Mech. Eng.*, No. WA/HT-2, 1, 1970.
- [53] Petrov, P; Mladenov, G; Michailov, V. *Proc. Int. Conf Electron Beam Technologies*, Varna, 1985, May 26-June 2, 183-189 (in Russian)
- [54] Michailov, V; Petrov, P. *Automatic Welding*, 1988, No.5, 13, (in Russian)
- [55] Petrov, P; Mladenov, G. *Proc. of Second Int. Conf EBT 88*, Varna May 31-June 4, 472-479, 1988, (in Russian)
- [56] Mladenov, G; Ledovskoy, V; Krivkov, B. On thermal model of EBW with deep penetrating beam, *Physics and Chemistry of Material treatment*, N4, 1974, 134 (In Russian).
- [57] Khuri, AI. *Analysis of multiresponse experiments, in Statistical design and analysis of industrial experiments*, Ed.S. Ghosh, 231-246.
- [58] Koleva, E., *Vacuum*, 2001, 62, 151-157.
- [59] Myers, RH; Karter, WH. Response surface techniques for dual response systems, *Technometrics*, 1973, 15, 301-317.
- [60] Khuri, AI. *Analysis of multiresponse experiments, in Statistical design and analysis of industrial experiments*, Ed.S. Ghosh, N.Y., Marcel Dekker, 1987, 231-246.
- [61] Myers, RH; Karter, WH. Response surface techniques for dual response systems, *Technometrics*, 1973, 15, 301-317.
- [62] Vuchkov, I; Boyadjieva, L. *Quality Improvement with Design of Experiments*, Kluwer Acad. Publishers, ed. Keller A, 2001.
- [63] Jang,, J; Sun, C; Mizutani, E. *NeuroFuzzy and Soft Computing*, Prentice Hall Publishing, 1997.
- [64] Koivo, HN. *Artificial Neural Networks in Fault Diagnosis and Control*, Control Engineering Practice, 1994, 2(1), 89-101.
- [65] Ungar, U; Powell, B; Kamens, S. Adaptive Networks for Fault Diagnosis and Process Control, *Computers and Chem. Engng.*, 1990, v.14, No. 4/5, 561-572.
- [66] Amari, S; Kasabov, N. Eds., *Brain-like Computing and Intelligent Information Systems*, Springer Verlag, 1997.
- [67] Chen, J. (1998), Systematic Derivations of Model Predictive Control Based on Artificial Neural Networks, *Chemical Eng. Communications*, 164, 35-39.
- [68] Sorsa, T; Koivo, HN; Koivisto, H. Neural Networks in Process Fault Diagnosis, *IEEE Trans. on Systems, Man, and Cybernetics*, 1991, 21(4), 815-825.
- [69] Tsai, CS; Chang, CT. Dynamic Process Diagnosis via Integrated Neural Networks, *Computers Chem. Engng.*, 1995, v. 19, Suppl., S747-S752.
- [70] D; Rumelhart, J. McClelland, (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, MIT Press, Cambridge, Mass., 1986.
- [71] Koleva, E; Vuchkov, I. Model-based approach for quality improvement of EBW applications in mass production, *Vacuum*, 2005, 77, 423-428.
- [72] Koleva, E; Vuchkov, I; Velev, K. Multiresponse Robust Engineering: Case with Errors in Factor Levels. *PLISKA Studia Mathematica Bulgarica*, 2009, 19, 193-206.

Chapter 3

AUTOMATION IN DETERMINING THE OPTIMAL PARAMETERS FOR TIG WELDING OF SHELLS

***Asif Iqbal^{*1}, Naeem Ullah Dar¹
and Muhammad Ejaz Qureshi²***

¹Department of Mechanical Engineering, University of Engineering & Technology,
Taxila, Pakistan

²College of Electrical & Mechanical Engineering, National University of Sciences &
Technology, Rawalpindi, Pakistan

ABSTRACT

Residual stresses and distortion are the two most common mechanical imperfections caused by any arc welding process and Tungsten Inert Gas (TIG) Welding is no exception to this. A high degree of process complexity makes it impossible to model the TIG welding process using analytical means. Moreover, the involvement of several influential process parameters makes the modeling task intricate for the statistical tools as well. The situation, thus, calls for nonconventional means to model weld strength, residual stresses and distortions (and to find trade-off among them) based on comprehensive experimental data.

Comprehensive Designs of Experiments were developed for the generation of relevant data related to linear and circumferential joining of thin walled cylindrical shells. The base metal utilized was a High-Strength Low Alloy Steel. The main process parameters investigated in the study were welding current, welding voltage, welding speed, shell/sheet thickness, option for trailing (Argon), and weld type (linear and circumferential).

For simultaneous maximization/minimization and trade-off among aforementioned performance measures, a knowledge base – utilizing fuzzy reasoning – was developed. The knowledge-base consisted of two rule-bases: one for determining the optimal values of the process parameters according to the desired combination of maximization and/or minimization of different performance measures; while the other for predicting the values of the performance measures based on the optimized/selected values of the various

^{*} Corresponding author: Email: asif.asifiqbal@gmail.com

process parameters. The optimal formation of the two rule-bases was done using Simulated Annealing Algorithm.

In the next stage, a machine learning (ML) technique was utilized for creation of an expert system, named as EXWeldHSLASteel, that could: self-retrieve and self-store the experimental data; automatically develop fuzzy sets for the numeric variables involved; automatically generate rules for optimization and prediction rule-bases; resolve the conflict among contradictory rules; and automatically update the interface of expert system according to the newly introduced TIG welding process variables.

The presented expert system is used for deciding the values of important welding process parameters as per objective before the start of the actual welding process on shop floor. The expert system developed in the domain of welding for optimizing the welding process of thin walled HSLA steel structures possesses all capabilities to adapt effectively to the unpredictable and continuously changing industrial environment of mechanical fabrication and manufacturing.

1. INTRODUCTION

The word Residual stresses and distortion are the two most common mechanical imperfections caused by any arc welding process and Gas Tungsten Arc Welding (GTAW) is no exception to this. Residual stresses are those stresses that would exist in a body if all external loads and restraints were removed. Weld induced residual stresses are produced in a structure as a consequence of local plastic deformations introduced by local temperature history consisting of a rapid heating and subsequent cooling phases. During the welding process, the weld area is heated up sharply compared to the surrounding area and fused locally. The material expands as a result of being heated [1]. The heat expansion is restrained by the surrounding cooler area, which gives rise to thermal stresses. The thermal stresses partly exceed the yield limit, which is lowered at elevated temperatures. Consequently, the weld area is plastically hot-compressed. After cooling down too short, too narrow or too small as compared to the surrounding area, it develops tensile residual stresses, while the surrounding areas are subjected to compressive residual stresses to maintain self-equilibrium [2].

Weld induced distortion can be defined as change in shape and/or dimension of a welded structure when it is free from any of the external forces of thermal gradients. The interaction of solidifying weld metal with the parent base metal, results in change in dimensions and shape of the weldments, generally referred to as welding distortions [3]. The residual stresses and the structure deformations are highly affected by the usage of welding fixtures during welding process and the amount of restraint determines the control of distortions and residual stress fields on the weldments [4]. Generally, there is a trade-off between magnitudes of residual stress and distortion and the amount of the restraint is determined as per structural design requirements.

Thin-walled shells comprise an important and growing proportion of engineering manufacture with areas of application becoming increasingly diverse, ranging from aircraft, missiles, ships, pressure vessels, bridges and oil rigs to storage vessels, industrial buildings and warehouses. Thin-walled shells are designed with advanced numerical analysis techniques and manufactured using sophisticated fabrication processes. The effects of geometrical/structural imperfections in thin-walled shells may introduce changes in the stresses that are nearly equal to the stresses due to the loads [5]. Permanent joining of thin-

walled shells, using welding process, is the critical most area of their manufacture. They are highly vulnerable, due to their slender structure, to catastrophic distortions caused by the generation of immense heat during the process. Application of restraints, in order to avoid distortions, on the other hand, leads to impartation of structure-weakening residual stresses.

Gas Tungsten Arc Welding (GTAW) or Tungsten Inert Gas (TIG) welding is one of the most well established processes of arc welding type. TIG welding has been the most widely accepted welding processes so far in the industry due to its availability and versatility of welding equipment, low cost equipment, excellent quality and skilled welders. The TIG welding process attains a good position in respect of the total cost specifically for thin walled structures because of the medium equipment cost and mainly due to low wire cost i.e. low deposition rates due to lower wire feed speeds [6]. Many parameters affect TIG welding quality, such as base metal, filler wire, weld geometry, electrode type, shielding gas type, welding current, and travel speed of the welding torch. The desired welding parameters are usually determined based on experience or handbook values. However, this does not ensure that the selected welding parameters result in near optimal welding quality characteristics for the particular welding system and environmental conditions.

1.1. Variables and Performance Measures in TIG Welding Process

Following are some of the basic parameters of welding process besides pre-heating, inter-pass temperature, post-heating and no. of weld passes etc:

1. **Material.** Base metal properties like material composition and material properties (like thermal conductivity, coefficient of thermal expansion, reaction with atmospheric oxygen, effect of flux residue, and crack sensitivity) are considered as the most influential parameter.
2. **Weld geometry.** It is used for the selection of welding process. The joint type may be butt, lap, fillet or T-joint. Bevel may be single-V, double-V or U shape. Weld geometry is directly influential upon weld quality.
3. **Welding Position.** It can either be flat, horizontal, vertical, or overhead etc. Mainly vertical and horizontal welding position is used. Weld bead geometry is affected by the position in which the work piece is held with respect to welding gun.
4. **Shielding Gas (lit/min).** It is a protective gas used to prevent atmospheric contamination. TIG welding process is mostly conducted in shielding. Shielding Gas Flow Rate has significant effect on weld bead shape which in turn effects the distortion, residual stresses, heat effected zone (HAZ) and mechanical properties of the material to be welded.
5. **Welding Speed (cm/min).** It is the parameter that varies the weld penetration and width of beads. Maximum weld penetration is at a specific welding speed and decreases as speed varies. The increased input heat per unit length due to reduced speed results increase in weld width and vice versa.
6. **Wire Feed Rate (cm/min).** It is the parameter that controls the speed of welding filler wire. It is normally attributed to increased resistance heating which itself is

increased with the increase in wire feed rate. The welding current varies with the change in wire feeding and the relationship is linear at low feeding rate.

7. **Material Thickness (mm).** Material thickness plays a vital role in process selection and parameters setting. Material thickness is used to decide the input heat required and to control the cooling rate. Higher thickness means higher cooling rate resulting increase in heat effected zone (HAZ) and hardness of weld metal.
8. **Welding Current (Amp).** It is one of the most important parameter that directly affects the penetration and lack of fusion by affecting the speed of welding. Welding current is the current being used in the welding circuit during the making of a weld. If the current is too high at a given welding speed, the depth of fusion or penetration will be too great. For thinner plates, it tends to melt through the metal being joined. It also leads to excessive melting of filler wire resulting in excessive reinforcement.
9. **Welding Voltage (V).** It is the parameter that directly affects the bead width. It also influences the microstructure and even the success and failure of the operation. Like current, welding voltage affects the bead shape and the weld deposit composition. Increase in the arc voltage results a longer arc length and a correspondingly wider, flatter bead with less penetration.

Following are some of the important performance measures of welding process, besides weld quality, toughness, hardness, ductility, HAZ and FZ etc:

1. **Weld Strength (MPa).** It is the most important performance measure that directly affects the weld efficiency and production cost. Mostly, the weld quality is based and judged by the weld strength and the strength of base metal. Many factors influence the weld strength including the base material, filler metal, weld type, joint type, weld method, heat input, and their interactions.
2. **Weld Induced Residual Stresses (MPa) & Distortions.** Residual stress is the most important welding performance measure. Both, residual stresses and distortions are the major concerns in welded structures. The residual stresses in weld region are normally tensile and close to the material yield stress due to the shrinkage of the weld during cooling. The residual stresses have a significant effect on the process of the initiation and further propagation of the fatigue cracks in welded elements. The fatigue life of the welded elements depends on the possible variations of the residual stress level and in many cases the residual stresses are one of the main factors, determining the engineering properties of structural components, and plays a significant role in fatigue of welded elements. In welding process, low values of residual stresses and distortions are desired.
3. **Welding Temperatures (oC).** The temperatures experienced by the metal produced by weld torch during the welding process are called as weld temperatures. The amount of heat input during welding process is very important as the high heat input results increase in heat affected zone (HAZ).

1.2. Current Challenges in TIG Welding of Shells

At present, following three challenges, related to the TIG welding of shells, constitute the focus of major research activities:

1. Weld induced imperfections like residual stresses and distortions are the major demerits of arc welding technology that adversely affects the weld efficiency. Thus, it is a primary need of the present time to search for the welding conditions that could significantly suppress the weld induced imperfections. By the term welding conditions it is meant here the different combinations of welding wire parameters (e.g., weld wire speed, wire composition, wire size etc.), welding parameters (e.g., welding speed, welding current, welding voltage, thickness of base metal & composition, weld type, weld geometry etc.), heating (pre-heating or post heating) and cooling (e.g., air or gas etc.).
2. The parameters that lead to enhanced weld strength do not necessarily provide minimum residual stresses or distortion. In addition, it is also well known that the parameters favorable for low distortion also cause increase in residual stresses. These two facts imply that the challenge sought is two-folded. The researchers are required, not only, to find the ways to minimize residual stresses and distortion but also to make sure that weld strength is not compromised. In other words, researchers have to find the trade-off among the two conflicting objectives: (a) maximize weld strength; and (b) minimize residual stresses and distortions.
3. It is also highly desired to have a fully automated system that should acquire knowledge from the data generated by the research activities and utilize that knowledge to: (a) work out the optimal welding conditions for achievement of desired objectives in a best possible way; and (b) predict the values of performance measures based upon welding conditions selected.

The chapter targets minimizing the weld induced structural imperfections and seeking trade-off between two of its most common types, i.e., distortion and residual stress, in GTAW (linear as well as circumferential weld) of thin-walled cylindrical shells. The base metal worked upon will be a common high strength low alloy steel (HSLA) and the optimization process will be based on a comprehensive Design of Experiments (DoE) that would get the results from actual experiments. The effects of following five input parameters (predictor variables) upon the welding performance measures will be sought: welding current, welding voltage, welding speed, shell thickness, and Argon trailing. Before going on to the actual work, it is pertinent to have a brief review of the most relevant literature.

In [7], a design of experiments approach was chosen as an efficient technique to maximize the information gained from the experimentation for the reduction of pores in welds by laser at a car production line as case study and an average reduction in the number of pores of 97 per cent was obtained. In [8], the researchers presented the use of response surface methodology (RSM) by designing a four-factor five-level central composite rotatable design matrix with full replication for planning, conduction, execution and development of mathematical models for predicting the weld bead quality and selecting optimum process parameters for achieving the desired quality and process optimization of Submerged arc welding (SAW) of pipes of different diameters and lengths. In [9], the authors established

relationships between the laser-welding parameters (laser power, welding speed and focal point position) and the three responses (tensile strength, impact strength and joint-operating cost) for butt joints made of AISI304. The optimization of the welding process was done by orthodox DoE techniques in order to increase productivity and minimize total operating cost.

In [10], researchers presented a specially designed test rig which was developed and used for assessment of thermal and residual stresses for given welding conditions characterized by the peak temperature and cooling time of the thermal cycle of high strength low alloy quenched and tempered steel. An induction coil used for programming the heating and cooling of small specimens for simulation of actual weld thermal cycles. The chosen range of peak temperature and cooling time produced varying effects on the temperature field, micro-structural state field, and mechanical field. This technique facilitated the study of important relationships between weld thermal cycles, phase transformations and residual stresses.

Many welding distortion mitigation methods have been developed by the researchers to eliminate weld induced imperfections. For this purpose, several researchers have used the trailing heat sink during welding to minimize distortion. This method is called dynamically controlled low stress no distortion (DC-LSND) welding, which was first developed and introduced by Guan et al. [11]. However, still its practical application and implementation is complex. In this method, a trailing heat sink is attached at some short distance behind the welding heat source and moved as the welding heat source. Usually this method is used to control the weld buckling of thin plates as the compressive stresses developed during welding of thin sections exceed the critical level of buckling stress. The welding longitudinal residual stresses are affected significantly with the application of trailing heat sink and the residual stresses remain below the critical buckling stress level and consequently minimize buckling. In [12], the two steel plates of AISI 316L of size 250x100x1.5 mm were welded by TIG welding with same parameters (3mm/s, 750 W) with and without the application of trailing heat sink (at fixed distance of 25mm from welding torch, CO₂ as cooling media of trailing). The plate welded without trailing application was severely buckled whereas the plate welded with trailing application was free of buckling. In [13], the researcher presented several approaches to analyze the effects of the cooling source parameters. It was determined, analytically, that the sensitivity of buckling depends upon stress levels and their distribution behavior and decreases with the decrease of width of compressive zone at the plate edges that can be achieved with the increase in tension zone width or compressive zone on the weld.

The analytical approaches were replaced by numerical approaches after the advent of finite element (FE) based numerical simulation techniques for modeling in welding. It is possible to account for nonlinear effects like temperature-dependent convection and radiation to the surrounding medium, plastic flow and volume expansion during possible final phase transformation with the use of FEM. Modeling of moving heat source for the analytical solution of transient temperature distribution in arc welding process presented by Rosenthal [14] was the first step towards the simulation of welding phenomenon. The author presented linear 2D and 3D heat flow in a solid of infinite size bounded by planes and also validated the model through experimentally measured temperature distributions during plate welding of different geometries. A predefined temperature at some specified locations of weld was used by Goldak et al. [15]. To overcome the issues in previously presented heat source model, Goldak et al. [16, 17] developed the most dominating heat source model with Gaussian heat source distribution, which is also known as Double Ellipsoidal Heat Source model and most widely utilized now-a-days. Rybicki et al. [18] presented a numerical study of multi-pass

welding regarding the effect of pipe wall thickness on welding residual stresses, which is of significant importance for relating residual stresses with geometrical size of the pipe. Basically, it was a parametric study in which basic FE model was validated for residual stresses measured experimentally and subsequently developed FE model was used for different welding parameters and geometrical dimensions of the pipe.

1.3. Application of Artificial Intelligence in Optimizing Welding Process

The requirement number 3 described in the sub-section 1.2 is a hot candidate for application of Artificial Intelligence (AI) tools. AI is a branch of science that imparts to machines the ability to think and reason. Precisely, it can be defined as the simulation of human intelligence on a machine, so as to make the machine efficient to identify and use the right piece of knowledge at a given step of problem solving [19]. The artificial intelligence (AI) is related to intelligent behavior i.e. perception, reasoning, learning, communicating, and acting in complex environments, in artifacts having long term goals, both engineering and scientific, of development of machines that can do as human or better [20]. The ultimate target of research in field of AI is to construct a machine that can mimic or exceed human mental capabilities including reasoning, understanding, imagination, and creativity [21]. In a very broad sense AI can be subdivided into two categories: (1) Knowledge-Based Systems (KBS); and (2) Computational Intelligence (CI).

KBS is a kind of non-conventional computer program in which knowledge is kept explicitly separate from the control module of the program. The module that contains the knowledge, in the form of rules and facts, is called knowledge-base, while the control module is called inference engine. The inference engine contains meta-knowledge i.e. the knowledge about how, where, and when to apply the knowledge. Expert System (ES) is a special kind of KBS that contains some extra frills like knowledge acquisition module and explanation module etc [21]. An expert system is a computer program designed to simulate the problem solving behavior of a human who is an expert in a narrow domain or discipline. An expert system is normally composed of a knowledge base (information, heuristics, etc.), inference engine (analyzes the knowledge base), and the end user interface (accepting inputs, generating outputs). The path that leads to the development of expert systems is different from that of conventional programming techniques. Expert systems are capable of delivering quantitative information or for use in lieu of quantitative information. Another feature is that these systems can address imprecise and incomplete data through the assignment of confidence values to inputs and conclusions. One of the most powerful attributes of expert systems is the ability to explain reasoning. ES possesses high potentials for optimizing the process parameters and improving the manufacturing efficiency/effectiveness.

CI is different from KBS in the sense that in CI the knowledge is not explicitly stated in form of rules or facts, rather it is represented by the numbers, which are adjusted as the system improves its efficiency. One of the common forms of CI is the Artificial Neural Network (ANN) [21]. A brief literature review regarding application of Expert System / Artificial Intelligence to the domain of welding process engineering is provided as under.

The term artificial intelligence was named by John McCarthy in 1956. In artificial intelligence (AI) field until early 1970s, the researchers acknowledged that the general

purpose problem solving methods developed since 1960s were not capable to tackle the to-day complex research and application oriented problems and felt that there was a need of specific knowledge related to a specific and limited domain of application rather than a general knowledge for many domains. This reason was made a base for the development of knowledge-based systems i.e. expert systems and this technology remained dominant in the field of AI. The history of numerous knowledge-based systems developed earlier can be found in [22]. A good and broad view definition of AI field by Tanimoto is as “Artificial Intelligence is a field of study that encompasses computational techniques for performing tasks that apparently require intelligence when performed by human. It is a technology of information processing concerned with processes of reasoning, learning, and perception” [22]. In 1970s, the areas emerged in the AI field were knowledge-based systems (expert systems), natural language understanding, learning, planning, robotics, vision and neural networks.

An expert system (ES) that uses a collection of fuzzy rules, facts and membership functions to draw conclusion and uses fuzzy logic for inferencing rather than boolean logic is called a fuzzy expert system (FES) [23]. In 1975, Lotfi A. Zadeh proposed the fuzzy set theories and fuzzy logic that deals with reasoning with inexact or fuzzy concepts. Fuzzy logic (FL) computes with words rather than with numbers whereas the fuzzy logic controller (FLC) controls with rules (IF-THEN) rather than with equations [23].

Traditionally, AI covers several application areas in manufacturing. Recently developed systems have demonstrated the importance of AI based software to produce intelligent engineering software that can make many routine engineering decisions for welding applications and guide a human user to optimum decisions for welding to save cost and human hours. Mostly, these systems utilize expert systems and neural networks technology to provide and predict accurate weld process models and engineering decision making capability [24]. Usually expert systems in welding include the application to select the suitable filler metal type and size, to determine the pre-heat and post-weld heat-treatment schedules, to determine welding parameters and others [24]. In [25], the authors presented a fuzzy expert system approach for the development of the classification of different types of welding flaws in the radiographic weld domain. The fuzzy rules were generated from the available examples using two different methods and the knowledge acquisition problem was carried by using two machine-learning methods by using a simple genetic algorithm to determine the optimal number of partitions in the domain space. In [26], the researchers reported that expert system technique is more fruitful approach to the automated generation of procedural plans for arc welding than previous algorithmic methods. The main purpose was to evaluate recent computing advances in the context of planning for arc welding and to extract more generic knowledge about the application of expert system techniques to advanced manufacturing problems. In [27], the authors developed an expert system for quenching and distortion control in a heat treatment process. The goals of this expert system were predicting results obtained under given quenching conditions and to improve the performance by supporting decision making. In [28], a genetic algorithm and response surface methodology was used for determining optimal welding conditions and desirability function approach was used for different objective function values. Application of the method proposed in this research revealed a good result for finding the optimal welding conditions in the gas metal arc (GMA) welding process. In [29], an integrated approach comprising the combination of the Taguchi method and neural networks for the optimization of the process conditions for GTA welding

was presented. Taguchi method was used for design of experiments and initial optimization with ANOVA for the significance of parameters of TIG welding (Electrode size, Electrode angle, Arc length, Welding current, Travel speed, and Flow rate). In [30] the authors presented a novel attempt to carry out the forward (the outputs as the functions of input variables) and reverse (the inputs as the functions of output variables) modeling of the metal inert gas welding (a multi-input and multi-output) process using fuzzy logic based approaches. The statistical regression analysis was used for the forward modeling efficiently. The developed soft computing-based approaches were found to solve the above problem efficiently. In [31] a prototype knowledge based expert system named WELDES was presented. WELDES was developed to identify the aluminum welding defects, to correlate them with the welding parameters (which cause them), and to offer advice regarding the necessary corrective actions for a ship industry.

1.4. Inadequacies of Previously Developed AI Based Automation Tools

Most of the previously developed AI based automation tools seem to be limited in effectiveness because of following three reasons:

1. The application area is not broad, in the sense that most of the tools do not cover all the influential aspects of a manufacturing process. It can be observed that the recommendation of any controllable process parameter has been provided based upon relationship between two or three given input parameters. In pragmatic conditions there are many more influential parameters that need to be cared for in recommending optimal values of any controllable parameter for the desired response.
2. They provide single-purpose consultation. They mostly consider one objective at a time for optimization. Some of the tools provide just the prediction of some performance measures based upon limited number of input parameters.
3. They lack dynamic characteristics. Most of the tools presented are static, in the sense that they lack automated mechanism for expanding their knowledge or increasing the application range with experience.

The chapter presents an expert system that optimizes the TIG welding (linear and circumferential) of thin walled shells of a high strength low alloy (HSLA) steel. The schematic of linear and circumferential weld of thin walled shell has been presented in Figure 1. The knowledge-base of the system is based on the data generated by the actual experiments. The presented expert system is a highly effective automation tool that provides the optimized values of the process parameters based on the combination of maximization and/or minimization of different objectives and also predicts the values of the performance measures based on the finalized settings of the process input parameters. Moreover, the expert system also possesses the capability of self-learning, self-correcting, and self-expanding, based on continuous feedback of the results to the system.

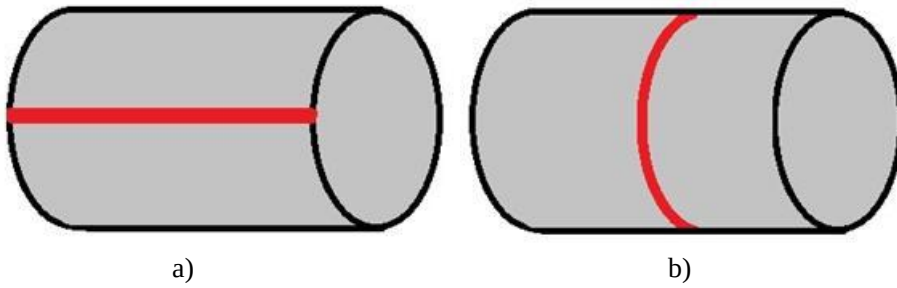


Figure 1. (a) Linear Weld and (b) Circumferential weld of a thin-walled shell

2. DESIGN OF EXPERIMENTS (DoE)

The traditional approach to experimentation require to change only one factor at a time (OFAT), while keeping others as constant and this approach doesn't provide data on interactions of factors which occurs in most of the process. The alternative statistical based approach called "two level factorial design" can uncover the critical interactions that involve simultaneous adjustments of experimental factors at only two levels: high (+1) and low (-1). The two level factorial design offers a parallel testing scheme which is most efficient than the serial approach OFAT. Two level experiments restrict the number of experiments to a minimum and the contrast between the levels gives the necessary driving force for the process improvement and optimization. The statistical approach to design of experiments (DOE) and analysis of variance (ANOVA), developed by R.A. Fisher in 1920, is an efficient technique for experimentation which provides a quick and cost effective method for complex problem solving with many variables [32].

2.1. Linear Welding of Shells

This section presents the details of experiments performed upon the experimental data of TIG welding of thin-walled, high strength low alloy (HSLA) steel cylinder (linear weld), for the purpose of analyzing and optimizing the welding parameters.

2.1.1. Predictor variables

Predictor variables are the welding process parameters that can also be represented as process input parameters or input variables. A 2^4 (4 factors, 2 levels, 16 test) full factorial design model (replicates 1, block 1, centre point per block 0 and order 4FI [factors interaction]) was used for the linear welding experiments. Tables 1, 2 and 3 show the low and high settings (or levels) for the predictor variables (or parameters) used in sixteen tests for the shell (cylinder) thickness of 3, 4 and 5mm, respectively. The practical range of the parameters (especially welding current) should be specific with respect to thickness of the material and the heat input (welding current, welding voltage and welding speed) required for the fusion.

Table 1. High and Low Settings of Factors (Predictor Variables) [t = 3mm]

Factor	Name	Units	Type	Low Actual	High Actual
A	Current	A	Numeric	170.00	210.00
B	Voltage	V	Numeric	10.50	13.50
C	Weld Speed	cm/min	Numeric	15.00	18.00
D	Trailing		Categorical	nil	Ar

Table 2. High and Low Settings of Factors (Predictor Variables) [t = 4mm]

Factor	Name	Units	Type	Low Actual	High Actual
A	Current	A	Numeric	200.00	220.00
B	Voltage	V	Numeric	10.50	13.50
C	Weld Speed	cm/min	Numeric	15.00	18.00
D	Trailing		Categorical	nil	Ar

Table 3. High and Low Settings of Factors (Predictor Variables) [t = 5mm]

Factor	Name	Units	Type	Low Actual	High Actual
A	Current	A	Numeric	230.00	270.00
B	Voltage	V	Numeric	10.50	13.50
C	Weld Speed	cm/min	Numeric	15.00	18.00
D	Trailing		Categorical	nil	Ar

Table 4. Design of 16 Experiments following Full Factorial (t = 3 mm)

		Factor 1	Factor 2	Factor 3	Factor 4
Std	Run	A:Current	B:Voltage	C:Weld Speed	D: Trailing
		A	V	cm/min	
12	1	210.00	13.50	15.00	Ar
5	2	170.00	10.50	18.00	nil
1	3	170.00	10.50	15.00	nil
3	4	170.00	13.50	15.00	nil
11	5	170.00	13.50	15.00	Ar
7	6	170.00	13.50	18.00	nil
8	7	210.00	13.50	18.00	nil
16	8	210.00	13.50	18.00	Ar
13	9	170.00	10.50	18.00	Ar
4	10	210.00	13.50	15.00	nil
9	11	170.00	10.50	15.00	Ar
10	12	210.00	10.50	15.00	Ar
2	13	210.00	10.50	15.00	nil
6	14	210.00	10.50	18.00	nil
15	15	170.00	13.50	18.00	Ar
14	16	210.00	10.50	18.00	Ar

Table 5. Design of 16 Experiments following Full Factorial (t = 4 mm)

		Factor 1	Factor 2	Factor 3	Factor 4
Std	Run	A:Current	B:Voltage	C:Weld Speed	D: Trailing
		A	V	cm/min	
12	1	220.00	13.50	15.00	Ar
5	2	200.00	10.50	18.00	nil
1	3	200.00	10.50	15.00	nil
3	4	200.00	13.50	15.00	nil
11	5	200.00	13.50	15.00	Ar
7	6	200.00	13.50	18.00	nil
8	7	220.00	13.50	18.00	nil
16	8	220.00	13.50	18.00	Ar
13	9	200.00	10.50	18.00	Ar
4	10	220.00	13.50	15.00	nil
9	11	200.00	10.50	15.00	Ar
10	12	220.00	10.50	15.00	Ar
2	13	220.00	10.50	15.00	nil
6	14	220.00	10.50	18.00	nil
15	15	200.00	13.50	18.00	Ar
14	16	220.00	10.50	18.00	Ar

Table 6. Design of 16 Experiments following Full Factorial (t = 5 mm)

		Factor 1	Factor 2	Factor 3	Factor 4
Std	Run	A:Current	B:Voltage	C:Weld Speed	D: Trailing
		A	V	cm/min	
12	1	270.00	13.50	15.00	Ar
5	2	230.00	10.50	18.00	nil
1	3	230.00	10.50	15.00	nil
3	4	230.00	13.50	15.00	nil
11	5	230.00	13.50	15.00	Ar
7	6	230.00	13.50	18.00	nil
8	7	270.00	13.50	18.00	nil
16	8	270.00	13.50	18.00	Ar
13	9	230.00	10.50	18.00	Ar
4	10	270.00	13.50	15.00	nil
9	11	230.00	10.50	15.00	Ar
10	12	270.00	10.50	15.00	Ar
2	13	270.00	10.50	15.00	nil
6	14	270.00	10.50	18.00	nil
15	15	230.00	13.50	18.00	Ar
14	16	270.00	10.50	18.00	Ar

With the increase in material thickness, the increase in heat input is required to fuse and weld. Three of these predictor variables (welding current, welding voltage and welding speed) are numeric while the other one (gas trailing) is categorical. Complete detail of the 16 runs following full factorial design has been presented in Tables 4, 5, and 6 for shell thickness of 3, 4 and 5mm, respectively. All the statistical analyses were performed using a commercial computing package named Design-Expert[®] 7.1.6, by Stat-Ease[®].

2.1.2. Response variables

Response variables are the performance measures, which can also be termed as output variables or output parameters. Following response variables will be measured in order to judge the process performance of thin walled HSLA steel welded structures:

1. Weld Strength (maximum value of tensile strength) – to be measured in MPa by testing of weld tensile samples.
2. Distortion (maximum value of weld-induced distortion in the shell at weld zone) – to be measured in mm.
3. Residual Stress: (maximum value of weld-induced stresses [Von-Mises] in the weld zone) – to be measured in MPa.

2.1.3. Fixed parameters

The welding position used was flat and single V joint geometry including angle of 70° with 1mm root face and 1mm root gap. The electrical characteristics used were DC current and straight polarity. Argon gas (99.999% Liquid) was used for shielding (25 lit/min) and for trailing (25 lit/min). The size of shielding nozzle was Ø 18mm. The sizes used for trailing were: diameter of trailing nozzle = Ø 1.3mm, distance from nozzle to sample = 5 mm, distance (centre to centre) between arc and trailing nozzle = 30mm, effective diameter of trailing = Ø 25mm. The material of backing fixture used was Copper and alcohol (99%) was used for joint cleaning after mechanical cleaning of both sides (50mm) of weld joint. Welding conditions used were humidity less than 70%, ambient temperature greater than 18°C and no draught in welding area. The material of shells used as base metal was HSLA steel 30CrMnSiA and filler wire used was H08. The chemical compositions of both of the materials have been provided in Tables 7 and 8, respectively. Table 9 presents the mechanical properties of the base metal in as-annealed condition [33]. After heat treatment (i.e. quenching and tempering), these mechanical properties of base metal reaches to ≤ 1600 MPa (Tensile Strength), ≤ 1300 Mpa (Yield Strength), $\leq 8\%$ (Elongation) and ≤ 48 HRC (Hardness). The length and outer diameter of the shells, for all the three sheet thicknesses, were fixed to 500mm and 300mm, respectively. The other welding parameters that were kept constant in all experiments are: pre-heat temperature = 175°C, inter-pass temperature = 150°C, tungsten electrode (3% thoriated) size = Ø 3.2mm, and welding wire (H08) size = Ø 1.6mm.

Table 7. Chemical Composition of 30CrMnSiA Steel (Base Metal)

	C	Cr	Si	Mn	V	Mo	Ni	P	S
Content (%)	0.28-0.32	1.0-1.3	1.5-1.7	0.7-1.0	0.08-0.15	0.4-0.55	0.25	≤ 0.01	≤ 0.013

Table 8. Chemical Composition of H08 (Filler Wire)

	C	Cr	Si	Mn	V	Mo	Ni	P	S	Al
Content (%)	0.08-0.12	1.4-1.7	1.1-1.3	0.9-1.1	0.05-0.15	0.4-0.6	1.8-2	≤ 0.006	≤0.005	≤0.10

Table 9. Mechanical properties of the Base Metal.

Tensile Strength (MPa)	Yield Strength (MPa)	Elongation (%)	Hardness (HRC)
700 – 800	500 – 600	20	20



Figure 2 (a) SAF TIGMATE 270 Power Source and (b) NERTAMATIC 300 TR

All the TIG welding experiments, described in this chapter, have been performed on SAF TIGMATE 270 AC/DC power source, SAF NERTAMATIC 300 TR and fully automatic torch control. TIGMATE 270 and NERTAMATIC 300 TR welding power sources, as shown in Figure 2, is a computerized waveform control technology for high quality TIG welds. The parameters can be controlled as desired. Automatic torch positioning system is used to control / locate the torch movement. Tack welded sheets are properly clamped (as per desired structural boundary conditions) with torch aiming at 90°.

2.1.4. Experimental results and analyses

Figures 3, 4, and 5 show the comparison of weld strength for aforementioned sixteen tests as described in Tables 4, 5, and 6, respectively. The maximum and minimum values of weld strength (Ultimate Tensile Strength) obtained with respect to thickness of the material are presented in Table 10.

Table 10. Maximum and minimum values of Weld Strength.

Thickness (mm)	Minimum (MPa)	Maximum (MPa)	Mean (MPa)	Std. Deviation
3	730.6	791	754.294	16.9567
4	722.3	780.4	743.994	16.3382
5	715	765.7	735.506	15.7428

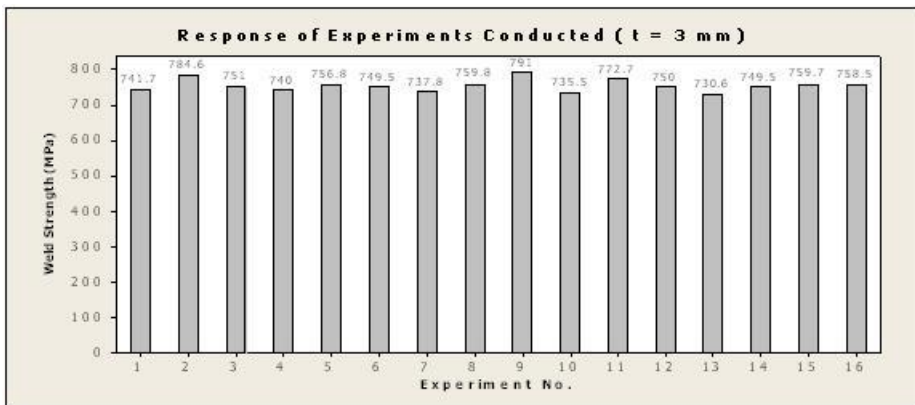


Figure 3 Weld Strength for Sixteen Experiments (t = 3 mm)

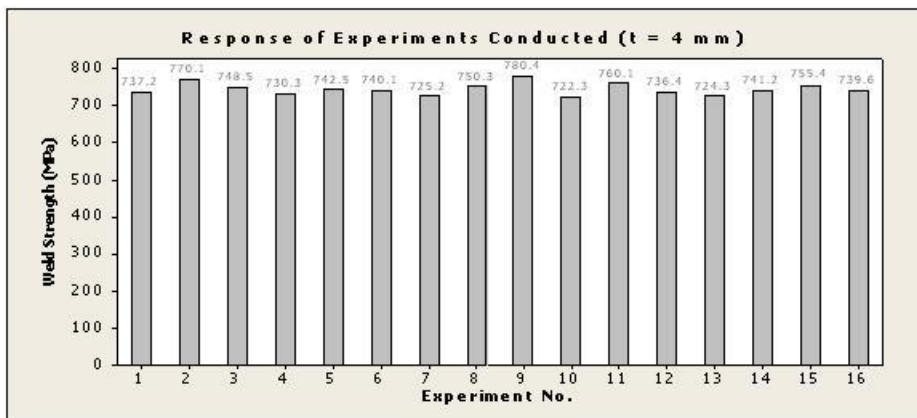


Figure 4 Weld Strength for Sixteen Experiments (t = 4 mm)

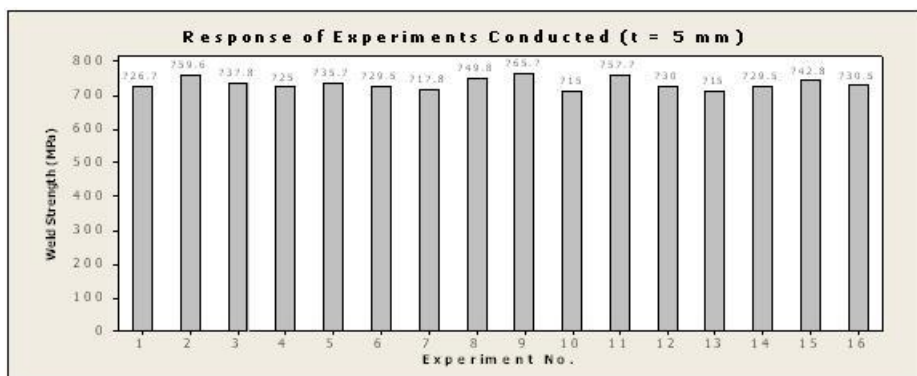


Figure 5. Weld Strength for Sixteen Experiments (t = 5 mm)

Analysis of Variance (ANOVA) performed on the experimental data suggested that the predictor variables can be arranged in the following order of decreasing significance of their effect on the response (weld strength):

1. Current
2. Weld Speed
3. Trailing
4. Voltage

For weld strength, all of the four predictor variables were found statistically significant. Further details of ANOVA applied to the aforementioned experimental data can be read from the reference [34].

The numerical optimization (using software “Design-Expert”) applied to the weld strength data suggests that for any sheet thickness value lying between 3 and 5mm (inclusive), the weld strength in TIG welding of HSLA steel can be maximized if the trailing is used along with low values of heat input i.e. low values of welding current and welding voltage and high value of welding speed. The predicted weld strength values are 783MPa, 772MPa and 762MPa for thickness 3mm, 4mm and 5mm, respectively at input combinations of: i) 170A, 10.5V, 18cm/min; ii) 200A, 10.5V, 18cm/min; and iii) 230A, 10.5V, 18cm/min, respectively.

Figures 6, 7, and 8 show the comparison of distortion (the maximum change in linear dimensions along any of the three axes) for the aforementioned sixteen tests as described in Tables 4, 5, and 6, respectively. The detailed mechanism for the measurement of distortion during welding of thin-walled shells can be studied from the reference [34]. The maximum and minimum values of distortion obtained with respect to thickness of the material are presented in Table 11.

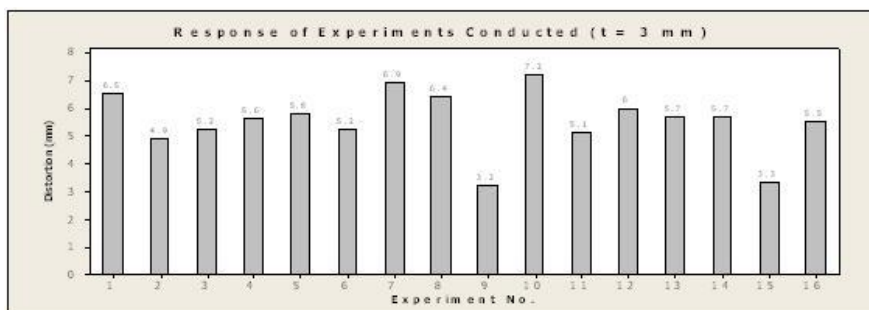


Figure 6. Distortion of Sixteen Experiments ($t = 3$ mm)

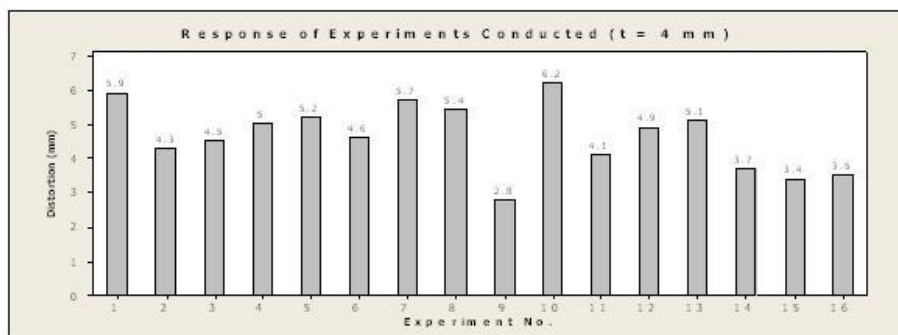
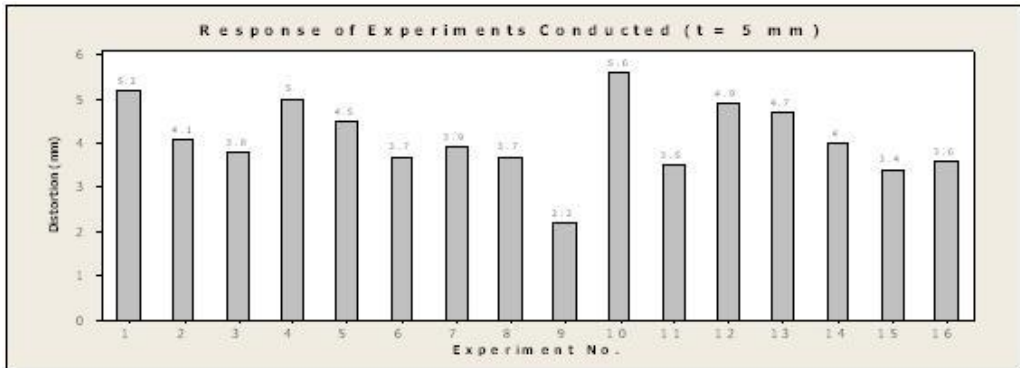


Figure 7. Distortion of Sixteen Experiments ($t = 4$ mm)

Table 11. Maximum and minimum values of Distortion

Thickness (mm)	Minimum (mm)	Maximum (mm)	Mean (mm)	Std. Deviation
3	3.2	7.2	5.512	1.09293
4	2.8	6.2	4.644	0.965
5	2.2	5.6	4.112	0.84

Figure 8. Distortion of Sixteen Experiments ($t = 5$ mm)

ANOVA performed on the experimental data suggested that the predictor variables can be arranged in the following order of decreasing significance of their effect on the response (distortion):

1. Weld Speed
2. Current
3. Voltage

The effect of the fourth predictor (Argon Trailing) on distortion was found statistically insignificant.

The numerical optimization applied to the distortion data suggests that for any sheet thickness value lying between 3 and 5mm, the distortion in TIG welding of HSLA steel can be minimized if the welding process is done at low values of welding current and welding voltage and high value of welding speed. The predicted weld distortion values are 3.7mm, 3.0mm and 2.7mm for thickness 3mm, 4mm and 5mm, respectively at input values of i) 170A, 10.5V, 18cm/min; ii) 200A, 10.5V, 18cm/min; and iii) 230A, 10.5V, 18cm/min, respectively.

Figures 9, 10, and 11 show the comparison of weld induced residual stresses (Von Mises) for the aforementioned sixteen tests as described in Tables 4, 5, and 6, respectively. The detailed mechanism for the measurement of the residual stresses during welding of thin-walled shells can be studied from the reference [34]. The maximum and minimum values of the residual stresses obtained with respect to thickness of the material have been presented in Table 11.

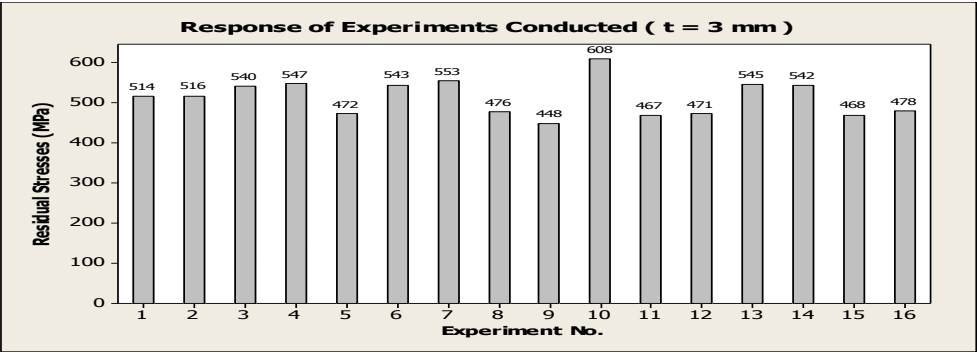


Figure 9. Residual Stresses of Sixteen Experiments (t = 3 mm)

Table 12. Maximum and minimum values of Residual Stresses (Von Mises)

Thickness (mm)	Minimum (MPa)	Maximum (MPa)	Mean (MPa)	Std. Deviation
3	448	608	511.75	44.471
4	366	505	425.25	39.211
5	335	452	384.12	32.087

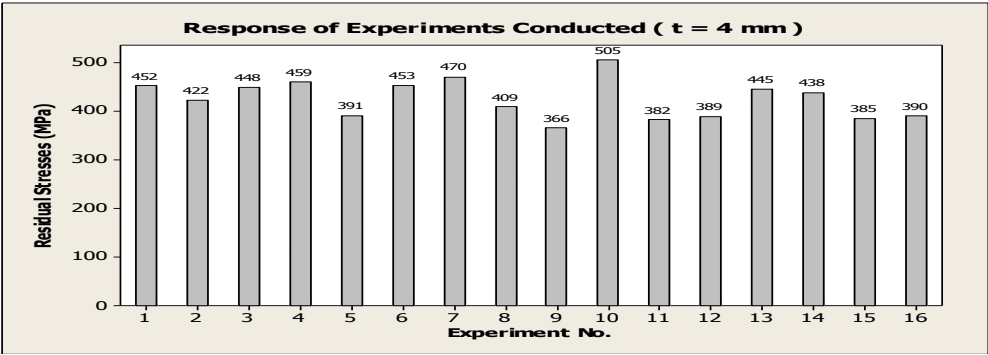


Figure 10. Residual Stresses of Sixteen Experiments (t = 4 mm)

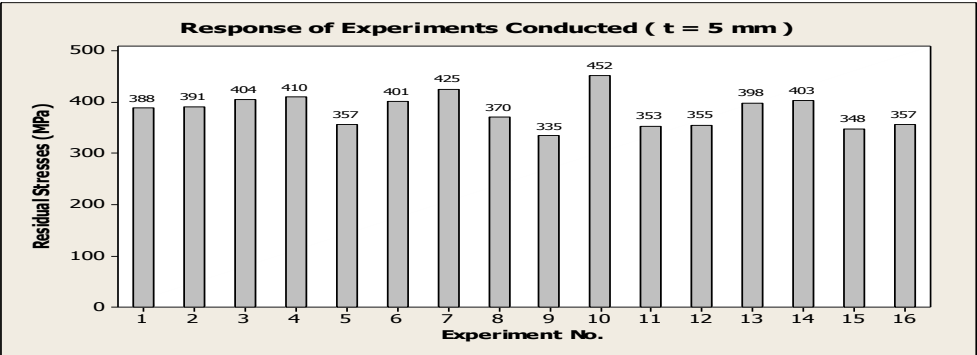


Figure 11. Residual Stresses of Sixteen Experiments (t = 5 mm)

ANOVA performed on the experimental data suggested that the predictor variables can be arranged in the following order of decreasing significance of their effect on the response (residual stresses):

1. Trailing
2. Voltage
3. Current
4. Weld Speed

For residual stresses, all of the four predictor variables were found statistically significant. Furthermore, the effect of Argon Trailing on residual stresses, as compared to the other two responses, was found extremely significant.

The numerical optimization applied to the residual stresses data suggests that for any sheet thickness value lying between 3 and 5mm, the residual stresses in TIG welding of HSLA steel can be minimized if the trailing is used along with low values of heat input i.e. low values of welding current and welding voltage and high value of welding speed. The predicted weld residual stresses values are 443MPa, 359MPa and 333MPa for thickness 3mm, 4mm and 5mm, respectively at input values of i) 170A, 10.5V, 18cm/min; ii) 200A, 10.5V, 18cm/min; and iii) 230A, 10.5V, 18cm/min, respectively.

2.2. Circumferential Welding of Shells

2.2.1. The predictor variables

Following is the list of significant TIG welding predictor variables with values that would be under study in the circumferential welding experiments to be performed on thin walled HSLA steel cylinders of different thicknesses (3, 4 and 5mm):

1. Welding Current (Amp) (170-270)
2. Welding Voltage (Volts) (10.5-13.5)
3. Welding Speed (cm/min) (15-18)
4. Argon Trailing (ON/OFF)

A 2^4 (4 factors, 2 levels, 16 test) full factorial design model (replicates 1, block 1, centre point per block 0 and order 4FI) was used for the circumferential welding experiments. Tables 13, 14, and 15 show the low and high settings (or levels) for the predictor variables used in sixteen tests for the cylinder thickness of 3, 4 and 5mm, respectively. Three of these predictor variables (welding current, welding voltage and welding speed) are numeric while the other one (gas trailing) is categorical.

Complete detail of 16 experiments following full factorial has been presented in Tables 16, 17, and 18 for cylinder thickness of 3, 4 and 5mm, respectively.

Table 13. High and Low Settings of Factors (Predictor Variables) [t = 3mm]

Factor	Name	Units	Type	Low Actual	High Actual
A	Current	A	Numeric	170.00	210.00
B	Voltage	V	Numeric	10.50	13.50
C	Weld Speed	cm/min	Numeric	15.00	18.00
D	Trailing		Categoric	nil	Ar

Table 14. High and Low Settings of Factors (Predictor Variables) [t = 4mm]

Factor	Name	Units	Type	Low Actual	High Actual
A	Current	A	Numeric	200.00	220.00
B	Voltage	V	Numeric	10.50	13.50
C	Weld Speed	cm/min	Numeric	15.00	18.00
D	Trailing		Categoric	nil	Ar

Table 15. High and Low Settings of Factors (Predictor Variables) [t = 5mm]

Factor	Name	Units	Type	Low Actual	High Actual
A	Current	A	Numeric	230.00	270.00
B	Voltage	V	Numeric	10.50	13.50
C	Weld Speed	cm/min	Numeric	15.00	18.00
D	Trailing		Categoric	nil	Ar

Table 16. Design of 16 Experiments following Full Factorial (cylinder thickness = 3 mm)

		Factor 1	Factor 2	Factor 3	Factor 4
Std	Run	A:Current	B:Voltage	C:Weld Speed	D: Trailing
		(A)	(V)	(cm/min)	
12	1	210.00	13.50	15.00	Ar
5	2	170.00	10.50	18.00	nil
1	3	170.00	10.50	15.00	nil
3	4	170.00	13.50	15.00	nil
11	5	170.00	13.50	15.00	Ar
7	6	170.00	13.50	18.00	nil
8	7	210.00	13.50	18.00	nil
16	8	210.00	13.50	18.00	Ar
13	9	170.00	10.50	18.00	Ar
4	10	210.00	13.50	15.00	nil
9	11	170.00	10.50	15.00	Ar
10	12	210.00	10.50	15.00	Ar
2	13	210.00	10.50	15.00	nil
6	14	210.00	10.50	18.00	nil
15	15	170.00	13.50	18.00	Ar
14	16	210.00	10.50	18.00	Ar

Table 17. Design of 16 Experiments following Full Factorial (cylinder thickness = 4 mm)

		Factor 1	Factor 2	Factor 3	Factor 4
Std	Run	A:Current	B:Voltage	C:Weld Speed	D: Trailing
		(A)	(V)	(cm/min)	
12	1	220.00	13.50	15.00	Ar
5	2	200.00	10.50	18.00	nil
1	3	200.00	10.50	15.00	nil
3	4	200.00	13.50	15.00	nil
11	5	200.00	13.50	15.00	Ar
7	6	200.00	13.50	18.00	nil
8	7	220.00	13.50	18.00	nil
16	8	220.00	13.50	18.00	Ar
13	9	200.00	10.50	18.00	Ar
4	10	220.00	13.50	15.00	nil
9	11	200.00	10.50	15.00	Ar
10	12	220.00	10.50	15.00	Ar
2	13	220.00	10.50	15.00	nil
6	14	220.00	10.50	18.00	nil
15	15	200.00	13.50	18.00	Ar
14	16	220.00	10.50	18.00	Ar

Table 18. Design of 16 Experiments following Full Factorial (cylinder thickness = 5 mm)

		Factor 1	Factor 2	Factor 3	Factor 4
Std	Run	A:Current	B:Voltage	C:Weld Speed	D: Trailing
		A	V	cm/min	
12	1	270.00	13.50	15.00	Ar
5	2	230.00	10.50	18.00	nil
1	3	230.00	10.50	15.00	nil
3	4	230.00	13.50	15.00	nil
11	5	230.00	13.50	15.00	Ar
7	6	230.00	13.50	18.00	nil
8	7	270.00	13.50	18.00	nil
16	8	270.00	13.50	18.00	Ar
13	9	230.00	10.50	18.00	Ar
4	10	270.00	13.50	15.00	nil
9	11	230.00	10.50	15.00	Ar
10	12	270.00	10.50	15.00	Ar
2	13	270.00	10.50	15.00	nil
6	14	270.00	10.50	18.00	nil
15	15	230.00	13.50	18.00	Ar
14	16	270.00	10.50	18.00	Ar

2.2.2. The response variables

Following response variables will be measured in order to judge the process performance of all the experiments designed for welding of shells of diameter and length equal to 300mm and thickness values of 3mm, 4mm, and 5mm.

1. Residual Stress (maximum value of weld-induced stresses [Von-Mises] in the weld zone) – to be measured in MPa.
2. Distortion (maximum value of weld-induced distortion in the cylinder in weld zone) – to be measured in mm.

The performance measure “Weld Strength” has not been included in the list because of the observation that response of this parameter to the aforementioned four predictor variables has been the same as that for the linear welding process.

The fixed parameters of the circumferential welding experiments are the same as that for the linear welding experiments.

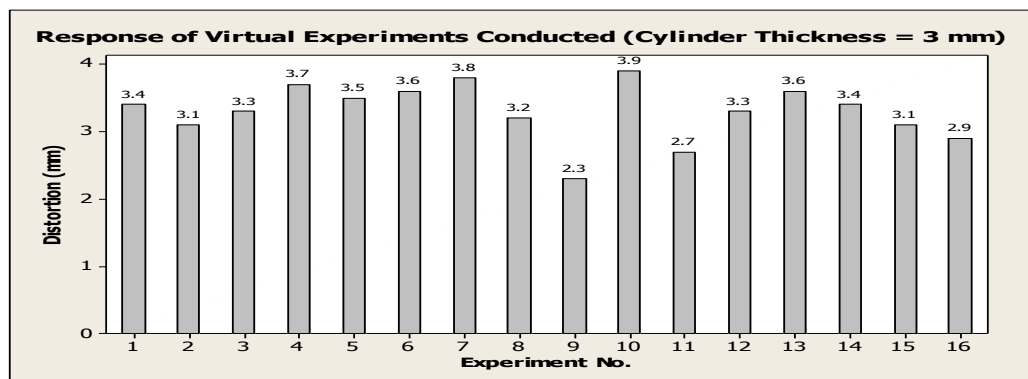


Figure 12. Distortion of Sixteen Experiments (cylinder thickness = 3 mm)

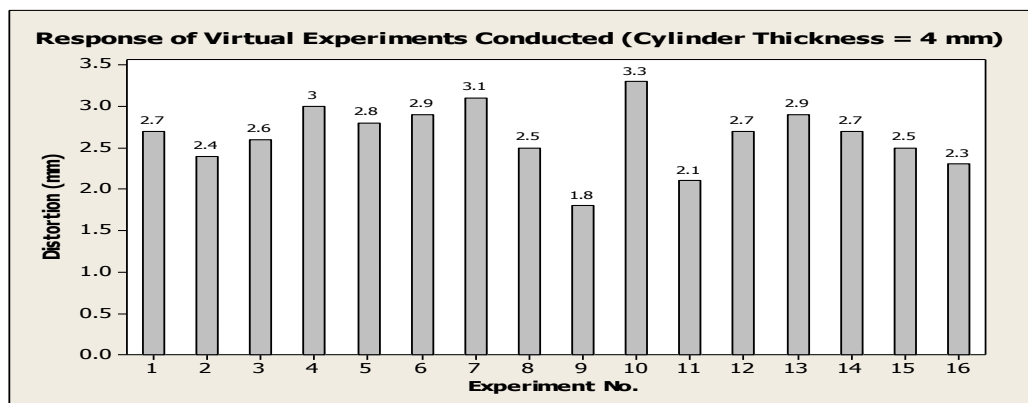


Figure 13. Distortion of Sixteen Experiments (cylinder thickness = 4 mm)

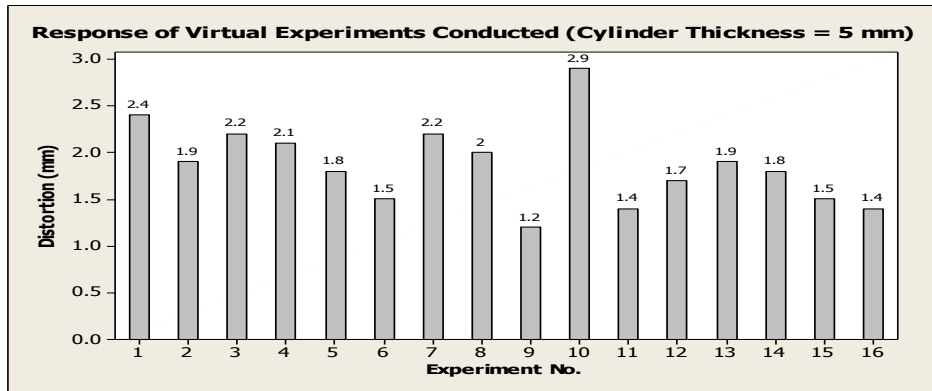


Figure 14. Distortion of Sixteen Experiments (cylinder thickness = 5 mm)

Table 19. Maximum and minimum values of Distortion

Thickness (mm)	Minimum (mm)	Maximum (mm)	Mean (mm)	Std. Deviation
3	2.3	3.9	3.3	0.417
4	1.8	3.3	2.64	0.379
5	1.2	3.2	1.87	0.434

2.2.3. Experimental results and analyses

Figures 12, 13, and 14 show the comparison of distortion for the aforementioned sixteen tests as described in Tables 16, 17, and 18, respectively. The detailed mechanism for the measurement of distortion during circumferential welding of thin-walled shells can be studied from the reference [34]. The maximum and minimum values of distortion obtained with respect to thickness of the material are presented in Table 19.

ANOVA performed on the experimental data suggested that the predictor variables can be arranged in the following order of decreasing significance of their effect on the response (distortion):

1. Trailing
2. Voltage
3. Current
4. Weld Speed

For distortion, all of the four predictor variables were found statistically significant.

The numerical optimization applied to the distortion data suggests that for any material thickness of cylinders value lying between 3 and 5mm, the distortion in TIG welding of HSLA steel can be minimized if the trailing is used along with low values of heat input i.e. low values of welding current and welding voltage and high value of welding speed. The predicted weld distortion values are 2.56mm, 1.96mm and 1.14mm for cylinder thickness 3mm, 4mm and 5mm, respectively, at input values of i) 170A, 10.5V, 18cm/min; ii) 200A, 10.5V, 18cm/min; and iii) 230A, 10.5V, 18cm/min, respectively.

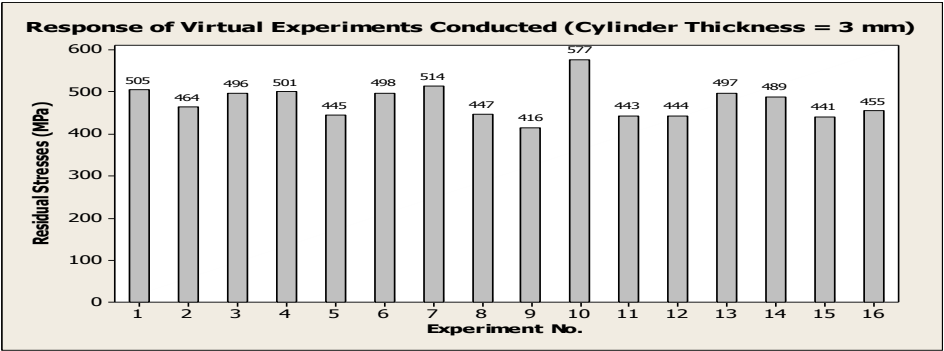


Figure 15. Residual Stresses of Sixteen Experiments (cylinder thickness = 3 mm)

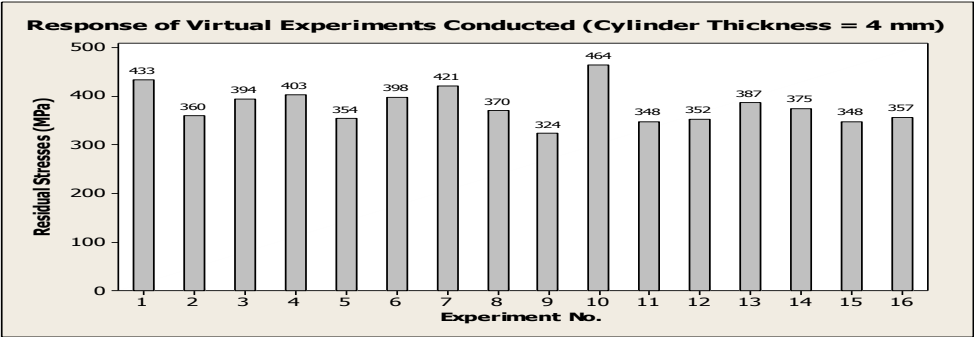


Figure 16. Residual Stresses of Sixteen Experiments (cylinder thickness = 4 mm)

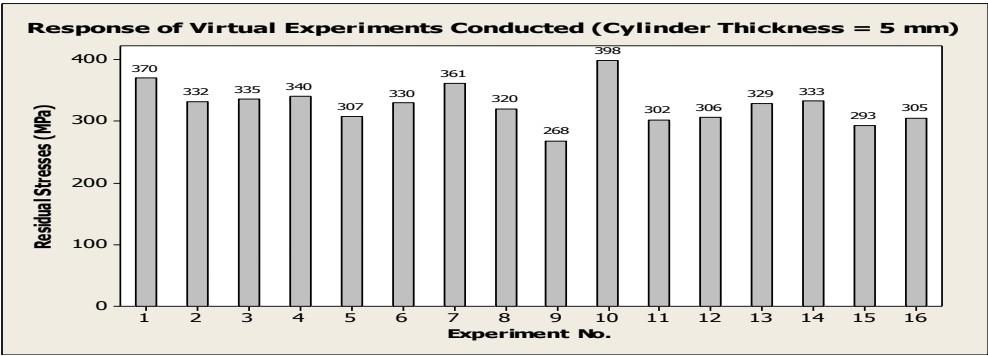


Figure 17. Residual Stresses of Sixteen Experiments (cylinder thickness = 5 mm)

Table 20. Maximum and minimum values of Residual Stresses (Von Mises)

Thickness (mm)	Minimum (MPa)	Maximum (MPa)	Mean (MPa)	Std. Deviation
3	416	577	477	39.9
4	324	464	380.5	36.8
5	268	398	326.8	31.61

Figures 15, 16, and 17 show the comparison of residual stresses for the aforementioned sixteen tests as described in Tables 16, 17, and 18, respectively. The maximum and minimum values of residual stresses obtained with respect to thickness of the material have been presented in Table 20.

ANOVA performed on the experimental data suggested that the predictor variables can be arranged in the following order of decreasing significance of their effect on the response (residual stresses):

1. Trailing
2. Current
3. Voltage
4. Weld Speed

For residual stresses, the effects of all the four predictor variables were found statistically significant.

The numerical optimization applied to the residual stresses data suggests that for any material cylinder thickness value lying between 3 and 5mm, the residual stresses in TIG welding of HSLA steel can be minimized if the trailing is not used along with low values of heat input i.e. low values of welding current and welding voltage and high value of welding speed. The predicted weld residual stresses values are 410MPa, 316MPa and 273MPa for cylinder thickness 3mm, 4mm and 5mm, respectively at input values of i) 170A, 10.5V, 18cm/min; ii) 200A, 10.5V, 18cm/min; and iii) 230A, 10.5V, 18cm/min, respectively.

3. KNOWLEDGE-BASED SYSTEM FOR OPTIMIZING TIG WELDING PROCESS

After completion of all welding analyses required to obtain the data by experimental work and statistical analyses related to optimization of welding process, the next process is to manage the available welding experimental data and optimization information at a single platform and to utilize some automated means to extract the useful information from that platform in most effective manner as knowledge. The selection of expert system is the best option for this requirement. Furthermore, the relationship among welding parameters and response is complex and it is very difficult to represent it using some mathematical model. In the following sections, the objectives of developing expert system and application to welding; the configuration; the utilization of fuzzy logic for reasoning mechanism; and the optimal formation of rule-base of the expert system are presented.

3.2. The Objectives of Expert System and Application to Welding

The expert systems are computer programs that embody narrow domain knowledge for problem solving related to that knowledge domain [35]. Generally, an expert system comprises of following three main elements: a knowledge base, an inference engine, and working memory. The knowledge base is a collection of knowledge which is expressed by

using some formal knowledge representation language, normally in form of facts and IF-THEN rules. Whereas, the inference engine is a generic control mechanism that uses the axiomatic knowledge present in the knowledge base to the task oriented data to reach at a conclusion. Furthermore, a program that contains meta-knowledge is called inference engine. Usually a knowledge base is very large, therefore, it is necessary requirement to have inference mechanisms that searches through the database and deduces results in a systematic and organized way [36].

During the execution of the expert system, the working memory is used to temporarily store the values of variables. The main components of an expert system are shown in Figure 18. The knowledge is explicitly kept separate from the control module in expert systems, while it is intertwined with the control mechanism in conventional programs. In this way, the expert system programs are better than conventional programs. It is very easy to add new knowledge in expert system due to the separation of knowledge from the control module during the expert system development phase or by experience of the program throughout use in its lifetime. This feature of mechanism mimics the human brain in which the control processes remain unchanged although individual behavior is continuously changing by addition of new knowledge by experience. This is the main feature that enables the expert system an ideal computer-based replacement of a human expert in the related domain.

The main objective of the research carried out in the welding domain and described in this chapter is the optimal settings of the welding process input parameters so as to maximize the weld strength and minimize the residual stresses and distortion without compromising the welding quality. The highly generalized information generated by the experimental work is very difficult to be utilized by the welder, operator, or engineers for solution of their highly specific welding problems.

In short, there is a dire need of a fast-acting informative tool that can recommend the optimal settings of the selected welding process parameters that would lead to accomplishment of desired objectives in best possible and efficient manner. Furthermore, the tool should also be capable of providing highly accurate predictions of the performance measures before the start of the actual process at shop floor. The expert system developed and presented in this chapter fulfils all these requirements and provides the highly specific information to the user at the expense of few seconds.

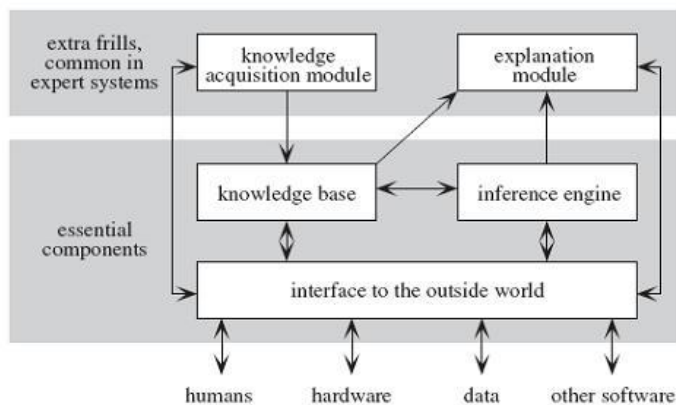


Figure 18. Main Components of an Expert System [21]

3.3. The Expert System Configuration

To cover the requirements of the current research work, the information available from the experimental data, ANOVA results and numerical optimization was used for the development of knowledge-base (or the rule-base). The presented expert system is dual functional as first it searches for the optimal selection and combination of the significant predictor variables in order to satisfy the desired objective; and secondly, it provides the predicted values of performance measures or responses for the selected combination of predictor variables or input parameters.

First consider the selection of five predictor variables only for the purpose of simplicity in description. These five predictor variables are the ones that were tested in the set of experiments explained in previous section, i.e. material thickness, welding current, welding voltage, welding speed, and choice of trailing. The shell thickness will be considered as a parameter that needs not to be optimized. This is so because thickness is the geometric property of the work piece and it cannot be changed unless the work piece is removed from the welding setup and changed. The configuration of the proposed expert system is shown in Figure 19.

3.3.1. Optimization and prediction modules

The knowledge-base consists of two sets of rules, each one of them being controlled and operated by a separate module as shown in Figure 19. The optimization module is the first one that takes charge and operates with relevant set of rules for the optimal selection and combination of four parameters (predictor variables): the welding current, welding voltage, welding speed and the trailing. The selection of the predictor parameters is made in accordance with the objective desired by the user, the material thickness provided, and the predictor variables pre-fixed by the user. After this, the prediction module takes charge and makes use of the finalized combination of predictor variables and the relevant set of rules in order to estimate the values of performance measures, i.e. weld strength, distortion, and weld induced residual stresses.

3.3.2. Expert system shell

As shown in Figure 19, it can be seen that the expert system shell consists of the user interface through which the input is taken from the user. The data fuzzifier fuzzifies the values of numeric parameters (predictor variables) according to the relevant fuzzy templates. The expert system shell also contains the working memory that consists of different variables, while the data defuzzifier is used to defuzzify the fuzzy sets of predictor variables (welding current, voltage, speed) and of performance measures.

As the expert system presented is a kind of production system that requires the control of forward-chaining inference mechanism for the extraction of conclusions from its knowledge-base, according to the set of asserted facts and rules. For this purpose, a forward-chaining expert system shell named Fuzzy CLIPS (Fuzzy extension of C Language Integrated Production Systems) – developed by National Research Council, Canada – was utilized for the development of this knowledge based system [36]. Fuzzy CLIPS provides its standard format for defining templates, facts, functions, rules, and modules, and whole of the knowledge-base is the combination of these elements.

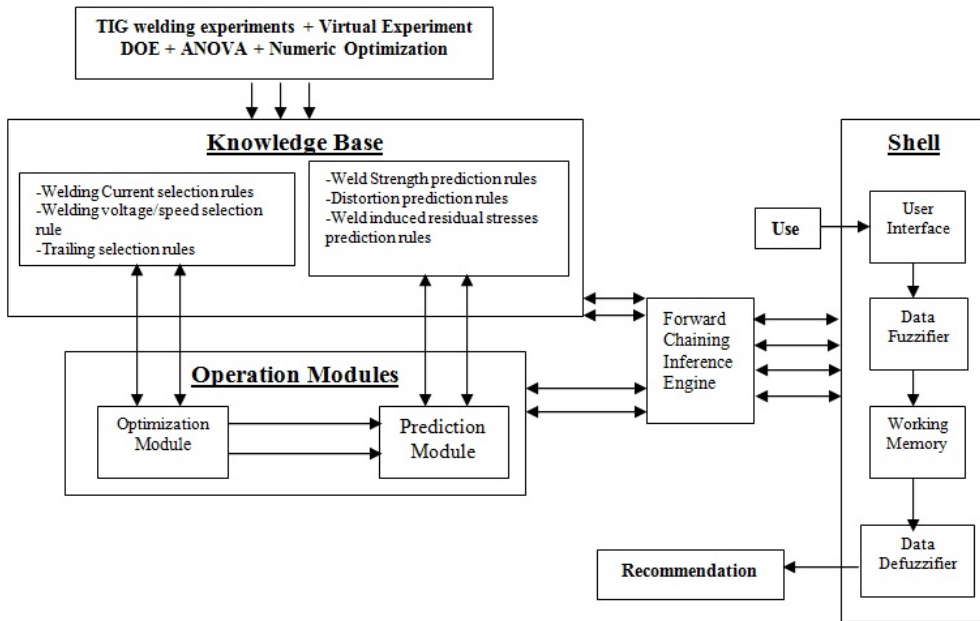


Figure 19. Configuration of the Expert System

3.3.3. The procedure

The flow chart of operating procedure of the expert system is shown in the Figure 20. The expert system process starts with the user's input of desired objective and the values of predictor parameters. It is mandatory for the user to fix the objective as well as the value of shell thickness, while the values of other four variables may or may not be fixed according to the welding problem-on-hand. The user may choose from the following three objectives:

1. Maximize weld strength.
2. Minimize residual stresses or distortion.
3. Achieve 1 and 2 simultaneously.

The selection of one objective as given above will lead to recommendation of different values of process inputs or predictor variables as compared to those of other, and consequentially, it will also lead to prediction of different values of the performance measures as per requirements of maximization or minimization. The objective number 3 provides the trade-off between the first two objectives. The values of material thickness and welding current (if fixed by user) are fuzzified according to the relevant fuzzy templates.

As the welding current has been proved, by ANOVA results, to be the most significant factor for weld strength / distortion / residual stresses, this factor is ought to be fixed ahead of others, if not already fixed by the user. The other three variables are also fixed in similar fashion.

After the fixation of predictor variables as mentioned above, the prediction module takes the charge and the values of three response variables, in accordance with the finalized values of predictor variables, are estimated simultaneously. The next step is data defuzzification, in which the fuzzy values of welding current, weld strength, distortion and residual stresses are

defuzzified in accordance with preferred defuzzification algorithm. Finally, in the last step, the recommendation of predictor variables and prediction of response variables are printed out.

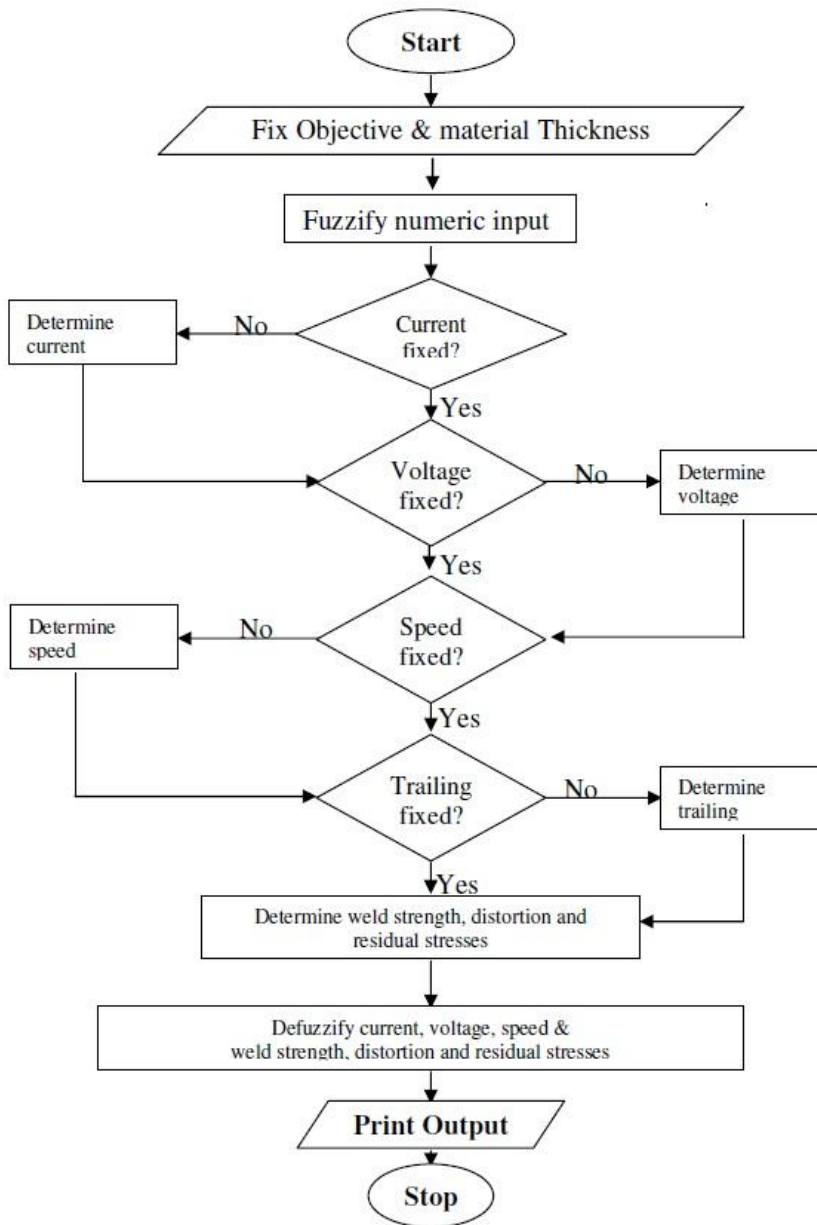


Figure 20. Flow chart representing the operational procedure of the expert system

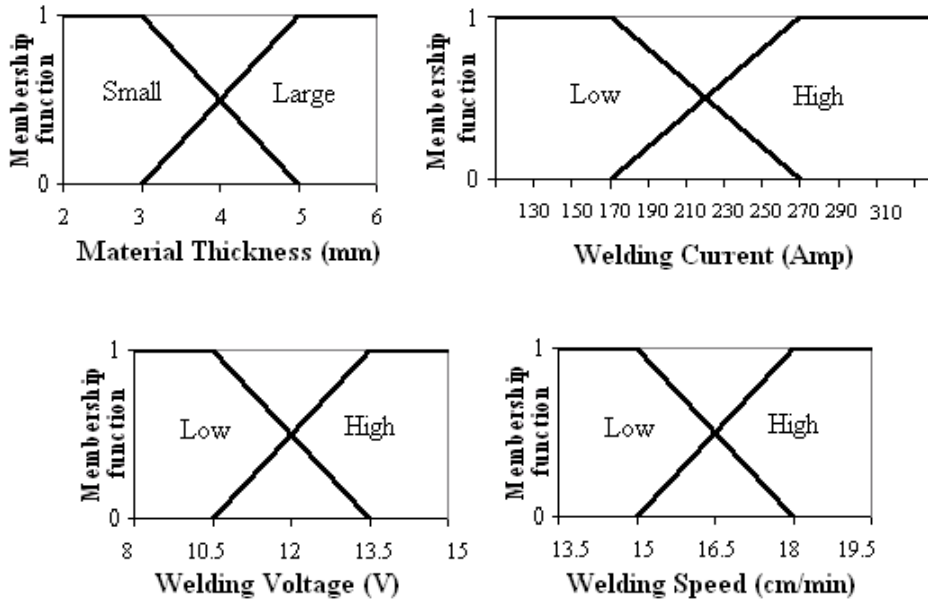


Figure 21. Fuzzy sets for the numeric input variables

3.4. Fuzzy Reasoning for the Expert System

The fuzzy logic is a discipline that has been successfully used in automated reasoning of expert systems [19]. In the real world system, there are some problems found in relationships between inputs and outputs like uncertainty, vagueness, ambiguity and impreciseness. These input and output relationship problems can be handled effectively by utilizing fuzzy logic treatments.

3.4.1. Fuzzy sets, Input Fuzzification, and Output Defuzzification

In the fuzzification, the precise or imprecise input data which are easily understandable by the human minds are converted into a kind of linguistic form, for example very low (weld strength) and highly distort (distortion) etc. The expert system then uses these fuzzified data to give answers to imprecise and vague questions and also describe the reality level of those answers. Figure 21 shows the fuzzy sets utilized for four predictor variables: material (sheet or cylinder) thickness, welding current, welding voltage and welding speed; while Figure 22 shows fuzzy sets for responses (weld strength, distortion and residual stresses).

Triangular shaped fuzzy sets for the response variables in Fuzzy CLIPS format are as follows:

```
(deftemplate Weld_Strength 680 800 MPa
  ( (very low (680 1) (700 1) (720 0) ) (low (700 0) (720 1) (740 0) )
    (medium (720 0) (740 1) (760 0) ) (high (740 0) (760 1) (780 0) )
    (very high (760 0) (780 1) (800 1) ) ) )
```



```
(deftemplate Distortion 1 7 mm
  ( (vlow (1 1) (2 1) (3 0) ); very low
    (low (2 0) (3 1) (4 0) ); low
    (med (3 0) (4 1) (5 0) ); medium
    (high (4 0) (5 1) (6 0) ); high
    (vhigh (5 0) (6 1) (7 1) ) ); very high
```

```
(deftemplate Residual_Stresses 100 700 MPa
  ( (vlow (100 1) (200 1) (300 0) ); very low
    (low (200 0) (300 1) (400 0) ); low
    (med (300 0) (400 1) (500 0) ); medium
    (high (400 0) (500 1) (600 0) ); high
    (vhigh (500 0) (600 1) (700 1) ) ); very high
```

The one predictor variable (trailing) is categorical, therefore, cannot be fuzzified. However, this variable is used as crisp variable in the fuzzy knowledge-base. This shows that in this expert system development, both crisp and fuzzy antecedents and consequents are freely mixed for the creation of the rules. The fuzzy rule application step provides the recommendation as a crisp value and/or fuzzy set, specifying a fuzzy distribution of a conclusion. But in welding process, the operator or welder needs a single discrete valued direction. Therefore, it is required to select a single point from fuzzy distribution that provides the best value. The process of reducing a fuzzy set to a single point is known as defuzzification [36]. There are two methods commonly used for defuzzifying the fuzzy sets i.e. center of gravity (CoG) or moment method and mean of maxima (MoM) method. The detail of both methods can be referred in [36, 37]. For this expert system development, the centre of gravity (CoG) method is used as defuzzification method for the reason that it provides smoothly varying output of response variables for gradually varying input values of material thickness, welding current, and voltage. Whereas the utilization of mean of maxima (MoM) method contained the risk of generating highly abrupt output values of response variables for small and gradual variations in material thickness, welding voltage, and welding current values that was observable at specific ranges of these two predictor variables.

3.4.2. Inference for aggregation of fuzzy rules

Generally, two kinds of methods are commonly used for yielding aggregation of fuzzy rules i.e. max-min inference method and max-product method. The max-min inference method is the default inference method for Fuzzy CLIPS. The application of max-min inference strategy is described in the following example.

Suppose a knowledge-base consists of following set of rules:

1. IF thickness is **Small** AND current is **Low** THEN weld strength is **Low**
2. IF thickness is **Small** AND current is **High** THEN weld strength is **Medium**
3. IF thickness is **Large** AND current is **Low** THEN weld strength is **Very Low**
4. IF thickness is **Large** AND current is **High** THEN weld strength is **Low**

Further suppose that it is required to predict the value of weld strength for work piece material thickness of 4.5mm and welding current of 190A, utilizing above-mentioned set of 4

rules and fuzzy sets provided in Figures 21 and 22. Figure 23 describes the input fuzzification process in which the welding current value of 190A has been converted to 2 fuzzy sets: Low (membership function $\mu_{Low} = 0.8$) and High ($\mu_{High} = 0.2$); while the material thickness of 4.5mm has also been converted into 2 fuzzy sets: Large ($\mu_{Large} = 0.75$) and Small ($\mu_{Small} = 0.25$). The fuzzy membership value for welding current can be expressed as: $\mu(\text{current}) = 0.8/\text{Low}, 0.2/\text{High}$. Similarly the fuzzy membership value for material thickness can be expressed as: $\mu(\text{thickness}) = 0.75/\text{Large}, 0.25/\text{Small}$.

All the four rules use AND operator in their antecedent parts. Considering the first rule in the list and applying the max-min strategy, the rule will yield a result (i.e. weld strength is Low) whose degree (or membership function) will be minimum of degrees of current (Low) and of thickness (Small). This can be expressed as follows:

$$\mu(\text{weld strength } Low) = \min \{ \mu(\text{current } Low), \mu(\text{thickness } Small) \}$$

$$\mu\left\{\frac{0.25}{Low}\right\} = \min \left\{ \mu\left(\frac{0.8}{Low}\right), \mu\left(\frac{0.25}{Small}\right) \right\}$$

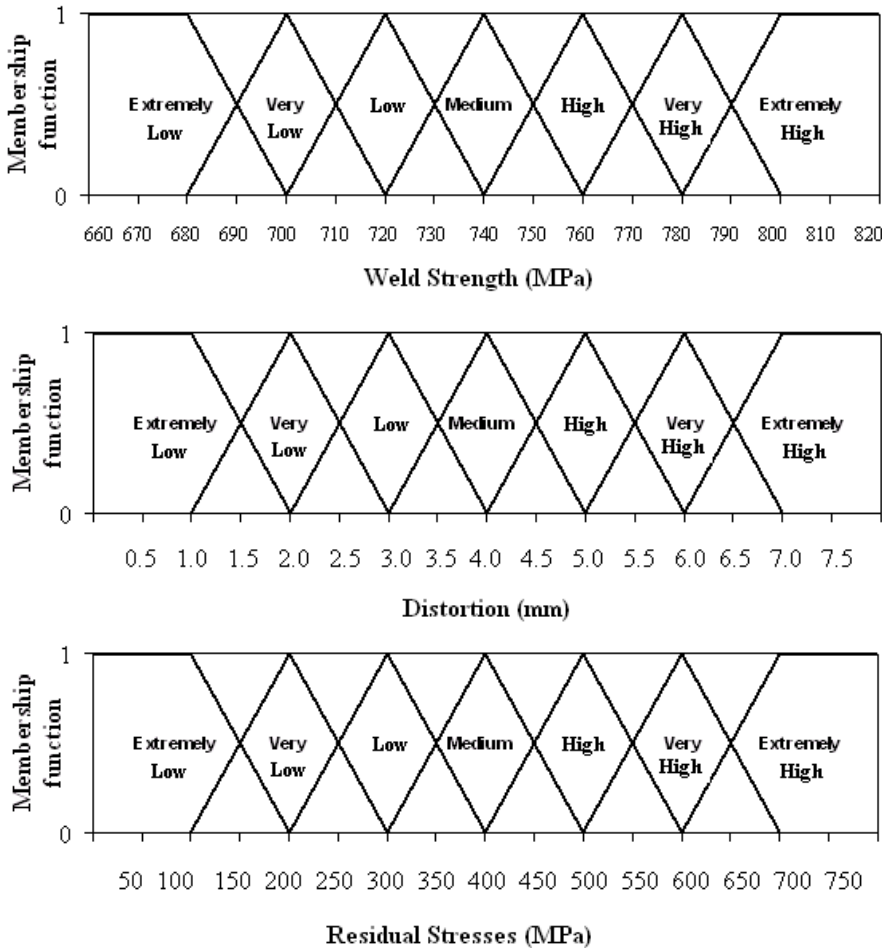


Figure 22. Fuzzy sets for the Responses

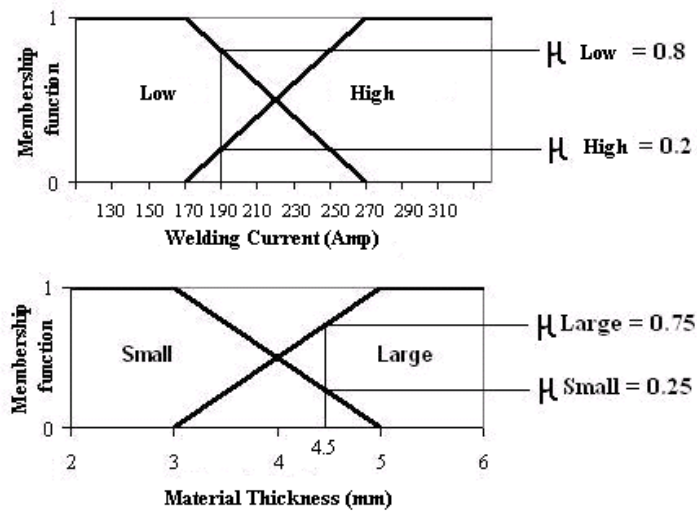


Figure 23. Fuzzification of input data

Using all possible combinations of two inputs and applying AND operation, we can have following fuzzy membership values for output variable weld strength (considering application of above listed four rules):

1. Low (0.8) and Small (0.25) will yield Low (0.25)
2. Low (0.8) and Large (0.75) will yield Medium (0.75)
3. High (0.2) and Small (0.25) will yield Very Low (0.2)
4. High (0.2) and Large (0.75) will yield Low (0.2)

Following the procedure of aggregation, in accordance with max-min strategy, Table 21 can be obtained, in association with the four rules. Applying OR operation to all fuzzy set values in Table 21 will yield the maximum value for the output fuzzy set, which is shown in Table 22. Defuzzified output which gives the value of weld strength can be obtained as follows:

$$\text{Weld Strength} = \frac{690 \times 0.2 + 700 \times 0.2 + 710 \times 0.25 + 720 \times 0.25 + 730 \times 0.5 + 740 \times 0.75 + 750 \times 0.5 + 760 \times 0.2}{0.2 + 0.2 + 0.25 + 0.25 + 0.5 + 0.75 + 0.5 + 0}$$

$$\text{Weld Strength} = 728.49 \text{ MPa}$$

Table 21. Weld Strength values from all the four rules

Fuzzy	Universe of weld strength (MPa)									
Subsets	670	680	690	700	710	720	730	740	750	760
Low	0	0	0	0	0.25	0.25	0.25	0	0	0
Medium	0	0	0	0	0	0	0.5	0.75	0.5	0
Very Low	0	0	0.2	0.2	0.2	0	0	0	0	0
Low	0	0	0	0	0.2	0.2	0.2	0	0	0

Table 22. Maximum fuzzy output from Table 21

670	680	690	700	710	720	730	740	750	760	770	780	790
0	0	0.2	0.2	0.25	0.25	0.5	0.75	0.5	0	0	0	0

3.5. Optimal Formation of Fuzzy Rule-Base

The relationship between inputs and output in a fuzzy system is characterized by set of linguistic statements which are called fuzzy rules [38]. The collection of rules is called rule-base and the combining the rule-base with list of facts is termed as knowledge-base. The number of fuzzy rules in a fuzzy system is related to the number of fuzzy sets for each input variable.

For the present case, there are two fuzzy sets each for material thickness, welding current, voltage and welding speed. Similarly, there are two possible values each for trailing (nil and Ar). However, for four variables (material thickness, welding current, welding speed, and trailing), the maximum possible number of rules for the prediction module of the expert system are 16 ($= 2 \times 2 \times 2 \times 2$). An important question arises here, “which weld strength sets, or distortion sets to be assigned to 16 possible combinations of input sets/values”? For a simple 2-inputs 1-output fuzzy model, the designer has to select the most optimum set of fuzzy rules from more than 10,000 combinations [38]. For the output variable of weld strength in the present case, there are 16 fuzzy rules with 7 possibilities each (7 fuzzy sets for weld strength). Thus, the total number of possible fuzzy rules combination will be $7^{16} = 3.323 \times 10^{13}$ for purpose of estimation of weld strength. Similarly, there are more possibilities for formulation of fuzzy rules for estimation of distortion and residual stresses. For the best possible combination of rules, the simulated annealing algorithm has been employed for assigning the most optimum fuzzy set of each of output variables to the 16 rules. The objective of rule-base optimization process is to minimize the estimation error (i.e., difference between predicted values of the output variable and its actual values).

3.5.1. Optimal formation using simulated annealing algorithm

Simulated annealing (SA) is a stochastic neighborhood search method, which is developed for combinatorial optimization problems [39]. It is based on the analogy between the process of annealing of solids and solution methodology of combinatorial optimization problems. It has capability of jumping out of local optima for global optimization. This capability is achieved by accepting with a probability the neighboring solutions worse than the current solution. The acceptance probability is determined by a control parameter “temperature”, which decreases during SA process. The details of SA can be found in [39]. The pseudo-code of the algorithm developed for optimization of fuzzy rules using SA technique is given in the following [40]:

[0] Initialize

[0.1] Set annealing parameters T_0 , $AT_{\min}$, $i_{\max}$, α , R_f

[0.2] Initialize iteration counter, $i = 0$

- [0.3] Generate initial rules combination and calculate estimation error value, i.e. rules [0], error [0]
- [1] Execute outer loop, i.e. steps 1.1 to 1.7 until conditions in step 1.7 are met.
- [1.1] Initialize inner loop counter $l = 0$, and accepted number of transitions $AT = 0$
- [1.2] Initialize rules combination for inner loop, rules $[i][0] = \text{rules } [i]$ and error $[i][0] = \text{error } [i]$
- [1.3] Execute inner loop, i.e. steps 1.3.1 to 1.3.5 until conditions in step 1.3.5 are met
- [1.3.1] Update $l = l + 1$
- [1.3.2] Generate a neighboring solution by changing randomly one rule, and compute estimation error for new rules combination (rules $[i][l]$ and error $[i][l]$)
- [1.3.3] Assign $q = \text{error } [i][l] - \text{error } [i][l - 1]$
- [1.3.4] If $q \leq 0$ or $\text{Random}(0, 1) \leq e^{-q/T_0}$ then
Accept rules $[i][l]$ and error $[i][l]$
Update $AT = AT + 1$
- Else* reject generated combination: rules $[i][l] = \text{rules } [i][l - 1]$, error $[i][l] = \text{error } [i][l - 1]$
- [1.3.5] If one of following conditions hold true: $AT \geq AT_{\min}$; OR $l \geq 5S^2$ (S – No. of fuzzy sets of output variable), then assign length of Markov chain $L[i] = l$. Terminate inner loop and go to 1.4, *else* continue the inner loop and go to 1.3.1
- [1.4] Update $i = i + 1$
- [1.5] Update: rules $[i] = \text{rules } [i - 1][L[i] - 1]$ and error $[i - 1][L[i] - 1]$
- [1.6] Reduce cooling temperature: $T[i] = \alpha \cdot T[i - 1]$
- [1.7] If one of following conditions hold true: $i \geq i_{\max}$; OR $(AT / L[i]) \leq R_f$; OR estimation error value does not reduce for last 20 iterations, then terminate the outer loop and go to 2, *else* continue outer loop and go to 1.1

[2] Print out the best rules combination along with minimum estimation error value and terminate the procedure

C++ was used to code the algorithm. The SA parameters were operated using following values: (1) starting annealing temperature (T_0) = 1300MPa; (2) rate of cooling (α) = 0.98; (3) maximum number of iterations ($i_{\max}$) = 100; (4) length of Markov Chain at each iteration (L) = $5 \times 7 \times 7 = 245$; (5) minimum acceptance ratio (R_f) = 0.01; (6) minimum number of accepted transitions at each iteration ($AT_{\min}$) = 100.

The objective function of the “optimization of fuzzy rules” problem is the *minimization of estimation error*, where the term “estimation error” can be defined as follows:

$$\text{Estimation\_error} = \frac{1}{l \times m \times n \times o} \left[\sum_{i=1}^l \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^o |WS * - WS^{est}| \right] \quad (1)$$

For equation (1): l, m, n, o = Number of levels (not the fuzzy sets) provided by the user for each of the four variables: material thickness, welding current, welding speed and trailing, respectively.

WS^* = Actual weld strength

WS^{est} = Weld strength estimated by the rule-base.

3.5.2. Results of optimal formation of rule-base

The optimal formation of fuzzy rule-base related to prediction of weld strength only has been presented. In the similar way, the rule-bases for prediction of other output variables can be optimized. Furthermore, the optimization of rules related to the optimizing module of the expert system is not in the scope of this section.

Initially, the random combination of fuzzy rules was made and the criteria of termination for algorithm depended upon fulfillment of one of three conditions provided in the algorithm pseudo-code. For estimation error, each transition of the rules of all iterations was tested in order to determine the optimal combination of fuzzy rules by using the data provided in the previous sections. The program continued processing for 30 iterations based upon SA algorithm until the criteria of termination was fulfilled, as the estimation error value did not improve for last 20 iterations.

The optimal combination of fuzzy rules was printed out at the termination of program as listed in Table 24 and the testing values of input variables resulted in least value of estimation error, i.e. 5MPa. Figure 24 shows the continuous improvement in estimation error through the iterations of this program run. The optimized rules for prediction of other output variables are listed in Table 25.

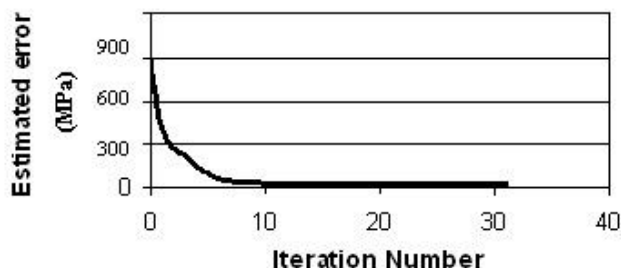


Figure 24. Decline of estimation error along number of iterations

Table 23. List of rules operated by the optimization module

Rule No.	Antecedents Objective	Consequent Thickness	Speed	Current	Trailing	
1	Any ¹	Any	Open ²	Any	Any	Speed Low
2	WS ³ or Both ⁴	Any	Any	Open	Any	Current High
3	Dist ⁵	Large	Any	Open	Ar or Open	Current Low & ⁶ High
4	Dist	Large	Any	Open	Nil	Current High
5	Dist	Small	Any	Open	Any	Current Low
6	WS	Any	Any	Any	Open	Trailing Nil
7	Dist or Both	Any	Any	Any	Open	Trailing Ar

¹Fixed with any level of the variable; ²Not fixed; ³Maximize weld strength;

⁴Achieve 1 & 2 simultaneously; ⁵Minimize distortion; ⁶Intersection operator

**Table 24. List of rules operated by the prediction module
[Consequents: Weld Strength & Distortion]**

Rule	Antecedents	Consequents					
No.	Thickness	Current	Speed	Trailing	Weld Strength		Distortion
1	Small	High	Low	Nil	very low	high & very high	
2	Small	High	Low	Ar	very low & low	medium	
3	Small	High	High	Nil	medium	low & medium	
4	Small	High	High	Ar	medium & high	very low	
5	Small	Low	Low	Nil	very low & low	low	
6	Small	Low	Low	Ar	low	medium	
7	Small	Low	High	Nil	high	low	
8	Small	Low	High	Ar	high & very high	very low	
9	Large	High	Low	Nil	very low	high	
10	Large	High	Low	Ar	low	low & medium	
11	Large	High	High	Nil	medium	low	
12	Large	High	High	Ar	high	very low	
13	Large	Low	Low	Nil	extremely low	low	
14	Large	Low	Low	Ar	extremely low	very low & low	
15	Large	Low	High	Nil	very low	low	
16	Large	Low	High	Ar	low & medium	very low	

3.5.3. The complete rule-base

In this sub-section, all the rules operated by the optimization module as well as the prediction module are listed. As the target of optimization module is to select the values of predictor variables (welding current, voltage, welding speed and trailing), which will best satisfy the desired objective, so all of the possible values of these variables (fuzzy or crisp) do not appear in the consequent parts of the optimization rules. However, the welding experiments and ANOVA results have shown that these non-appearing values of the variables do not satisfy any of three objectives in any combination of predictor variables. The complete list of rules operated by optimization module is given in Table 23.

Whereas the prediction module is assigned to generate the best possible estimate of all the three response variables for any given combination of four predictor variables whether all of the four predictor variables have been fixed by the user or any combination of these has been determined by the optimization module. Table 24 enlists these 16 rules with two consequents displayed: weld strength and distortion. Table 25 displays the other consequents – residual stresses – for the same 16 rules arranged in same order as in Table 24. Table 24 and 25 show the rules that were developed by Simulated Annealing Algorithm for maximum precision in predicting the values of output variables.

Table 25. List of consequents (Residual stresses) for the antecedents enlisted in table 24 (See the fuzzy sets provided in sub-section 3.4.1)

Rule No.	Residual Stresses
1	high & very high
2	medium
3	low & medium
4	very low
5	low
6	medium
7	low
8	very low
9	high
10	low& medium
11	low
12	very low
13	low
14	very low & low
15	low
16	very low

3.6. Application Example

Consider the application of the presented fuzzy expert system for optimization of parameters and prediction of performance measures in TIG welding process. Suppose it is required to find optimal values of welding current, voltage and welding speed in order to attain lowest possible distortion, when HSLA steel plates of thickness 5mm, is to be welded with Ar trailing. It is also desired to have prediction of weld strength, distortion, and residual stresses for the recommended welding conditions.

For this case, the user provides following input to the expert system: objective as ‘minimize distortion’; material thickness as 5mm; and trailing as Ar. After processing, the expert system prints out the following recommendations and predictions:

It is recommended to use welding current of 230A.

It is recommended to use welding voltage of 10.5V.

It is recommended to use welding speed of 18cm/min.

It is predicted that weld strength will be 765.7MPa.

It is predicted that distortion of plates will be 2.2mm.

It is predicted that weld induced residual stresses will be 335MPa.

4. HIGH LEVEL AUTOMATION FOR DEVELOPING KNOWLEDGE-BASED SYSTEM

The knowledge-based system (expert system) developed, in the previous section, consumed a considerable amount of effort and time but still its scope remained limited. It covered the effects of only four input parameters at the expense of formulation of 23 rules, 16 of them employing 3 output variables and also needed to be optimized using a cumbersome optimization algorithm. In order to expand the scope of the system, the developer would have to redo the same hectic efforts in order to incorporate the incoming knowledge from experimental work on welding in knowledge-base. Such type of requirement and situation represents a picture of a major barrier in the way of successful application of knowledge-based systems at industrial level. In this way, there is strong need to have a computer-based consultation system that can develop and expand its scope of application by itself without requiring knowledge engineering skills of the developers.

4.1. Self-Development of Knowledge-Based System

Only few research papers are available that have focused and described the ability of self-learning imparted to the knowledge-based systems. In broad aspect, the self learning field is called as machine learning in which the computer programs learn from their own experience upon utilization. A self-learning and self-testing fuzzy expert system applicable to control system was presented in [41]. The main feature of the expert system provided is to check the completeness and correction of the knowledge-base. The program was developed based upon the results of actions it performs in such a manner that the system extracts fuzzy rules from the set of input-output data pairs and keeps on correcting its rules. However, the paper does not cover the idea of expanding the scope or applicability of the expert system. In [42], the author presented a general framework for acquisition of knowledge using inductive learning algorithm and genetic algorithm.

In manufacturing, a few papers can be found that describe the application of machine learning to the field of metal cutting only. In [43], the authors presented a machine learning approach for building the knowledge-base from the numerical data and proved to be useful for classifying the dielectric fluids in Electric Discharge Machining. In [44], the author presented partially the application of pattern recognition and ANN for acquiring the knowledge in order to monitor the condition of tool in a plate machining process. In [45], the authors presented the use of ANN for picking up the experience of machinists and data from the machining handbook to predict the values of cutting speed and feed for a given turning process. In [46], the authors presented the utilization of Support Vector Regression, a statistical learning technique, to diagnose the condition of tool during a milling process. Now it is obvious that machine learning approach has been utilized on a very limited scale for optimization of few process parameters or for the purpose of tool condition monitoring as given in above review.

In this section, development of a fuzzy expert system for optimizing the welding process will be presented that have the capability of self-learning, self-correcting and also self-

expanding. Following are the salient features of the presented self developing expert system [47].

1. Predicts the values of output process variables based upon values of input process variables.
2. Suggest the best values of input process variables to maximize and/or minimize the values of selected set of output process variables.
3. Adjusts newly entered variable at any stage of development automatically.
4. Self learns and corrects according to the new data set provided.
5. Generates fuzzy sets for newly entered process variable and regenerates sets for other variables according to newly added data automatically.
6. Generates the rules for the knowledge-base automatically.
7. Solve contradictory rules with conflict resolution facility.
8. Deletes outdated data from the database.

The first two features represent the main objective of the expert system while the other features describe the automation required for the system to self developing. This self developing expert system offers numerous benefits as given in the following:

1. Scope of the expert system can be expanded according to the requirements.
2. Minimum human involvement would be required for updating knowledge-base.
3. Higher precision upon more utilization of expert system.
4. No requirement of optimal formation of rule-base and automatic generation of rules.
5. The application of self-developing expert system is expected to be highly adaptive to the rapidly changing industrial environment.

The main components of the self-development mode of the expert system are: data acquisition module; fuzzy sets development module; and rule-base (optimization and prediction) development module [47]. In the following sub-sections, the objectives, functionality, and algorithms for these modules, are described. The section will follow with explanation of data structures and coding techniques for programming these modules. Two comprehensive examples will be presented to show the functioning of the automated expert system at the end of section.

4.2. Data Acquisition and Interface Development Module

This module facilitates the automation of intake, storage, and retrieval of data and development of the interface. The data may be the specifications of a new variable or the values of input and output variables resulted from experiments or empirical models. The data is stored in a file on the hard disk after intake.

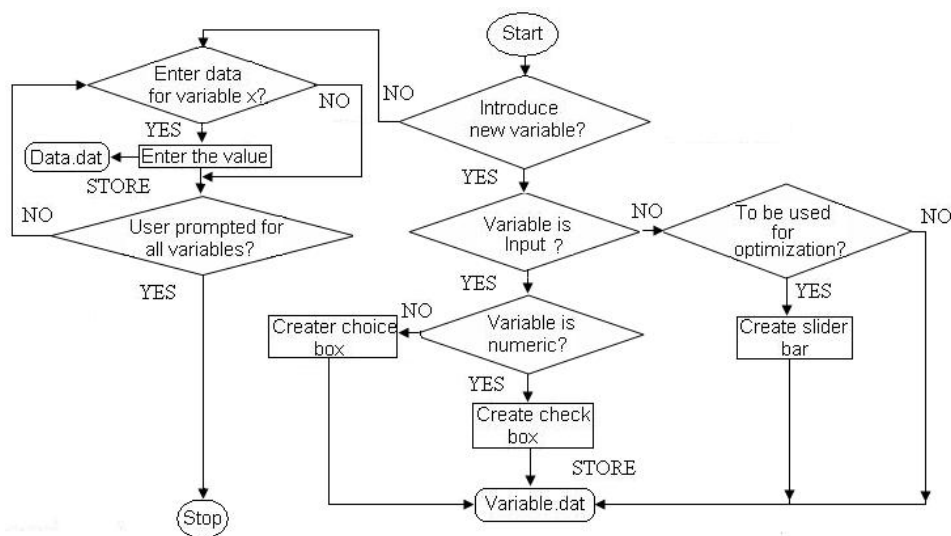


Figure 25. Flow chart for data acquisition and interface development module

Figure 25 shows flow chart of the data acquisition algorithm. The algorithm mostly constitutes of interaction with the user and consists of two parts: (1) introduction of a new process variable (predictor or response); and (2) addition of new practical data related to the variables already in use by the expert system. In part 1 the expert system collects the information about new variable regarding its type (input/output and numeric/categorical). Input variables can be numeric or categorical. If it is numeric a check box is created at the interface of the expert system asking whether the variable should be prefixed or not, otherwise a choice box, displaying all the possible values of the categorical variable, is created. Output variable can only be numeric and for each new output variable the user is enquired whether or not to include it for optimization purpose. If yes, a slider bar for that variable is created at the interface. From the slider bar, the user can specify whether to maximize or minimize the variable and also to how much desirability the objective needs to be satisfied. Specifications of the new variable are stored in file *Variable.dat*.

In part 2 the system prompts the user for practical data related to the variables in use. It is not compulsory for the user to enter data for all the variables but each data record should consist of data values related to at least two input variables and one output variable. Before further processing all the data records are loaded to a linked list named as *Set*.

4.3. Self-Development of Fuzzy Sets Module

This module deals with three processes: (1) Development of fuzzy sets for newly entered numeric variables; (2) Rearranging the fuzzy sets for already entered variables according to newly entered data records; and (3) Development of two fuzzy sets (*low* & *high*) for each output variable that is selected for optimization purpose. The set *low* represents minimization requirement and the other one represents maximization. The design of the sets for process 3 is fixed and is shown in figure 26, while the design of first two processes is dynamic and based on data values of respective variables. Desirability values shown in figure 26 are set by the

user using a slider bar available on the interface of the expert system. The figure shows that for any value below 5% means desirability is of totally minimizing the output variable, and total desirability of maximization is meant if the value is above 95%. Desirability of 50% means optimization of that output variable makes no difference.

Figure 27 shows a customized flowchart for the methodology used for the self-development of fuzzy sets. The user has to decide maximum allowable number of fuzzy sets for input as well as for output variables. The larger is the number of fuzzy sets the better are the optimization/prediction results but longer is the processing time. Thus, there is a trade-off between accuracy of results and processing time in the selection of maximum allowable number of fuzzy sets. Moreover, it has also been observed that increasing the number of sets beyond fifteen does not significantly affect the accuracy of results. Thus it has been fixed for the development of the expert system that the maximum number of fuzzy sets cannot be more than 10 for input variables and 15 for output variables. Following is the description of instructions, for developing fuzzy sets for any numeric variable x , as contained in the flowchart:

From the practical data records all the values of x are copied to a linked list $L1$ and sorted in ascending order. The list may also contain repeat values of the variable.

If x is input variable or an output variable with number of distinct values appearing in all data records lesser than the maximum allowable number of fuzzy sets (say $N2$), then all its distinct values from $L1$ are copied to another linked list $CL1$. Repeat values are not copied but the column “Appearances” is incremented accordingly. $CL1$ is then sorted in descending order of the number of appearances of the values.

Either top $N1$ (maximum allowable number of fuzzy sets for an input variable) or all of the values, whichever is smaller, are copied to another linked list $L2$, as shown in the figure 27.

To each of the values contained by $L2$ a separate triangular fuzzy set is assigned in Fuzzy CLIPS format. The logic involved in the methodology is that a value (of input variable) that has higher frequency of appearance in the data records possesses higher priority for allocation of a fuzzy set.

If x is an output variable with number of distinct values appearing in all data records greater than $N2$, then all the distinct values are copied from $L1$ to $CL1$ and for each of the values contained in $CL1$, *neighbor distance* is computed using following formula:

$$Neighbor\_Distance = \begin{cases} Value[i+1] - Value[i]; & \text{if } (i = \text{first}) \\ Value[i] - Value[i-1]; & \text{if } (i = \text{last}) \\ \frac{1}{2}(Value[i+1] - Value[i-1]); & \text{otherwise} \end{cases} \quad (2)$$

Respective *neighbor distance* is assigned to each of the values in $CL1$ and the list is sorted in descending order of *neighbor distance*.

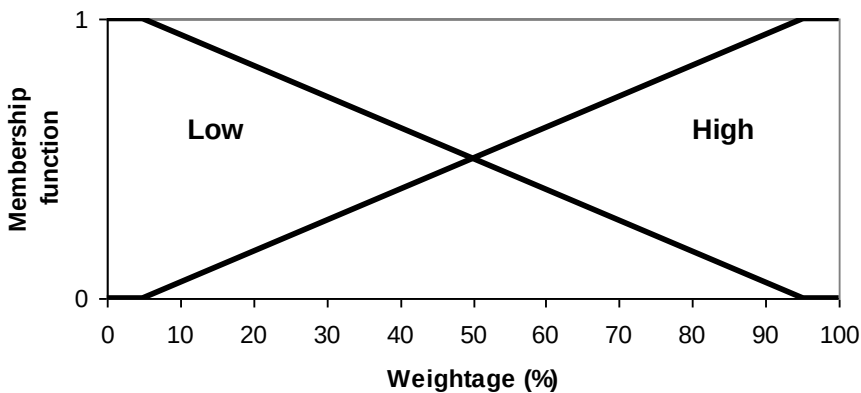


Figure 26. Fuzzy sets for maximization and/or minimization of output variable.

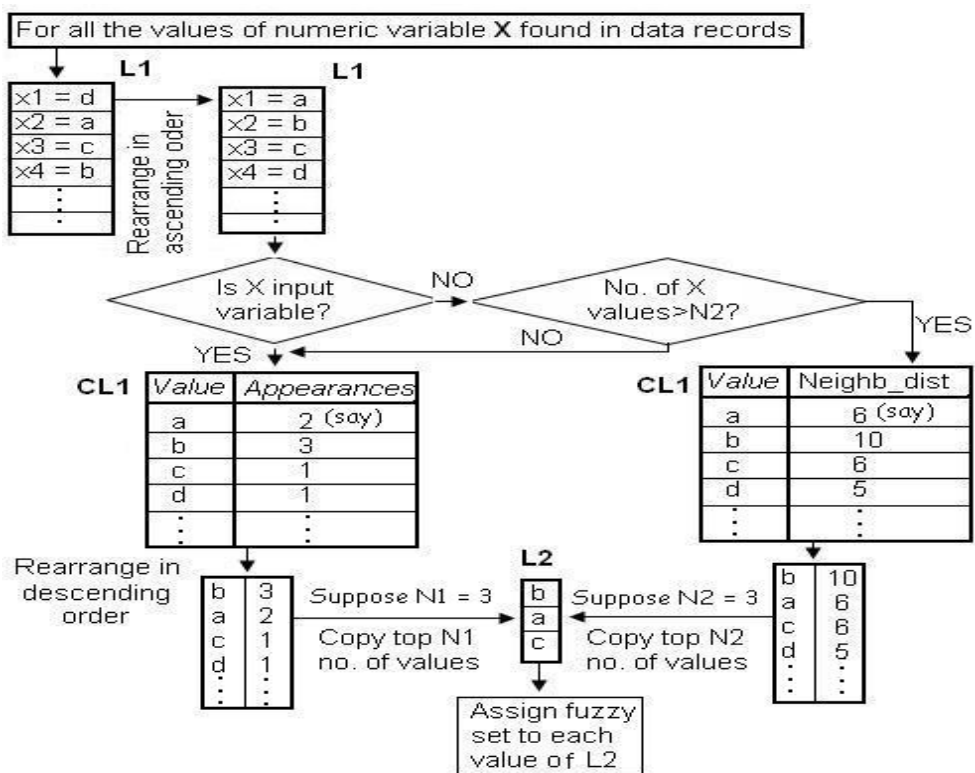


Figure 27. Customized flow chart for auto-development of fuzzy sets

Top N2 values are copied from CL1 to a linked list L2 and separate triangular fuzzy set is assigned to each of the values contained by L2. The idea utilized in this procedure is that any value (of output variable), in the list, possessing higher difference from its successor and predecessor, owns higher priority for allocation of a fuzzy set.

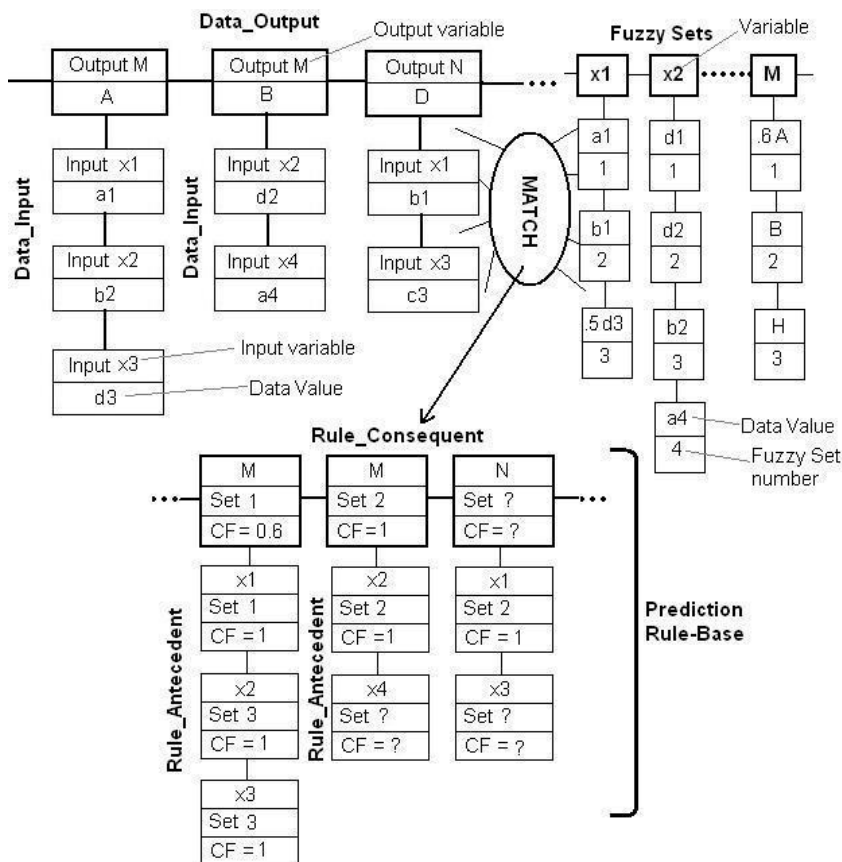


Figure 28. The framework for self-development of prediction rule-base

4.4. Self-Development of Prediction Rule-Base Module

This sub-section covers two parts: (1) automatic development of rules for prediction of process's performance measures, based on the data records provided by the users; and (2) conflict resolution among self-developed contradictory rules. In expert system's execution the priority of rule's firing is based on accomplishment of antecedent part of the rule and then on salience of respective rules specified by the rule-base developer. So, the sequence of appearance of the rules in the CLIPS file is absolutely immaterial. In this context, the development of prediction rule-base will be described before that of optimization rule-base. Figure 28 provides the graphical description of the algorithm for the development of prediction rule-base. Following is the brief description of the algorithm:

In the linked list *Set* there would be data records that contain data values of more than one output variables. The multiple output variables from such records are detached and for each output variable the record of relevant set of input variables is maintained in a doubly linked list named *Data_output*. Each node of this list contains value of output variable and to each node there is also connected a linked list *Data_input* that

contains respective data related to input variables. Thus, *Data_output* is the list of data records with one output variable per record.

The objective of the algorithm is to convert each node of *Data_output* (also consisting of list of related values of input variables *Data_input*) into a rule. This is achieved by finding and assigning most suitable fuzzy sets to all of the values involved in each node of *Data_output*.

The list *Data_output* is navigated from first node to last and for all of its values the closest values in fuzzy sets of respective variables are matched. If the match is perfect then certainty factor (*CF*) of 1 is assigned to the match of the data value and the fuzzy set. If suitable match of any fuzzy set for a given data value is not found then the data value is assigned the intersection of two closest fuzzy sets. This results in formation of prediction rules-base containing the number of rules equal to number of nodes in the linked list *Data_output*.

All the rules are stored in a doubly linked list, named *Rule_Consequent*, each node of which represents a rule. Each node of *Rule_Consequent* contains assigned fuzzy set of output variable and also a linked list (*Rule_antecedent*) containing assigned fuzzy sets of all the relevant input variables. To each rule is assigned a priority factor called *salience*, whose value is in direct proportion to the number of input variables involved in that rule. This emphasizes that a rule containing larger number of variables in its antecedent part enjoys a higher priority for firing.

4.4.1. Conflict resolution among contradictory rules

As new data are to be entered at free will of users, there is always a possibility that some anomalous data might be entered that could lead to development of some opposing rules. So it is utmost necessary to develop a mechanism that would detect such possible conflict among contradictory rules and would provide a way for its resolution. Figure 29 presents flow-chart of the algorithm that provides mechanism for conflict resolution.

The mechanism of conflict resolution algorithm can be described as follows:

Each and every rule of the prediction rule-base is compared to all the other rules of the same rule-base.

If, in the consequent parts of any two rules, following two conditions hold true: (1) response variables are same; and (2) assigned fuzzy sets are different, then it is checked whether the antecedent parts of both the rules are same (i.e., same predictor variables with same fuzzy sets assigned). If yes, these two form a pair of contradictory rules.

The user then is inquired regarding which one of the two contradictory rules needs to be abandoned. The *CF* value of that rule is set to zero.

Same procedure is continued for whole of the rule-base. At the completion of the process, all the rules possessing *CF* values greater than zero are printed to the CLIPS file: *Sets_Rules.clp*.

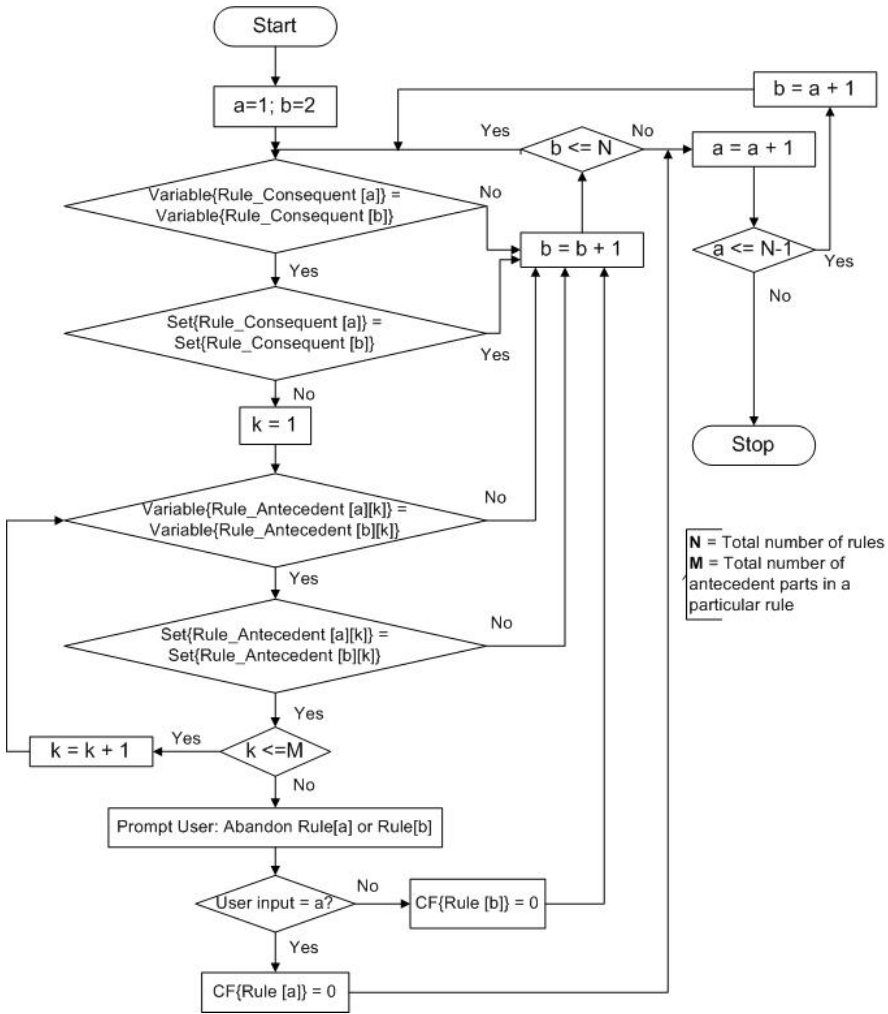


Figure 29. The algorithm for conflict resolution among contradictory rules.

4.5. Self-Development of Optimization Rule-Base Module

In this sub-section an algorithm is presented that leads to automatic generation of optimization rule-base. The optimization rule-base is responsible for providing optimal settings of input variables that would best satisfy maximization and/or minimization of the selected output variables. Figure 30 presents the graphical description of the methodology developed.

The idea utilized in this algorithm is that for maximization of any output variable ideal fuzzy sets should be selected for all the numeric input variables, which, on average, would generate maximum value of that output variable. For minimization purpose, those fuzzy sets, for respective input variables, should be selected that would result in smallest possible value of the output variable available in the data records. The procedural operation for automatic generation of rules for optimization purpose is based on following outline.

A question arises whether the rule-base generation procedure ensures optimality of the welding processes or not. Suppose that relationship between a response variable and a predictor variable is linear and also that the effect of interaction, between that predictor variable and other predictor variables, on the response variable is insignificant. For this case it means that if increase in value of that particular predictor variable causes increase or decrease in value of the response variable, it will cause the same effect to that response variable regardless of different combinations of other predictor variables. This suggests that if a fuzzy set of a predictor variable has been worked out (from the already developed prediction rule-base) as the one that contributes in generation of highest/lowest fuzzy set of a response variable, it will contribute in the same strength and same way regardless of any combination of fuzzy sets of different predictor variables. This ensures that the suggested values (or fuzzy sets) of predictor variables will deliver the optimal values of response variables. Now suppose that the relationship between the response variable and the predictor variable is not linear. For this case the optimization rule-base may not always suggest the optimal results because of the existing nonlinearity in the relationship. This shortcoming can be effectively addressed by providing additional practical data, related to the variables already in use, to the expert system. For the rare case in which interaction among different predictor variables exists, the optimality of the results can be enhanced by providing practical data related to the influential predictor variables that are not already covered by the expert system. Whatever the case may be, it must be kept in mind that the processes are optimized within the range of the data values provided to the system.

4.6. Application Examples

The fuzzy expert system presented in this section has been named as EXWeldHSLASteel (EXpert system for Welding of High Strength Low Alloy Steel of thin walled Shells). This sub-section describes the application examples showing the self-development of the knowledge-base and interface of EXWeldHSLASteel. The first example illustrates a fledgling knowledge-base that was self-developed from a very limited experimental data provided to it, while the second one portrays a veteran knowledge-base that reached this stage by continuously learning from the data that was supplied to it at different stages. The knowledge-base developed in second example covers all the experimental and statistical results of TIG welding experiments, presented throughout this chapter. The third example covers the verification of the EXWeldHSLASteel predictions by comparing them with the experimental results.

Consider limited experimental data provided in Table 26 that has been taken from the previous sections. The values for trailing, the fourth ingredient of the experiments, have been intentionally not included in the table. If the knowledge-base is developed based entirely upon these data, it is very likely that the expert system may provide anomalous results because of the fact that the other influential welding parameters (e.g., welding current etc.) have not been taken care of.

4.6.1. Example 1: A fledgling knowledge-base

Suppose the expert system is asked to develop its knowledge-base and update its interface based upon the data provided in Table 26 and it is also asked to include weld strength, but not the distortion, as output variable for optimization. Following is the detail of triangular fuzzy sets, in Fuzzy CLIPS format, developed itself by the expert system:

```
(deftemplate Obj_Weld_Strength 0 100 percent
  ( (Low (0 1) (5 1) (95 0) )
    (High (5 0) (95 1) (100 1) ) ) )
(deftemplate Thickness 2 6 mm
  ( (S1 (2 1) (3 1) (5 0) )
    (S2 (3 0) (5 1) (6 1) ) ) )
(deftemplate Welding_Voltage 9 15 V
  ( (S1 (9 1) (10.5 1) (13.5 0) )
    (S2 (10.5 0) (13.5 1) (15 1) ) ) )
(deftemplate Welding_Speed 13.5 19.5 cm/min
  ( (S1 (13.5 1) (15 1) (18 0) )
    (S2 (15 0) (18 1) (19.5 1) ) ) )
(deftemplate Weld_Strength 670 810 MPa
  ( (S1 (670 1) (725 1) (737.8 0) )
    (S2 (725 0) (737.8 1) (749.5 0) )
    (S3 (737.8 0) (749.5 1) (751 0) )
    (S4 (749.5 0) (751 1) (759.6 0) )
    (S5 (751 0) (759.6 1) (784.6 0) )
    (S6 (759.6 0) (784.6 1) (810 1) ) ) )
(deftemplate Distortion 0.5 7.5 mm
  ( (S1 (0.5 1) (3.6 1) (4.1 0) )
    (S2 (3.6 0) (4.1 1) (4.9 0) )
    (S3 (4.1 0) (4.9 1) (5.1 0) )
    (S4 (4.9 0) (5.1 1) (5.2 0) )
    (S5 (5.1 0) (5.2 1) (5.6 0) )
    (S6 (5.2 0) (5.6 1) (7.5 1) ) ) )
```

The first template is the one defining sets for maximization and minimization of weld strength, the process that has already been explained in section 4.3. The next three templates belong to input numeric variables, namely thickness, welding voltage and welding speed. The maximum allowable number of fuzzy sets for output variable was set to 6, thus, the last two templates have selected the best 6 values out of 8 for assignment of fuzzy sets. Following is the detail of six rules, self-developed by the expert system and operated by its optimization module:

Table 26. Data for the fledgling knowledge-base

No	Thickness	Voltage	Speed	Weld strength	Distortion
	(mm)	(V)	(cm/min)	(MPa)	(mm)
1	3	10.5	18	784.6	4.9
2	3	13.5	18	749.5	5.2
3	3	10.5	15	751.0	5.1
4	3	13.5	15	740.0	5.6
5	5	10.5	18	759.6	4.1
6	5	13.5	18	729.5	3.7
7	5	10.5	15	737.8	3.6
8	5	13.5	15	725.0	5.0

```

(defrule optimization1 (declare (salience 1000))
  (Obj_Weld_Strength High)
  (or (not (Thickness ?)) (Thickness S2))
  (assert (Thickness S2)))
(defrule optimization2 (declare (salience 1000))
  (Obj_Weld_Strength High)
  (or (not (Welding_Voltage ?)) (Welding_Voltage S1))
  (assert (Welding_Voltage S1)))
(defrule optimization3 (declare (salience 1000))
  (Obj_Weld_Strength High)
  (or (not (Welding_Speed ?)) (Welding_Speed S2))
  (assert (Welding_Speed S2)))
(defrule optimization4 (declare (salience 1000))
  (Obj_Weld_Strength Low)
  (or (not (Thickness ?)) (Thickness S1))
  (assert (Thickness S1)))
(defrule optimization5 (declare (salience 1000))
  (Obj_Weld_Strength Low)
  (or (not (Welding_Voltage ?)) (Welding_Voltage S2))
  (assert (Welding_Voltage S2)))
(defrule optimization6 (declare (salience 1000))
  (Obj_Weld_Strength Low)
  (or (not (Welding_Speed ?)) (Welding_Speed S1))
  (assert (Welding_Speed S1)))

```

Out of these six rules the first three perform the maximization operation, while the others perform minimization. Let us consider the first rule, whose first line consists of declaration of name of rule and its salience. The salience value is very high because the optimization rules are supposed to fire before prediction rules. The next two lines constitute the IF part of the rule and connected by AND operator. The antecedent part can be read as, “IF the objective is weld strength *high* AND Thickness is not fixed or Thickness is S2”. The symbol “=>” represents the term “THEN”. The consequent part of the rule can be read as, “Thickness is

S2”. Following is the detail of eight rules, self-developed by the expert system and operated by its prediction module:

```

(defrule prediction1 (declare (salience 15) (CF 1))
(Thickness S1)
(Welding_Voltage S1)
(Welding_Speed S1)
(assert (Weld_Strength S2 AND S3) CF 0.6918 (Distortion S6) CF 1))
(defrule prediction2 (declare (salience 15) (CF 1))
(Thickness S1)
(Welding_Voltage S2)
(Welding_Speed S1)
(assert (Weld_Strength S3) CF 1 (Distortion S4) CF 1))
(defrule prediction3 (declare (salience 15) (CF 1))
(Thickness S1)
(Welding_Voltage S1)
(Welding_Speed S2)
(assert (Weld_Strength S6) CF 1 (Distortion S2) CF 1))
(defrule prediction4 (declare (salience 15) (CF 1))
(Thickness S1)
(Welding_Voltage S2)
(Welding_Speed S2)
(assert (Weld_Strength S5) CF 1 (Distortion S1) CF 0.7826))
(defrule prediction5 (declare (salience 15) (CF 1))
(Thickness S2)
(Welding_Voltage S1)
(Welding_Speed S1)
(assert (Weld_Strength S2) CF 0.243697 (Distortion S3) CF 1))
(defrule prediction6 (declare (salience 15) (CF 1))
(Thickness S2)
(Welding_Voltage S2)
(Welding_Speed S1)
(assert (Weld_Strength S1) CF 1 (Distortion S5) CF 1))
(defrule prediction7 (declare (salience 15) (CF 1))
(Thickness S2)
(Welding_Voltage S1)
(Welding_Speed S2)
(assert (Weld_Strength S5) CF 1 (Distortion S2) CF 1))
(defrule prediction8 (declare (salience 15) (CF 1))
(Thickness S2)
(Welding_Voltage S2)
(Welding_Speed S2)
(assert (Weld_Strength S4) CF 1 (Distortion S3) CF 0.90625))

```

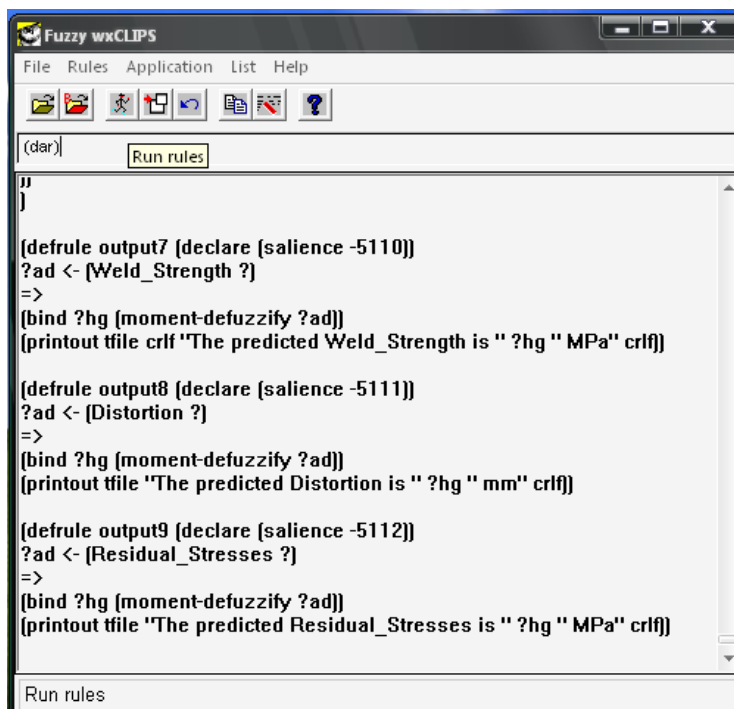


Figure 31. Continued

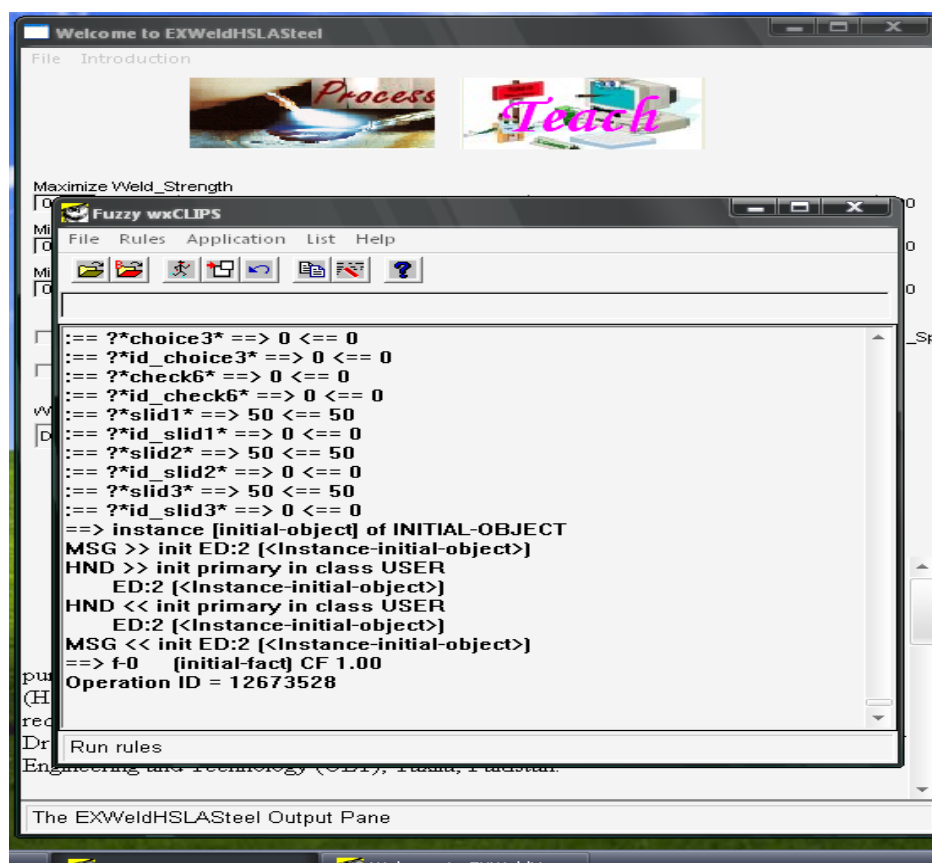


Figure 31. Process of interface of expert system from fuzzy CLIPS

Considering 2 fuzzy sets each for thickness, voltage and welding speed, the total number of prediction rules is 16. Saliency of each rule is equal to 15 (= number of input variables in the rule $\times$ 5). First line of each rule consists of the name of rule, its saliency and calculated certainty factor (CF). The next three lines form the antecedent part of rule, while the last line is the consequent part. In consequent parts of all the rules, two assertions have been made, one for weld strength and other one for distortion.

Figure 31 shows the process of interface of the expert system from fuzzy CLIPS and Figure 32 shows the interface of the expert system related to the fledgling knowledge-base. In Figure 32, top of the interface shows two buttons, one is for processing the optimization and prediction of welding process, while the second one is for self-development of expert system for optimizing welding process according to new data provided to it.

The slider bar provides the user whether to maximize or minimize the selected output variable and by how much weightage. Check-boxes are for numerical input variables asking the user whether to pre-fix them or optimize them according to the desired objective(s). These are followed by the choice-boxes for categorical input variables providing the possible choices for respective variables, including the option of leaving them open for optimization (i.e. "Do not know").

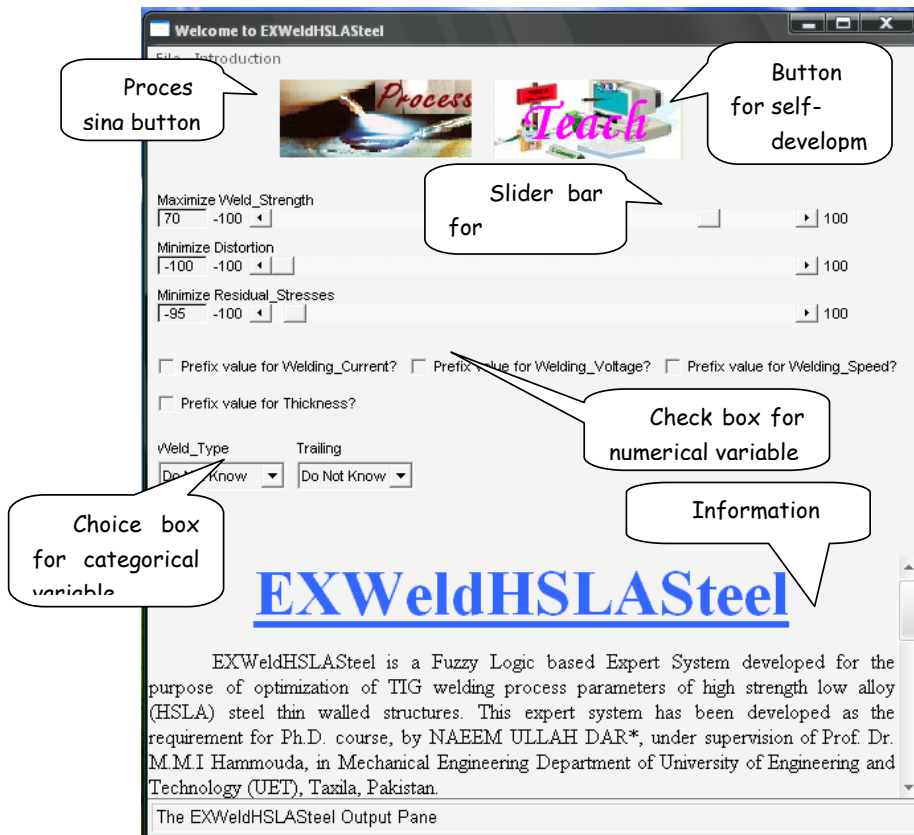


Figure 32. Interface of expert system representing fledgling knowledge-base

At bottom of the interface there is information pane that initially displays the introduction of EXWeldHSLASteel and then, after processing, it displays the results of optimization and prediction processes. Suppose EXWeldHSLASteel is provided with following input:

Objective: maximize weld strength with weightage of 95%

Thickness of material prefixed to 3.5mm.

Welding voltage and welding speed: *open* for optimization.

Pressing the *Process* button starts the processing of expert system and finally following results are displayed in the information pane:

The recommended welding speed is 17cm/min.

The recommended welding voltage is 1 A.

The predicted weld strength is 755MPa.

The predicted distortion is 4.3mm.

4.6.2. Example 2: A veteran knowledge-base

The veteran knowledge-base consists of all the data obtained from the welding experiments of HSLA steel shells for weld strength, distortion, and residual stresses fed to the knowledge-base. Figure 33 shows the interface of the expert system related to that knowledge-base.

Three output variables, namely: weld strength, distortion and residual stresses are selected for simultaneous optimization purpose. The interface contains three slider bars for this purpose. It can be further observed that the expert system at this stage is dealing with six input variables, four of them numeric and two categorical. Suppose the expert system is provided with following input:

Simultaneously maximize/minimize following performance measures: (1) maximize weld strength minimize with weightage of 70%; (2) minimize distortion with weightage of 100%; and (3) minimize residual stresses with weightage of 95%.

Prefix the value of work piece material thickness to 5 mm.

Prefix the value of welding current to 230 A.

Weld Type is Linear.

Leave the other input variables: welding voltage, trailing, and welding speed as *open* in order to be optimized.

Figure 33. Interface of expert system representing veteran knowledge-base

EXWeldHSLASteel provides following results, as displayed in information pane of the interface:

The recommended trailing is Ar.

The recommended value of welding speed is 17cm/min.

The recommended value of welding voltage is 11V.

The predicted value of weld strength is 780MPa.

The predicted value of distortion is 2.0mm.

The predicted value of residual stresses is 350MPa.

It is to be considered that the maximized value of weld strength is very satisfactory considering the fact that very high value of thickness was prefixed. Residual stresses value minimized by EXWeldHSLASteel seems quite high because of the fact that weightage of this objective was small as compared to other opposing objectives.

4.6.3. Example 3: Verification of EXWeldHSLASteel predictions

For the verification of EXWeldHSLASteel predictions against the welding parameters that already not fed for the maximization/minimization of responses of weld strength/distortion & residual stresses as given in Table 27 with prefixing the thickness and welding current, the responses are compared with experimental results as given in Table 28. The maximum variations observed between responses values are between 3-8% only.

Table 27. Welding Parameters for EXWeldHSLASteel Predictions

S.No.	Sheet Thickness (mm)	Welding Current (A)	Welding Voltage (V)	Welding Speed (cm/min)	Trailing	Weld-Type
01	3.5	200	11.5	17	Ar	Linear
02	4.5	220	11.5	17	Ar	Linear
03	3.5	200	11.5	17	Ar	Circumferential
04	4.5	220	11.5	17	Ar	Circumferential

Table 28. Comparison of Responses against welding parameters in Table 27

S.No.	Exweldhsla Steel			Experiment		
	Weld Strength (MPa)	Distortion (mm)	Residual Stresses (MPa)	Weld Strength (MPa)	Distortion (mm)	Residual Stresses (MPa)
01	746	3.00	460.0	724	3.25	485
02	763	2.59	395.4	778	2.48	420
03	742	2.36	428.0	-	2.21	405
04	763	2.33	318.6	-	2.15	298

4.6.4. The limitations of exweldhslasteel

The examples mentioned above present the compliance, efficacy, and adaptability of the developed expert system. Besides numerous advantages, EXWeldHSLASteel has also few

minor limitations. To ensure the effectiveness and reliability of the expert system, it is utmost important that the welding experimental data provided to EXWeldHSLASteel, for purpose of further self-development, should be based upon some statistical *DoE* technique. If this is not taken care of then the system may provide anomalous optimization results and it may also fail to provide predictions of some of the welding performance measures desired.

By providing more and more welding experimental data (based upon *DoE* technique) related to the input variables already in use by EXWeldHSLASteel, adds to its accuracy and reliability and providing welding experimental data related to some newly introduced variable, adds to its scope and span of application. If the data related to new welding input variable is based upon some fractional factorial design rather than full factorial design, it might compromise the accuracy of optimization and prediction results. If this situation is unavoidable then the accuracy and reliability of EXWeldHSLASteel can be enhanced to a certain degree by reducing the maximum allowable number of fuzzy sets of input numeric variables.

5. CONCLUSIONS

In this research work, expert system tool has been successfully applied for optimization of parameters and prediction of performance measures related to TIG welding process of thin walled HSLA steel shells. The optimization of parameters is performed based upon objective(s) of maximization and/or minimization of certain combination of performance measures. At the completion of optimization process the finalized settings of input variables are used to predict the values of the performance measures. This expert system tool possesses high potentials for reducing production cost, cutting down lead-time, and improving the product quality at expense of few seconds that the expert system would take to process.

The important feature of this research work is the success in imparting self-developing abilities to the fuzzy expert system for welding process optimization. The presented expert system is capable of auto-managing data, self-developing fuzzy sets, self-generating rule-base, automatically updating expert system interface, and providing conflict resolution among contradictory rules. These abilities make the expert system exceedingly adaptable to continuously changing high-tech industrial environments, without need of human intervention in the field of welding of thin walled structures.

The developed tool for optimization of welding process parameters and prediction of responses consumes only few seconds to give desired solution before the start of process on shop floor and this may be used in shipbuilding, aerospace and nuclear industries, oil and gas engineering and in other areas before the manufacturing of structural elements.

REFERENCES

- [1] Xiangyang, L. *Influence of Residual Stress on Fatigue Failure of Welded Structures*. PhD Thesis, North Carolina State University, 2002.
- [2] Radaj, D. *Heat Effects of Welding*, Springer-Verlag, 1992.
- [3] Masubuchi, K. *Analysis of Welded Structures*, Pergamon Press, 1980.

- [4] Fanous, IFZ; Younman, MYA; Wifi, AS. *ASME J Press Vess- T*, 2003, 125(4), 432-439.
- [5] Teng, JG; Lin, X; Rotter, JM; Ding XL. *J Eng. Struct*, 2005, 27(6), 938-950.
- [6] Dar, NU; Qureshi, EM; Khan, I; Malik, A.M. *Welding Quality and Cost: A Comprehensive Comparative Study. Proc. Conf. Adv. Design Manuf.*, Harbin, China, 2006
- [7] Tanco, M; Ilzarbe, L; Viles, E; Alvarez, M.J. *J Eng. Manuf.* 2008, 222(8), 1035-1042.
- [8] Gunaraj, V; Murugan, N. *J Mater. Process. Technol.*, 1999, 88 (1-3), 266-275.
- [9] Benyounis, KY; Olabi, AG; Hashmi, MS. *J. Opt. Laser Technol*, 2008, 40(1), 76-87.
- [10] Volden, L; Gundersen, O; Rarvik, G. *Development of Residual Stresses in High Strength Low Alloy Steel*; Proc. 9th Int. Offshore Polar Eng. Conf., Brest, France, May 30 - June 4, 1999.
- [11] Guan, Q; Zhang, CX; Guo, DL. *Weld. World*, 1994, 33(4), 308-313.
- [12] van der Aa, EM; Richardson, IM; Hermans, MJM. *Welding with a Trailing Heat Sink: How to Optimize the Cooling Parameters?*, *Trends in Welding Research*, Pine Mountain, GA, ASM International, 2005.
- [13] van der Aa, EM. *Local Cooling during Welding: Prediction and Control of Residual Stresses and Buckling Distortion*, *PhD thesis*, Delft University of Technology, 2007.
- [14] Rosenthal, D. *The Theory of Moving Heat Source and its Application to Metal Treatment*, *Transactions ASME*, 1946.
- [15] Goldak, J; Zhou, J; Breiguine, V; Montoya, F. *JWRI*. 1996, 25(2), 1851-1889.
- [16] Goldak, J; Chakravarti, A; Bibby, M. *Metall. Trans, B*, 1984, 15(B), 299-305.
- [17] Goldak, J; Bibby, M; Moore, J; House, R; Patel, B. *Metall. Trans, B*, 1986, 17(B), 587-600.
- [18] Rybicki, EF; McGuire, PA; Merrick, E; Wert, B. *J Press Vess-T*, 104, 204-209.
- [19] Konar, A. *Introduction to Artificial Intelligence and Soft Computing*, in: *Artificial Intelligence and Soft Computing*, CRC Press LLC, FL, 2000.
- [20] Nilsson, NJ. *Artificial Intelligence: A New Synthesis*, Morgan Kaufmann Publishers, USA, 1998.
- [21] Hopgood, AA. *Intelligent Systems for Engineers and Scientists*, 2nd Edition, CRC Press LLC, FL., 2001
- [22] Gonzalez, AJ; Dankel, DD. *The Engineering of Knowledge-Based Systems: Theory and Practice*, Prentice Hall, NJ, 16-22, 1993.
- [23] Ganesh, M. *Introduction to Fuzzy Sets and Fuzzy Logic*, Prentice Hall, NJ, 147-174, 2006.
- [24] *ASM Handbook (Welding, Brazing and Soldering)*. Vol. 6, 1025, 1059.
- [25] Liao, TW. *Expert Syst. Appl.*, 2003, 25, 101-111.
- [26] Taylor, A. *Int. Journal Prod. Res.*, 1989, 27(11), 1855-1862.
- [27] Varde, AS; Maniruzzaman, M; Rundensteiner, EA; Sisson Jr, R.D. *The QuenchMiner Expert System for Quenching and Distortion Control*, *Proc. ASM 2nd Int. Heat Treat. Soc. Conf.*, Indianapolis IN, 2003.
- [28] Kim, D; Rhee, S; Park, H. *Int. J Prod. Res.*, 2004, 40, 1699-1711.
- [29] Lin, HL; Chou, CP. *Sci. & Technol, Welding & Join.* 2006, 11(1), 120-126.
- [30] Ganjigatti, JP; Pratihari, DK. *J Intell. Fuzzy Sys.*, 2008, 19(2), 15-130.
- [31] Tsoukalas, V; Kontesis, M; Badogiannis, E; Papachristos, D; Fragiadakis, N. *Prototype of an Expert System for Aluminum Welding*; Proc. 5th WSEAS Int. Conf. Comput. Intell., *Man-Machine Sys. & Cyber*; Venice, Italy, 2006, 78-83.
- [32] Tanco, M; Viles, E; Pozueta, L; Are All of Experiments Approaches Suitable for Your Company?, *World Congress on Engineering*, London UK, 2008, 1202-1207.

- [33] Malik, AM; Qureshi, EM; Dar, NU; Khan, I. TIG Welding Process: Experimental Validation of Simulated Results. *Proc. Int. Conf. Adv. Design & Manuf.*, Harbin, China, 2006.
- [34] Dar, NU. Expert System for Optimization of Welding Process of Thin-Walled HSLA Steel Structures. *Ph.D. Thesis*, UET Taxila, Pakistan, 2009.
- [35] Pham, DT; Pham, PTN. Computational Intelligence for Manufacturing, in: *Computational Intelligence in Manufacturing Handbook*, CRC Press LLC, Florida, 2001.
- [36] Orchard, RA. Fuzzy CLIPS, Version 6.04A; *User's Guide*, National Research Council, Canada, 1998.
- [37] http://www.iit.nrc.ca/IR_public/fuzzy/fuzzyJDocs/APIdocs/nrc/fuzzy/FuzzyValue.html.
- [38] Hashmi, K; Graham, ID; Mills, B; Hashmi, MSJ. *J. Mater. Process. Technol.*, 2003, 142, 152-162.
- [39] Laarhoven, PJM; Arts, EHL. *Simulated Annealing: Theory and Applications*; Kluwer Academic Publishers, Dordrecht, Netherlands, 1987.
- [40] Iqbal, A; He, N; Li, L; Dar, NU. Simulated annealing assisted optimization of fuzzy rules for maximizing tool life in high-speed milling process; *Proc. 5th IASTED Int. Conf. Artif. Intell. & Appl.*, Innsbruck, Austria, 2006, 335-340.
- [41] Lekova, A; Batanov, D. *Comput. Indust.*, 1998, 37, 135-141.
- [42] Castro, JL; Castro-Schez, JJ; Zurita, JM. *Fuzzy Sets Syst.*, 2001, 123, 307-320.
- [43] Filipic, B; Junkar, M. *Comput. Indust.*, 2000, 43, 31-41.
- [44] Monostori, L. *Eng. Appl. Artif. Intell.*, 2003, 16, 277-291.
- [45] Al Assadi, HMAA; Wong, SV; Hamouda, AMS; Ahmad, MMMH. *J Mater. Process. Technol.*, 2004, 155-156, 2087-2092.
- [46] Cho, S; Asfour, S; Onar, A; Kaundinya, N. *Int. J. Mach. Tools & Manuf.*, 2005, 45, 241-249.
- [47] Iqbal, A; Dar, N.U; He, N; Hammouda, MMI; Li, L. *J. Intell. Manuf.* (published online), DOI: 10.1007/s10845-009-0252-3.

Chapter 4

FRICION STIR WELDING: FLOW BEHAVIOUR AND MATERIAL INTERACTIONS OF TWO SIMILAR AND TWO DISSIMILAR METALS AND THEIR WELDMENT PROPERTIES

Indra Putra Almanar¹ and Zuhailawati Hussain^{*2}

¹School of Mechanical Engineering,

²School of Materials and Mineral Resources Engineering,
Universiti Sains Malaysia, Engineering Campus, Nibong Tebal, Penang, Malaysia

ABSTRACT

In friction stir welding of two similar and dissimilar metals, the work materials are butted together with a tool stirrer probe positioned on the welding line. The work materials in the welding area are softened due to heat generation through friction between the probe and the surface of the work materials. Upon the softening of the work materials, the friction will be diminished due to the loss of frictional force applied between the tool stirrer probe and the softening surface of work materials. The probe then penetrates the work material upon the application of the axial load and the tool shoulder confines the working volume. In this configuration, the advancing and retreating zones are created relevant to the direction of the probe rotational direction. At the same time the leading and trailing zones are also created relevant to the direction of motion of the tool. These zones determine the flow behavior of the softened work materials, which determine the properties of the weldment. Since the chemical, mechanical, and thermal properties of materials are different, the flow behavior of dissimilar materials becomes complex. In addition, material interaction in the softened work materials influences material flow and mechanical intermixing in the weldment. This review discusses the fundamental understanding in flow behavior of metal during the friction stir welding process and its metallurgical consequences. The focus is on materials interaction,

* Corresponding author: Email: zuhaila@eng.usm.my, Telephone: 604-5995258 Fax: 604-5941011.

microstructural formation and weldment properties for the similar and dissimilar metals. Working principles of the process are explained beforehand.

1. INTRODUCTION

The quality of metal weldment depends on how it is formed. In fusion welding such as electric arc welding, oxy-acetylene, etc., the weldment is formed by placing molten filler material in between the melting areas of base metals to be joined in order to fuse them together as molten nugget which when solidified becomes the weldment (Figure 1). This technique is widely used in construction works, piping and some other applications because it is easy to operate. However, the heat history of fusion produces some disadvantages because porosity is likely to be formed due to gas entrapment in the molten nugget. Solidification of the molten nugget also significantly changes the microstructure of the weldment, which deteriorates the quality of the welded structure.

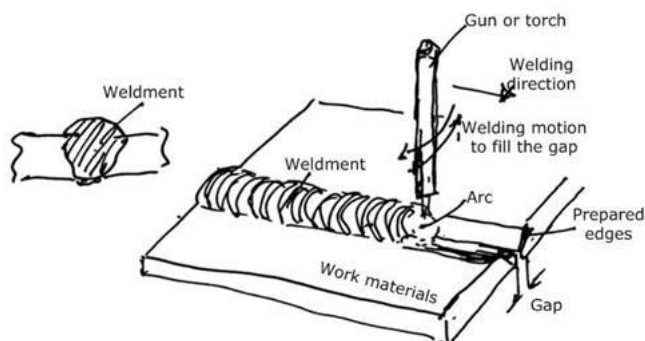


Figure 1. Typical fusion welding with filler material in butted configuration

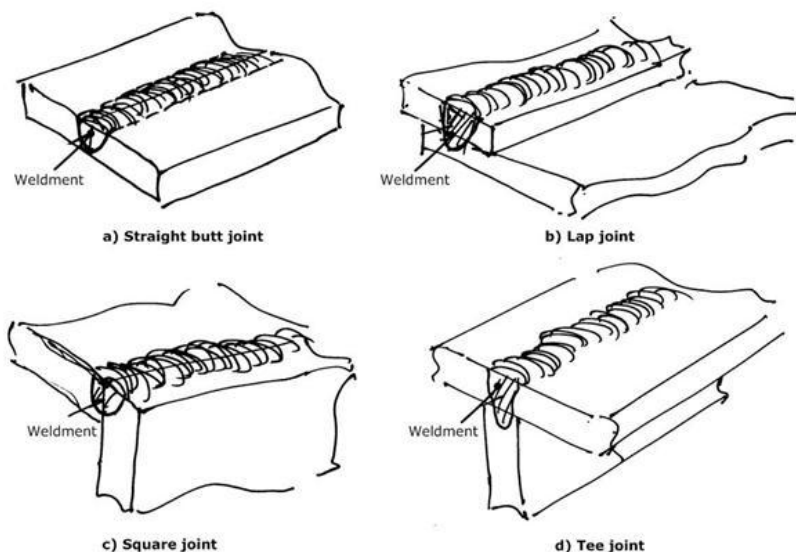


Figure 2. Configurations that are possible for FSW operations.

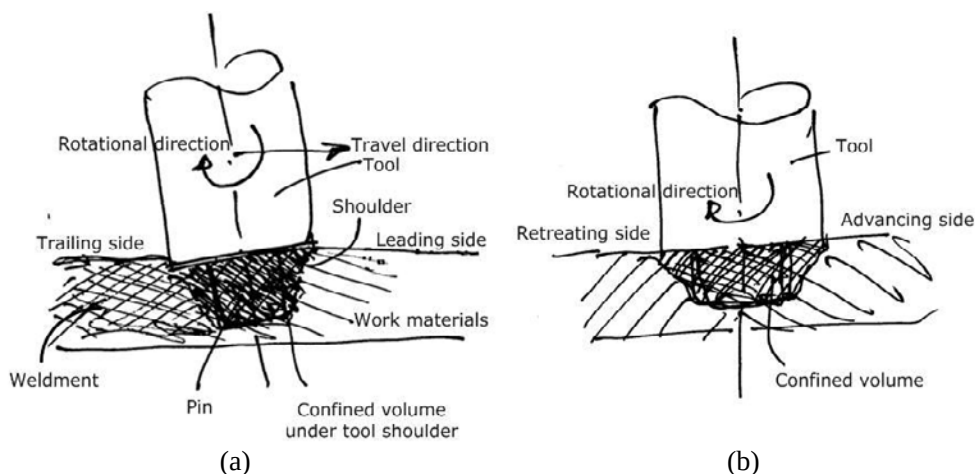


Figure 3. (a) Leading and trailing sides, and confined volume viewed from retreating side and (b) advancing and retreating sides, and confined volume viewed from leading side

Friction stir welding (FSW) process does not have the above-mentioned drawbacks. This is because the weldment is formed through a mechanical bonding of materials below their melting temperatures. This welding technique forms the weldment by using the materials taken from the areas to be joined. Joint configurations that can be used by this technique are butt, lap, square or tee (Figure 2).

In this technique, the success of the weldment formation depends on the flow behavior of softened work materials inside the confined volume under the tool shoulder in different regions around the rotating pin (Figure 3). This is due to the different characteristics of each region in the confined volume enclosed by the shoulder and the peripheral of the softened work materials.

Even though the chemical, mechanical, and thermal properties of materials are the same, the flow behavior of two similar materials is complex. When two dissimilar materials are used, the situation becomes more complex. In addition, material interactions in the softened stage influence the flow behavior of materials and thus, the quality of mechanical bonding in the weldment. Further understanding in material interactions in joining of similar and dissimilar materials should also consider the possibility of the formation of brittle intermetallics and low melting point eutectic. This chapter discusses the fundamental understanding of flow behavior of material during the friction stir welding process and their metallurgical consequences. Attention is given to material interactions, microstructural formation and weldment properties for joining two similar and dissimilar metals. However, working principles related to the process itself should be well understood beforehand.

2.0. WORKING PRINCIPLE OF FRICTION STIR WELDING

FSW is a welding technique that uses heat generated by mechanical friction between the rotating tool and the stationary work materials. Thus, to begin with, there must be surfaces that are moving under different relative velocities and being complemented by normal force acting on them in order to produce heat energy. The softening process of the work materials

depends on the amount of normal force applied and the difference in relative velocities given. The higher the force and the difference in relative velocities, the higher the frictional energy generated. When the heat generated is already reaching a sufficient amount, the work materials will then be softened. Once the surfaces become soft, the normal force will lose its function to keep the mechanical friction in producing heat. Then the coefficient of friction becomes lower which means that there is not much heat generated by the mechanical friction anymore. Thus, this situation will automatically ensure that there will be no more heat energy that could be generated by mechanical friction in order to reach a temperature high enough to melt the work materials.

Thus, to follow the basic principles explained above, work materials to be welded should be firmly fixed against heavy mechanical friction that would be applied on them (Figure 4). It can be in lapped, butted, squared or teed configurations. Moreover, since this is a non-filler material technique, there should not be any part of work materials missing in the welding line during the weldment formation. Thus, since there will be no filler material supplied in order to form the weldment, any shortages of work materials will produce a cavity in the weldment.

The areas to be welded are then softened by heat generated by mechanical friction created by the cylindrical-shouldered tool with cylindrical pin rotating at a constant speed and under an axial load, positioned on the work materials. The heat softens the materials and subsequently the rotating pin penetrates until the shoulder touches the surface of the work materials. The shoulder is then kept in an intimate contact with the surface of the work materials in order to provide a confined volume underneath the shoulder. This volume is required for the stirring process of the work materials in the formation of weldment. At this stage, the welding process has just begun. However, the weldment has not been formed yet.

The weldment formation in FSW occurs in the trailing side of the tool traveling direction where the materials from the advancing side is mixed mechanically in the confined volume with the materials from the retreating side by the stirring action of the tool pin and the shoulder. Thus, the weldment can only be formed when the rotating cylindrical-shouldered tool with the pin inside the confined volume of work materials accumulates the mixture of work materials in trailing side upon the travels of the tool along the welding line.

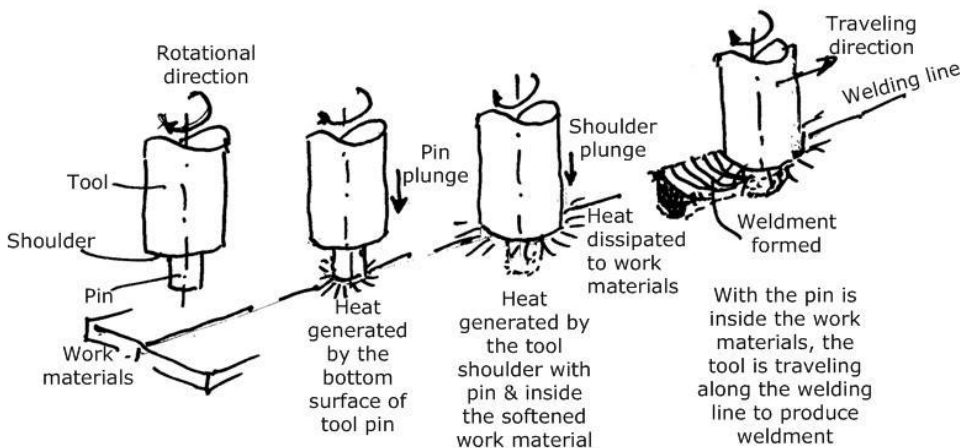


Figure 4. Operational sequence of FSW in butted configuration.

2.1. Heat Generation by Mechanical Friction Principle

A cylindrical-shouldered tool with threaded or plain cylindrical pin, whose length is slightly less than the thickness of the work materials for butt welding, or slightly less than the thickness of two overlapping work materials for lap welding, rotating at a constant speed under an axial pressure, is positioned on the surface of work materials to be welded. The rotating tool is then pushed onto the surfaces of the work materials. The lower surface of the rotating pin makes the first contact with the surface of work materials and upon axial pressure, the heat generation by mechanical friction in thermo mechanical joining process is started.

Although there will be heat losses during the thermo mechanical joining process as depicted in Figure 5, heat that is generated by the mechanical friction between the surfaces of the rotating-traveling tool pin and shoulder against the work materials inside the confined volume will be considered as one of two main sources of heat that contributes to the welding process (Figure 6). The other main source is the heat generated during material deformation inside the confined volume around the tool.

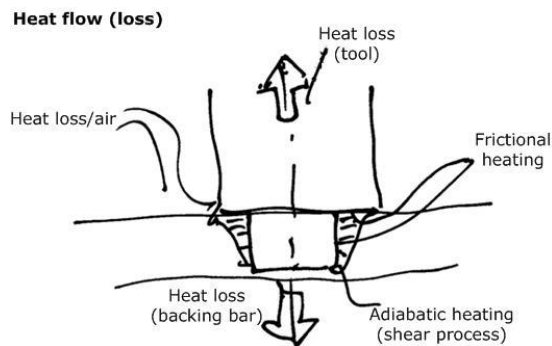


Figure 5. Thermo-mechanical joining process.

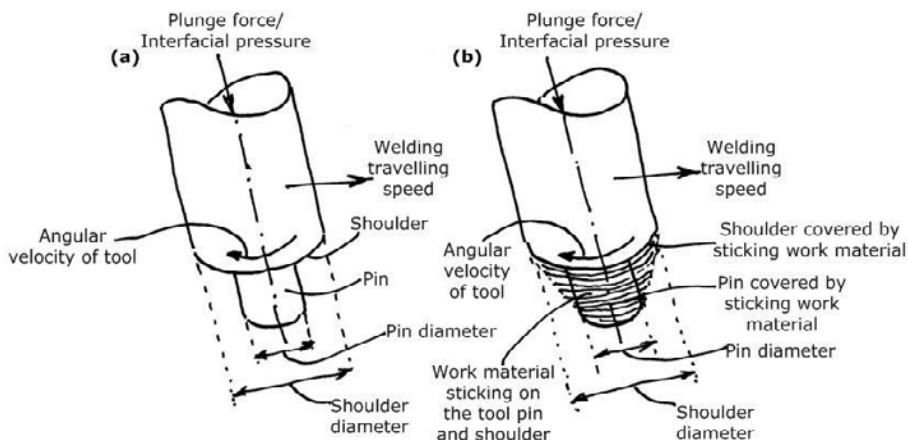


Figure 6. Typical tool configuration used in FSW (a) parts of the tool used to generate heat during mechanical friction and (b) workpiece sticks on the tool pin and shoulder.

The bottom surface of the rotating pin makes the first contact with the surface of work materials and upon the plunge force acting on tool axis, the heat generation by mechanical friction is started. When the bottom surface of the rotating pin is sliding on the surface of the work materials, the amount of heat energy Q generated by the mechanical friction can be expressed as:

$$Q_{pin-bottom} = 2/3[\pi P \mu \omega (R_{pin-bottom}^3)] \quad (1)$$

Where:

- $R_{pin-bottom}$ = the bottom radius of tool pin (mm)
- P = plunge force or interfacial pressure (N/mm^2)
- μ = friction coefficient (dimensionless)
- ω = the angular velocity of the tool (rpm)

This is the set up that allows the generation of heat by friction to soften the surface of work materials, sufficient for the rotating pin to penetrate the work materials.

When the surfaces of the work materials are softened, the rotating pin, which is still under the axial pressure, penetrates into the work materials until the rotating shoulder slides on the surface of the work materials. The rotating shoulder is now in the position of making an intimate contact with the surfaces of the work materials.

Since the amount of surface area of the rotating shoulder, which is sliding on the work materials, is bigger than the surface area of the lower bottom of the pin, the introduction of the rotating shoulder onto the surface of work materials will intensify the generation of heat by friction. Thus, the amount of heat energy Q generated by the mechanical friction of the shoulder and the bottom of the pin is:

$$Q_{shoulder} = 2/3[\pi P \mu \omega (R_{shoulder}^3)] \quad (2)$$

Where:

- $R_{shoulder}$ = the radius of tool shoulder (mm)

When the work materials become soft, the interfacial pressure P loses its ground. Thus, in order to keep the softened materials under the tool shoulder intact, the tool shoulder is kept in the intimate interfacial position. Interfacial pressure is then not fully functioning, the value of friction coefficient μ becomes uncertain. However, because the sliding-sticking interaction between the surface of the shoulder and the surface of the work materials some heat may still be generated by sliding friction, and some of the heat comes from sticking friction.

In sticking friction, the geometrical shape of the rotating tool pin is changed because the confined volume is now sticking and covering up the shoulder and the pin (Figure 6b). The shape of the tool is modified and there is no shoulder anymore. When the tool with the modified pin is rotating and traveling to perform welding operation, the work materials will be splashing out. The amount of heat generated by mechanical friction of the surface of the modified pin and the work materials becomes enormous. This is because the large surface area of the modified rotating-traveling pin makes an intensive sweeping work on the advancing and retreating sides of the work materials and accumulate the materials in the

trailing side in order to form weldment. However, since plenty of work materials were splashed during the welding process, the weldment produced will prone to have cavities due to the shortage of materials to form the weldment.

If the shape of the modified pin is considered as a conical frustum, the surface area of the modified pin is:

$$\pi(R_{\text{shoulder}} + R_{\text{pin}})\sqrt{(R_{\text{shoulder}} - R_{\text{pin}})^2 + h^2} \quad (3)$$

where

R_{shoulder} = the outer radius of the tool shoulder (mm)

R_{pin} = the radius of the tool pin (mm)

h = the length of the tool pin (mm)

Thus, the amount of heat energy Q generated by the mechanical friction of the modified pin is:

$$Q = 2/3[\tau\omega\pi(R_{\text{shoulder}} + R_{\text{pin}})\sqrt{(R_{\text{shoulder}} - R_{\text{pin}})^2 + h^2}] \quad (4)$$

Where:

τ = torque acting on rotating-traveling tool (Nm)

2.2. How the Weldment Is Formed

In FSW, the weldment is formed in the trailing side while the tool is moving along the welding line. The weldment is the accumulated materials swept from advancing and retreating sides which are mechanically mixed by the stirring action of the rotating-traveling tool pin inside the confined volume. Thus, the weldment cannot be formed if the tool is not traveling.

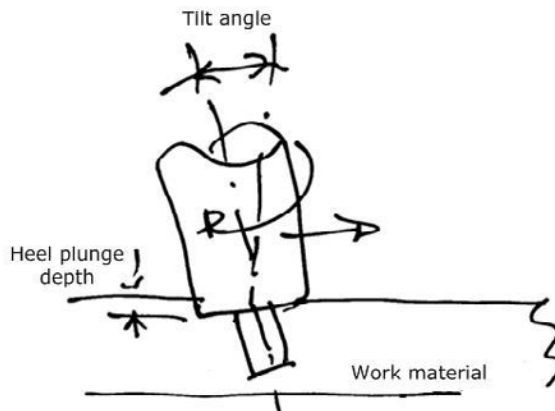


Figure 7. Tilt angle in tool positional configuration.

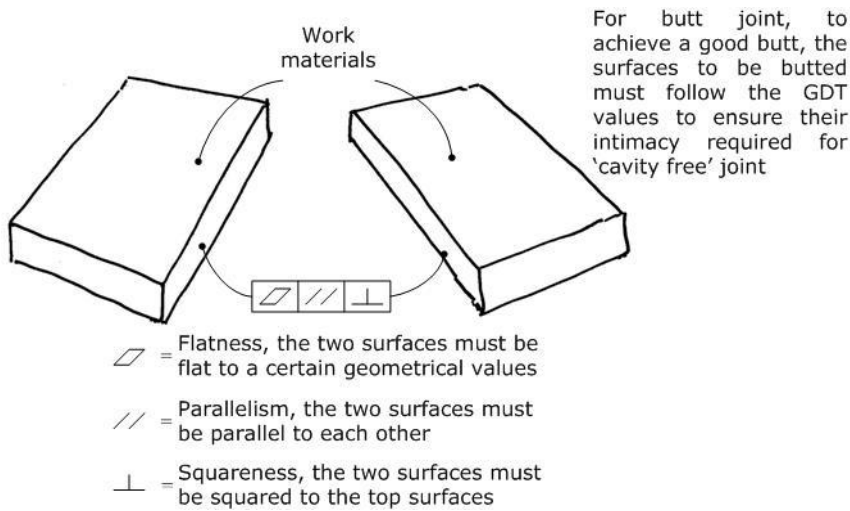


Figure 8. Geometrical, Dimensional and Tolerance (GDT) conditions to be fulfilled for butted joint in FSW process.

To pack together the accumulated swept materials in the trailing side, the tool normal position is slightly tilted backward to produce the heel plunge depth beneficial for the weldment compaction (Figure 7). Another benefit gained from the tilting backward of the tool normal position is the materials in the leading side will not be scraped by the outer rim of the tool shoulder that will reduce the volume of the weldment. This will open up the opportunity for cavities to be formed in the weldment.

In butt welding where the butted line is used as the welding line, the resulting weldment is formed from the materials taken from both sides of the butted work materials without any additional filler. Thus, attention should be given to the intimacy of the contact between the shoulder and the surface of the work materials because once the shoulder is not touching the surface of the butted materials, the confined volume is broken and some softened welding materials from the confined volume will escape through the gap between the shoulder and the surface of work materials. In addition, it must be ensured that the butted surfaces of the work materials should be flat, square and parallel to each other because there should not be any gap existed in the welding line. Thus, the work materials should be firmly fixed vertically and laterally as shown in Figure 8. These situations should be considered because the process is a non filler-addition process, which means that the lost or the shortage of the welding materials will create cavity in the weldment at the volume equal to the volume of missing or shortage of materials.

2.3. Stirring of Soft Metal

Upon the softening of the edges of work materials to be welded, the next action to be taken is to stir those materials in order to form the weldment. Thus, since the friction stir welding process is relying upon the success of the formation of stirred work materials in their soft states, a confined volume for the soft metals to be stirred should be formed. Thus, the shoulder which is positioned in an intimate contact with the surface of the work materials

with the pin is inside the metals is the ideal confined volume. The tool is then moved against the work materials, or the other way round, at a constant traveling speed along the welding line. Mechanical frictional heat, which is generated between the welding tool (shoulder and pin) and the work materials, along with, the heat generated by the mechanical mixing-shearing processes and the heat within the materials generated adiabatically are the heat sources that cause the materials to stay soft during stirring. These materials cannot reach their melting point because with this technique, there is no more additional heat from other source available to achieve melting. Thus this process is cited as a solid-state process.

As the rotating pin is traveling along the welding line, the leading edge of the pin forces the plasticized material from the leading advancing zone to enter the leading retreating zone (Figure 9). The material is then stirred with the material from the leading retreating zone and the mechanically stirred materials are then pushed to the trailing zone by the subsequently produced stirred materials while the rotating slightly slanted backward shoulder is applying a substantial forging force to consolidate the soft stirred metals produced. Thus, the welding of the stirred work materials is facilitated by severe plastic deformation in the solid state, where dynamic recrystallization of the work materials is involved [1]. After cooling, the soft stirred metals become the weldment.

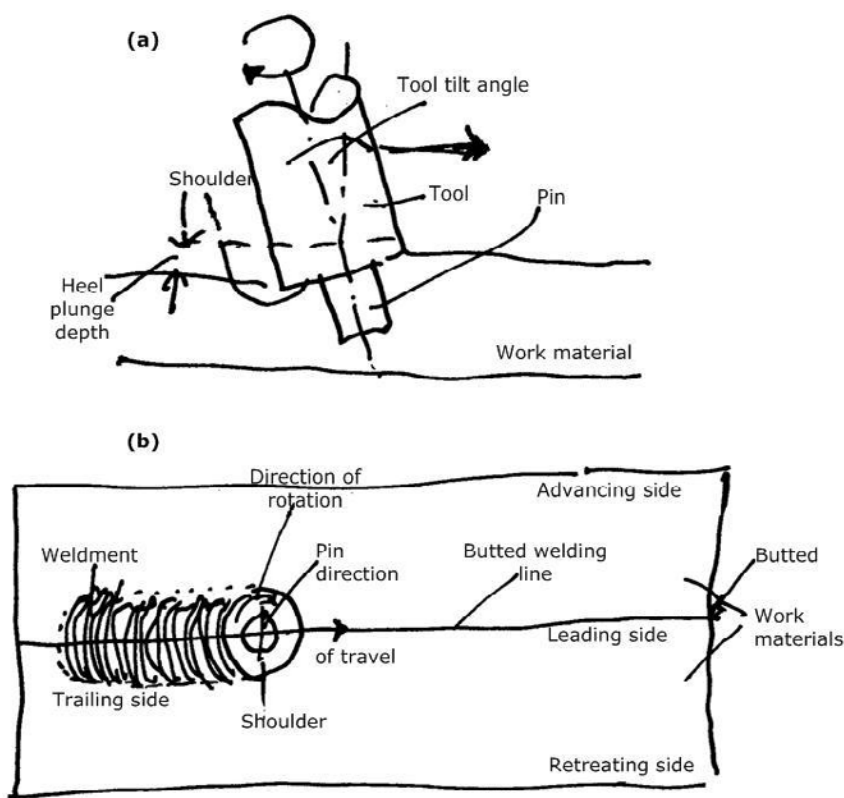


Figure 9. (a) The position of FSW rotating tool on the welding line of the two butted work materials, (b) the establishment of sides during FSW process.

3.0. OPERATIONAL CONSIDERATION

The success of friction stir welding depends on several conditions such as the dimensional accuracy of the butted edges of work materials, the cleanliness of the welding zone the degree of softness of stirred materials, the quality of the confined volume, the design of the pin and shoulder of the tool as well as the rotational and travelling speeds. The quality of the butted edges of work materials means that the edges should be pre-machined and clamped firmly to ensure that there is no gap formed between the butted two edges during the welding process being performed. This is required because the welding process is of non-filler material technique where any shortage of materials during the weldment formation will give results in the formation of cavity.

The degree of materials softness is also important because if it is too soft it means that the stirred materials temperature is too high. Although the materials temperature is still below the melting temperature, the weldment microstructure will change significantly compared to the base metals' since the grain growth, the formation of brittle intermetallic phase and phase transformation will likely to occur. The packing quality of the confined volume is determined by the ratio of shoulder and tool pin diameter. The bigger the ratio between the shoulder and the pin diameter, the bigger is the confined volume. This is good with respects to the assurance of the tightness of the confined volume. However, this will promote an extra surface contact between the tool shoulder and the work materials which can create an extra heat due to excessive amount of mechanical friction. Thus, the work materials will be too soft. The rotational and travelling speeds are two critical variables to be chosen for the welding process. The rotational speed of the tool should be set in such a way in combination with the travelling speed with the main intention to shorten the welding time without scarifying the quality of the formed weldment [2].

3.1. Work Materials Considerations

This process is cited as a solid-state process since the metallic bonding is formed at working temperature below their solidus lines. The work materials cannot reach their melting points because other than frictional and adiabatic heat sources, there is no other heat source to generate more heat. The weldment produced is formed from the materials taken from both sides of the butted work materials without any additional filler. For solid state joining, the mechanism involves in the formation of metallic bonding among the mixed soft work materials, as well as between the mixed soft work materials and the base metals is atomic diffusion.

Thus, in FSW, in order to enhance the diffusion so that a sufficient metallic bond can be formed between two butted work materials, the surfaces of work materials, with sufficient degree of geometrical accuracy, are deformed using an external mechanical force applied on a stirring tool to facilitate maximum contact between the soft work materials. At the same time, the rotating tool exposes new and fresh softened metals from both advancing and retreating sides, which are free from contaminations especially oxide films. Oxides film, dirt, oil or grease on the surfaces of the work materials, and metal inclusions in the work materials at welding area will contaminate the weldment that inhibit the atomic diffusion and

consequently limits the strength of metallic bonding. Thus, contamination reduces the quality of weldment. The work materials to be welded should also be considered for their cleanliness and their compatibilities in their melting temperatures. The welding of the work materials is facilitated by severe plastic deformation, where dynamic recrystallization [1] of the work materials is involved since the plastic deformation takes place at elevated temperature which is more than the recrystallization temperature of the workpiece materials. Thus, metallurgical aspects of FSW involve plastic deformation, diffusion and annealing of workpiece that promote recrystallization.

3.2. Tool Considerations

Since this process is non-filler technique, the tool must be able to travel smoothly along the welding line with the shoulder is in intimate contact with the surface of butted work materials, without creating any splash. Any incident of the splash will reduce the amount of soft metal to be formed as a weldment and as a result, cavity will be formed in the weldment. However, the intimate contact alone is not enough to ensure the splash-free operation. The welding tool should also be positioned tilted 2-3° backward in the traveling direction to provide volume for the agitated materials in front of the pin (see Figure 7). This edge will provide the compacting and forging actions on the materials accumulated in the rear side of the pin under the tool shoulder. Upon solidification, these materials become the weldment.

Selection of tool to be used in FSW process should be made carefully because the configuration of the tool determines the quality of weldment produced and welding speed that could be achieved. Tool is used to generate heat by mechanical friction against the work materials in order to soften the work materials. Tool is also used to provide the confined volume to accommodate the stirring process of the soft work materials in order to form weldment. Thus, tool should be made from materials superior in physical and mechanical properties compared to work materials.

The tool should be made of material that is strong, tough and hard wearing. Moreover, to minimize heat loss during welding, the tool should be made of material with low thermal conductivity. These are important features that a tool should have because it will affect the profile of the confined volume. When the tool with low thermal conductivity is used, the softened materials inside the confined volume will not stick on the surface of the pin and shoulder. This will keep the sliding interaction between the tool pin and shoulder against the work materials that provide a constant heat source sufficient to soften the work materials uniformly under the constant interfacial pressure P (Figure 10).

However, when the tool with higher thermal conductivity is used, the softened materials inside the confined volume will tend to stick on the surface of the tool pin and shoulder. Thus the geometrical shape of the tool pin will change from a fixed cylindrical to a random conical shape. It is random because the amount of materials that is sticking on the surface of the tool pin and shoulder is not constant.

In its function to generate heat required to soften the work materials, tool should have sufficient amount of surface areas to generate heat by friction. Large diameter of tool pin is good to produce an ample amount of mechanical friction in the beginning of the welding process. The large diameter of tool pin is also good for the formation of weldment where the

larger weldment will be produced. Larger weldment means larger amount of work materials that are mechanically mixed. However, the size of pin diameter should also be in good proportion with the size of diameter of shoulder.

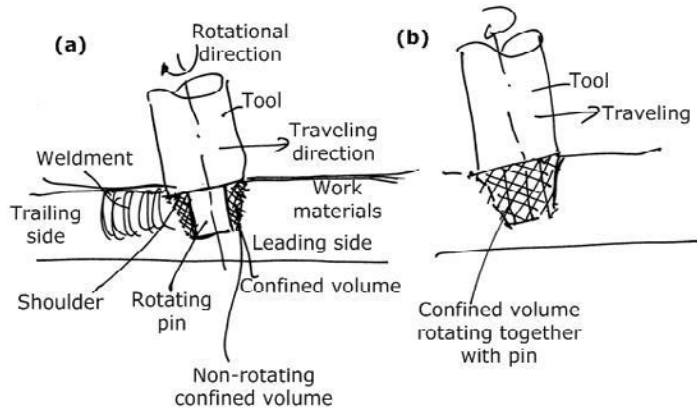


Figure 10. Tool pin profiles: a) in sliding condition and b) in sticking condition.

To increase the performance of FSW process, some effort has been made to give different geometrical shape to the tool pin [3]. Threaded tool pin, squared, triangular etc. have been claimed to improve the welding performance. For example, for the triangular pin profile, when the tool is rotated at say 300 rpm, the profile made on the tool will only create an empty volume around the profile with sweeping and mixing frequencies at $3 \times 300/60$ Hz which is equal to 15 Hz since the triangular pin has three corners (Figure 11). This means that the tool can sweep and mix the work materials from advancing and retreating side 15 times in a second and accumulate the mixture in the trailing side with in a lamellae structure with a frequency of 15 Hz.

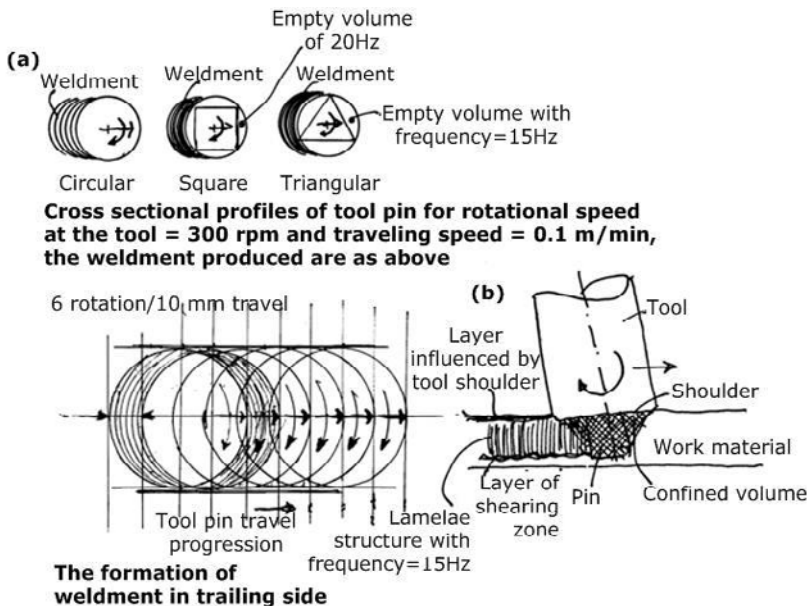


Figure 11. a) Cross sectional profiles of tool pin with its performance. b) lamellae structure of 15Hz.

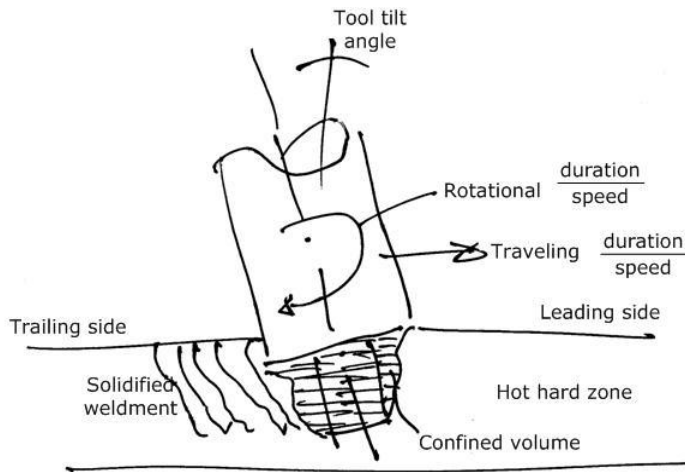


Figure 12. Location of confined volume and formation of weldment at the trailing zone.

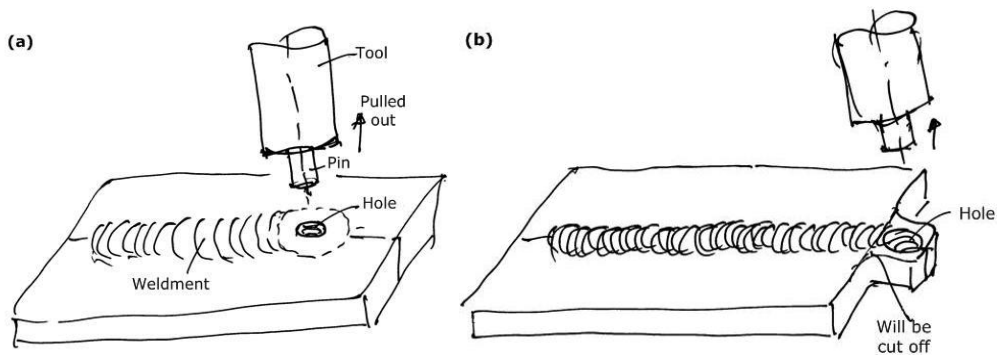


Figure 13. (a) Volume of displaced materials that is equal to the volume of penetrated pin that would never be replaced in FSW and (b) technique to locate the exit hole outside the weldment.

3.3. Confined Volume

When the shoulder makes an intimate contact with the surfaces of work materials, a confined volume is formed (Figure 12). Large diameter shoulder provides more frictional area with work materials that also provide a larger confined volume. More heat would be generated by mechanical friction as well as adiabatically in confined volume that is sufficient to keep the work materials softened. Thus, when used in high angular velocity, the temperature generated will be higher. The stirring action facilitated by the tool pin will be more intensive. The softness of the work materials being stirred will then be higher, which mean that there will be less shear strain during the formation and compaction of weldment. However, there will not be enough frictional energy nor adiabatic energy could be produced to melt the work materials. The system remains in solid state.

When the pin is inside the work materials, some equal amount of volume of work materials with the volume of the inserted part of the pin will be displaced out. Thus, since the FSW is a non filler material welding process, when the weldment is formed with the pin is

still positioned inside the work materials, the amount of displaced metal has not been replaced and would never be replaced. Thus, the end of welding process should not be performed inside the joining working area because when the tool is pulled out of the work materials, a hole of the size of the tool pin will be left on the work materials (Figure 13). Thus, to avoid this situation, the welding process should be terminated outside the welding area where the part with the hole will be cut off.

3.4. Setups and Work Material Holding

In metal joining, bonding between atom-atom which forms weldment involves metallic bonding where the free electron and positively charged metallic atoms are attracted to each other. In order to ensure this kind of bonding with stable energy is established, the atoms must be brought together into a certain distance. Thus, for a solid state joining, atomic diffusion plays an important role for the successful intermixing of the materials from advancing and retreating sides which subsequently form the weldment by metallic bonding. Similar mechanism occurs in liquid state joining. However, the intermixing of atoms in molten liquid is performed in a much easier way since the atoms can diffuse easily because the molten metal has low fluidity. In solid state joining such as forge welding, diffusion welding, friction welding or explosive welding, the workpiece interface is conditioned to expose fresh soft metal to enable atom to atom contacts. Similarly, in FSW a tool which generates heat for softening the metal workpiece is also used to intermix the soft metal by bringing the atoms close to each other to establish bonding. Thus, to achieve this condition, setups and work material holding must be carefully provided.

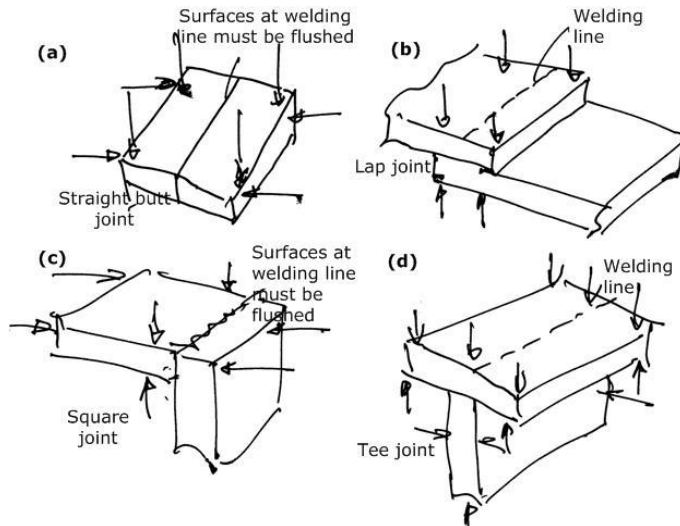


Figure 14. Work materials holding techniques to ensure successful weldment formation.

The setups of FSW are simple. As shown in Figure 14, principally in all configurations, the pair of work materials should be clamped firmly in order to make sure that the member of

the pair does not move relative to each other hence a gap will be existed between them (for joint configuration a, b, c and d) or the surfaces to be joined are not flushed (for joint configuration a and c), which will make tool shoulder fail to make an intimate contact with the surfaces of work materials, thus the confined volume becomes leaking.

3.5. The Tool Rotational and Traveling Speeds

During welding, rotational speed of the tool and the amount of surface area of the axially loaded tool shoulder in contact with the surface of work materials determine the amount of heat generated by mechanical friction to soften the work materials. The higher the rotational speed and the axial load applied, the faster the materials to be softened, and when the materials are softened, the axial load will be diminished and at a predetermined position, the rotating tool will stop to penetrate the work material. Now, the rotating shoulder, which is still in an intimate contact with the surface of the work materials, generates enough heat through the sticking-slipping interaction. The rotating pin inside the confined volume is also making contact with the softened materials inside the confined volume, which bound to also generate some heat by mechanical friction. Up to this stage, the formation of weldment has been started in a very minimum way. Under the confined volume and stick-slip interaction between the shoulder and the softened work materials, some material from advancing side has penetrated into the retreating side and mixed with the material in the retreating side and vice versa. No accumulation of the mixture in the trailing side because since the tool is not traveling, the trailing side is not existed yet.

Once the rotating tool moves leading and trailing sides are established in front and behind the rotating pin respectively. The formation of weldment is then started where the mixture of materials from advancing and retreating sides are accumulated in trailing side, which were left empty by the rotating pin when it moves forward.

Table 1. The effects of tool rotational and traveling speed on the longitudinal microstructure and welding time

Tool Speed		Weldment longitudinal microstructure produced	Welding time
Angular	Traveling		
Low	Low	Regularly structured lamellae	Long
Low	High	Coarsely structured lamellae	Short
High	Low	Fine lamellae structure	Long
High	High	Randomly lamellae structure	Short

The rotating tool travels along to form weldment behind. This is applicable for all possible FSW configurations (butt, lap, square and tee). The rotating pin inside the confined volume sweeps materials from advancing and retreating sides and the materials are then mixed and accumulated in the trailing side. The speed of travel of the rotating tool should then be arranged in such a way in order to produce well-formed weldment. This is important because the formation of well-formed weldment is dependent on the rotational speed and

traveling speeds of the tool. Thus, the combination of tool rotational and traveling speeds, quality of weldment formed, and welding time under a well proportioned tool pin and shoulder are tabulated in Table 1.

4.0. REGIONS ESTABLISHED IN WORK MATERIALS

Since the quality of weldment is determined by the microstructures produced along the welding line, it is important to map out the sides established on the work materials once the rotating pin inside the confined volume starts its travel along welding line. The sides have their own microstructure characteristics because of the difference in the nature of the work done by the rotating-traveling tool pin inside the confined volume under the rotating-traveling tool shoulder.

4.1. Sides Established during FSW

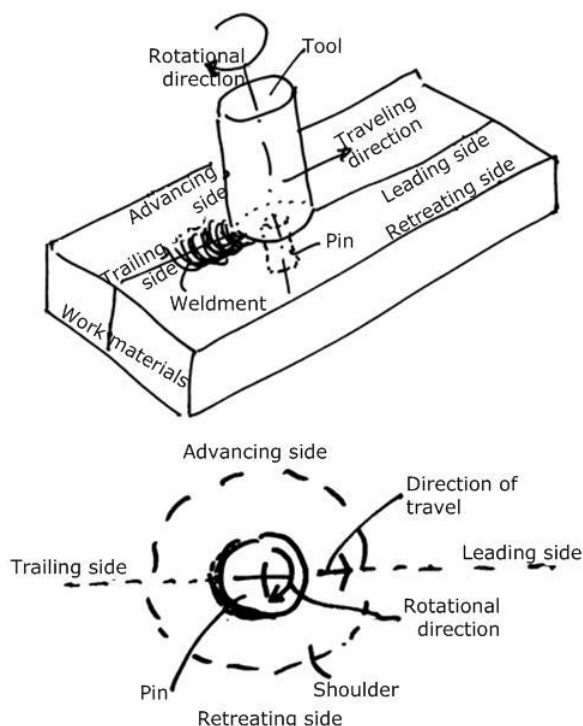


Figure 15. Map of sides locations established during FSW process with reference to the tool rotational and traveling directions.

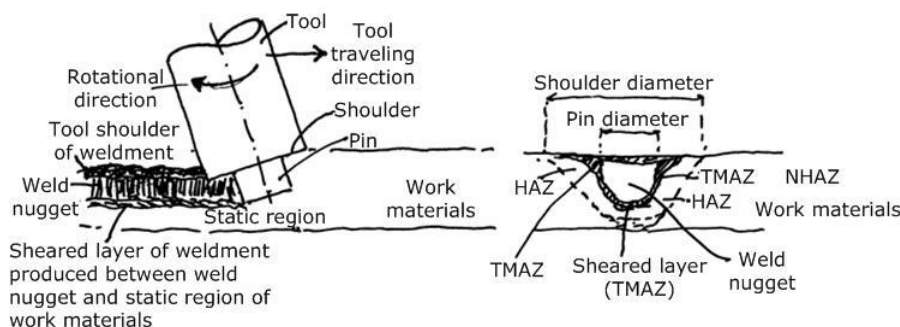


Figure 16. Formed weldment cross sectional views in longitudinal and lateral directions.

The locations of advancing, retreating, leading and trailing sides established in accordance with the combination of rotational and traveling directions of the tool along the welding line. Figure 15 shows the map of the sides locations. When the tool rotates in clockwise direction, viewed from the top of trailing side, the advancing side is located on the left hand side when the tool travels forward. The retreating side will be located on the right hand side of the tool travel direction. Obviously, the leading and trailing sides will be located in front and behind the tool respectively. They produced their own microstructure characteristics resulting from the nature of the flow of soft materials during stirring process (Figure 16).

When the rotating tool pin penetrates the softened surface of work materials, there will be no sides established yet. However, for the butt, square and tee joint configuration, it is obvious that the advancing and retreating sides will be either on the left or right side of the welding line respectively. It will be established which one is which once the rotating tool starts its travel. Thus, once the rotating tool moves forward, leading and trailing sides, which are located in the front and behind the shoulder of moving tool respectively, and the advancing and retreating sides, which are on the left and the right sides of the traveling direction of the rotating traveling tool respectively are established. The advancing and retreating sides together with the leading and trailing sides and the static side below the bottom surface of tool pin are the boundaries of the confined volume.

During the welding process, the stirring action performed by the tool pin inside the confined volume is to sweep material from advancing-leading sides and transport the material to the retreating side. In the same time, pin is also sweeping the material from the leading-retreating side and together with the transported work material from advancing side are mixed and further transported to the retreating-trailing edge of the rotating tool and accumulate the mixture in the vacant volume behind the rotating-traveling tool pin in the trailing side. The accumulated material in trailing side is recognized as the weldment formed.

4.2. Zones in Work Materials after FSW Process

The butted configuration of work materials is used to represent the zones established in welded work materials after FSW process. Due to different heat load, deformation and flow behavior experienced by both butted work materials in the advancing and retreating sides as well as in the leading and trailing edge, the zones in the cross section of joined materials (or

weldment) can be recognized as weld nugget, thermo-mechanically affected zone (TMAZ), heat affected zone (HAZ) and non-heat affected zone (NHAZ) (Figure 17).

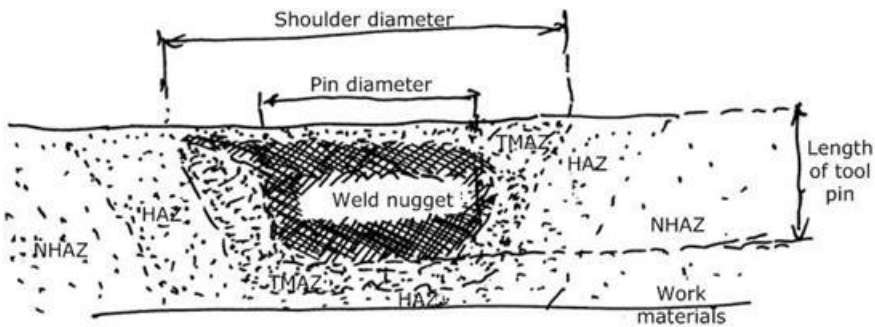


Figure 17. Illustration of zones in work materials after FSW process.

The development of the zones can be described as follows:

- 1) The weld nugget is the mixture of materials swept from advancing and retreating sides that are accumulated in the trailing side of rotation tool pin in the confined volume, under the rotating shoulder. These materials experience plastic deformation during the transportation from advancing side to retreating side through the extrusion process performed by the rotating pin in the confined volume of soft metals and passed to trailing edge. Trailing side of the shoulder then forges these materials.
- 2) TMAZ is a region built when the soft metal under the confined volume at advancing side has interaction with the confined volume, which causes shearing that produced heat. TMAZ is the interface region between the base metal and weld nugget in advancing region. The work materials in TMAZ experience less plastic deformation compared to the work materials in the nugget. Since TMAZ is the interfacing region, its size is much smaller compared to the size of nugget.
- 3) The next region that is not affected by deformation but affected severely by the propagated heat is known as heat affected zone (HAZ) that is located outside TMAZ. This area experiences the changing of microstructure and properties due to the exposure to high temperature during FSW, which induces annealing effect. This area is located between the TMAZ and the base metals.
- 4) Non-Heat Affected Zone (NHAZ) is the parts of work materials that are not affected by the heat propagated from the welding area during welding. The microstructure of the materials in these areas are not changed, remained the same as before welding process was performed.

5.0. MECHANISM OF STIRRING

The success of weldment formation is entirely dependent upon the success of stirring action performs by the rotating-traveling tool shoulder and the pin inside the confined volume of soft metal. The weldment is not going to be produced if the rotating tool is not moving in forward direction. This is because there will be no input and output of the soft metal entering

and leaving the confined volume. The quality of the input material is influenced by the amount of heat provided to soften the welding material in front of the tool pin and the speed of travel of the tool along the welding line. If the temperature in front of the tool pin is high which means that the soft metal has a higher fluidity, the stirring action performed inside the confined volume is not producing enough mechanical shearing in order to have intermixing of the work material since the shearing action in the material is less intensive compared to the mixing action performed when the material is of lower fluidity.

The stirring mechanism of materials inside the confined volume is complex. When the tool is rotating and traveling, the confined volume under the shoulder can be divided vertically into three regions: i) the intimate contact region, ii) the stirring region, iii) the shearing region and iv) the static region as depicted in Figure 18.

i). **The Intimate Contact Region** is the region of confined volume in the work materials located just underneath the rotating tool shoulder, having an intimate contact with the surface of rotating-traveling shoulder. The intimate contact can be in the mode of a) sliding, b) stick-slip, or c) sticking, depends on the heat conductivity of the tool material used, the roughness of the surface of the shoulder, the rotational and traveling speeds, and the physical properties of the work materials.

a) In the sliding mode, the degree of softness of the materials in the confined volume can still maintain a strong material bond with the material outside the confined volume. In this situation, the rotational motion of the tool shoulder does not create a rotational motion of the work materials in confined volume. In sliding mode, the geometrical shape of confined volume is similar to a hollow cylinder. When the rotating tool starts to travel along the welding line, the surface of the tool inside the hollow part starts to push the inner surface of the hollow cylinder. Mechanical friction is generated between the outer surface of the tool pin and the inner surface of the hollow confined volume. Sides on the work materials are then established. Since materials are in a confined volume, materials from the advancing side as well from the retreating side will be swept by mechanical friction performed by the tool pin. In the combinatorial effort performed by the tool pin under rotational and traveling speed, the swept materials are then stirred to form mechanical bonding between the stirred materials.

All of these operations are performed inside the confined volume and materials being transported from one side to the next through narrow slits, parallel to the axis of the pin, existed between the rotating-traveling tool pin and the softened work materials inside the confined volume. The materials from the intimate contact region will fill the empty volume left by the trailing edge of the rotating shoulder, together with the materials from advancing and retreating sides that fill in the empty volume left by the rotating-traveling pin in the trailing side. These materials form the weldment. This is an ideal condition in FSW process where the best microstructure configuration can be obtained as the indication of the best quality of weldment produced.

b) When stick-slip mode occurs, the materials in the intimate contact region inside the confined volume sometimes lost its material bond with the materials on the surface of work materials outside the confined region (Figure 19). When this

happens, the configuration of confined volume will be intermittently changes from the slipping mode configuration to the sticking configuration.

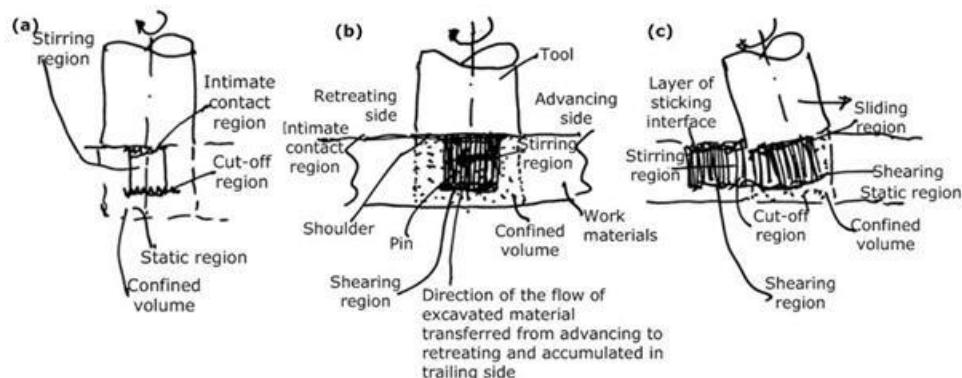


Figure 18. a) Sliding mode of interfacial contact between the surface of tool shoulder with work materials with different regions in sliding mode of FSW, b) front view and c) side view.

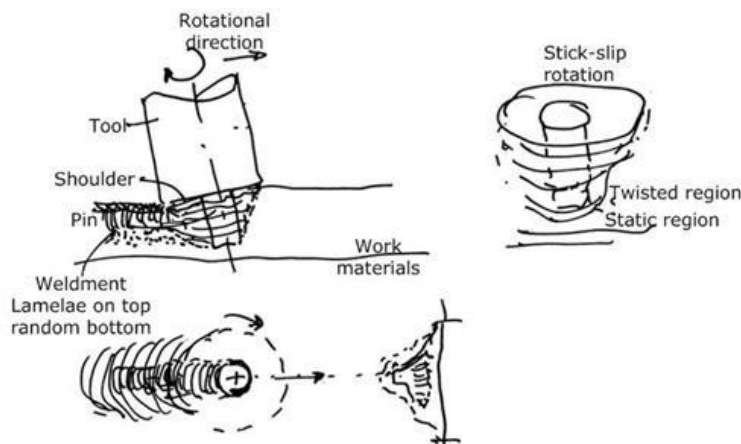


Figure 19. The stick-slip mechanism of stirring.

- (c) The weldment produced in this mode will be intermittently changes their microstructure configuration, thus the quality of the weldment is not as superior as what produced in slipping mode. The name stick-slip is given to this region because in here, the materials are rotating in the sticking–slipping conditions, depends on the state of interfacial contact between the shoulder and the surfaces of work materials. The materials in this region tend to stick to the surface of the shoulder when the coefficient of kinetic friction is less than the coefficient of static friction. The two contact surfaces will stick until the sliding force reaches the value of the static friction. The surfaces will then slip over one another with a small-valued kinetic friction until the two surfaces stick again. The simplest model for explaining this mechanism of friction, known as 'stick-slip,' is the case of a spring with a mass attached as seen in Figure 20. In this setup there is a mass attached to a coiled spring being pulled by a tension force so that the

spring moves at a constant velocity. The surface upon which the mass rests has a coefficient of kinetic friction that is much less than the coefficient of static friction.

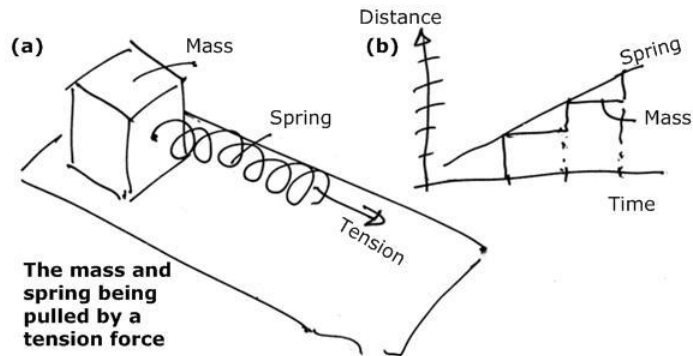


Figure 20. The mass and the spring being pulled by a tension force.

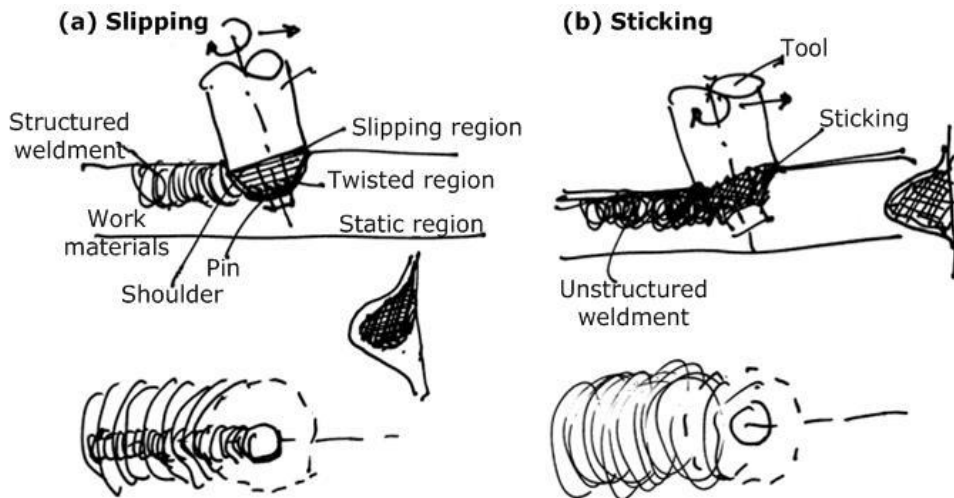


Figure 21. a) Slipping and b) sticking conditions of materials in confined volume and their resulting weldment.

The mass is pulled by using spring as a mediator in one unit of distance (Figure 20.a). When the tension is enough to overcome the force of static friction, the mass begins to move. Because the kinetic friction is far less than the static friction, the mass moves at a velocity faster than that of the spring, rapidly restoring the spring to its unstretched length. This causes the mass to once again come to rest to start the entire process over again. The mass will again remain at rest until the tension exceeds the static friction causing the block to move forward another unit of distance until the mass stops because of the compression of the spring back to its unstretched length. By performing this run at numerous spring velocities and making plots of position versus time, the trend begins to

show is that the faster the spring velocity, the motion of the mass becomes less jerky. Also, the motion of the mass becomes less jerky if the two coefficients of friction approach the same value.

The stick-slip situation is undesirable because it produces uneven weldment (Figure 21). The quality of the weldment becomes low. Thus, it would be advisable to rotate the tool at a rotational speed that the materials under the axially loaded shoulder will not stick or stick-slip to the shoulder.

- (c) In the sticking mode, the materials under the confined volume are just sticking to the rotating-traveling pin and shoulder (Figure 22). The materials are covering up the shoulder and the pin and modified the geometrical shape that acts as pin without shoulder. Although the confined volume is still existed, that is sticking to the tool shoulder and pin, but since the excavation of materials from advancing and retreating sides are performed outside the confined volume, the materials from the region where the shoulder makes the intimate contact with the surface of work materials will be splashed out of the welding zone. Since the FSW is the non-filler material process, the material deficit during the formation of weldment will leave cavities in the weldment.

(ii) The Stirring Region

In this region, the work material from advancing-leading side is swept and transported to retreating side. During transportation, this material is mixed together with work material swept from retreating-leading side by using the stirring action of the rotating-traveling tool pin inside the confined volume under the tool shoulder. The stirring action is the function of rotational and traveling speed of the tool pin (Table 1). This region has boundaries: the intimate contact region as the upper boundary, the shearing region as the bottom boundary, and the advancing, leading, retreating and trailing sides as the peripheral boundaries (Figure 23).

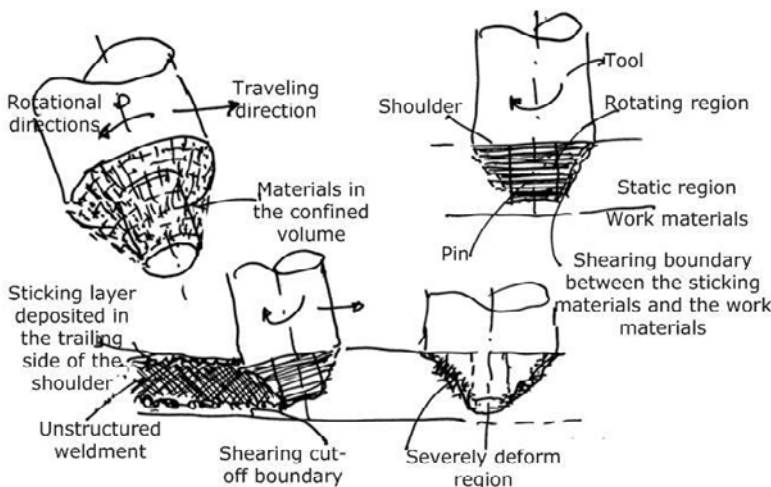


Figure 22. Sticking condition of materials in confined volume and their resulting weldment.

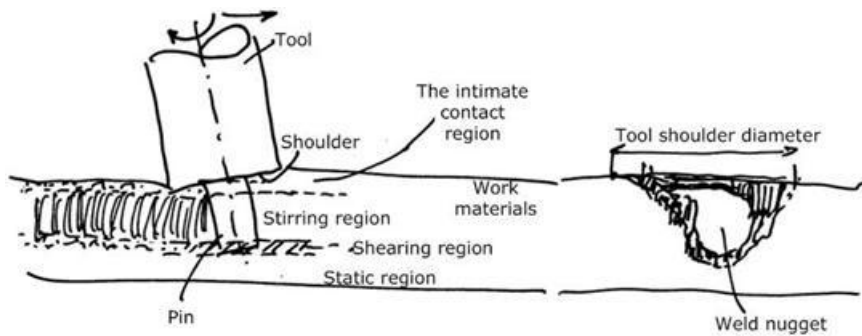


Figure 23. Regions of work materials during the formation of weldment.

This region is best existed when the materials in the confined volume is in the sliding mode, and intermittently in stick-slip mode where the materials accumulated in the trailing side have lamellae structure. In the sticking mode, the geometrically modified pin shape without confined volume is governing this region where the production of lamellae structure weldment cannot be promoted. The stirring region is difficult to be established when the length of the tool pin is not sufficient.

(iii) The Shearing Region

This is the transitional region between the stirring region and the static region where shearing process takes place between the materials being transported from advancing side to the retreating sides and accumulated in trailing side and the work materials statically present at the bottom side of the tool pin. In this region materials shearing process takes place that generate heat utilized partly to soften the materials in confined volume during welding.

When sliding mode occurs in the intimate contact region, pin-full-length materials will be swept by the rotating tool pin from the advancing side and transported through the leading side to the retreating side. Here, the rotating pin will mix the materials vertically with the material swept from the retreating side. The mixture, in lamellae structure, will then transported to and accumulated in the trailing side to form weldment. In the stick-slip mode as well as in sticking modes, the shearing region always been existed as the region of transition between the moving work materials in the confined volume.

(iv) The Static Region

The region of the work materials underneath the bottom of rotating-traveling tool pin, which are not influenced by the work done by the pin is called the static region. This region is the bottom boundary of the confined volume required for stirring process of work materials by the tool pin in order to form weldment. The depth of this region can be minimized when the length of the tool pin is about equal to the thickness of work materials. In this configuration the shearing region is eliminated. However, a backing plate should be used underneath the work materials to prevent leakages of confined volume.

6.0. FLOW BEHAVIOR

Inside the confined volume, the soft metals of two similar or dissimilar materials are mixed mechanically to form a weldment. The success of the mechanical mixing depends on several conditions such as the availability of a confined volume, processing condition, material characteristics, the transportation of the soft metals in different zones and the accumulation of the soft mixture in the trailing zone (Figure 24).

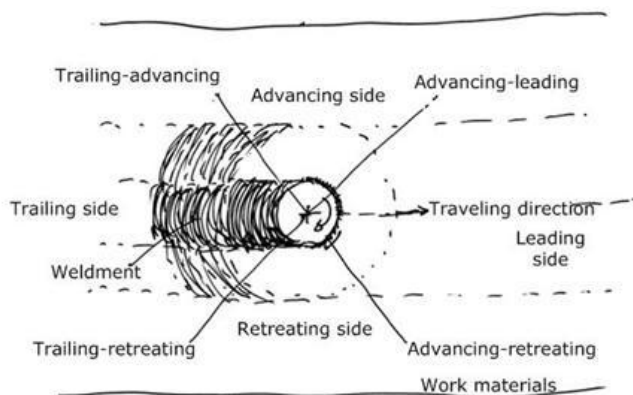


Figure 24. Transportation of soft metal in the different zones

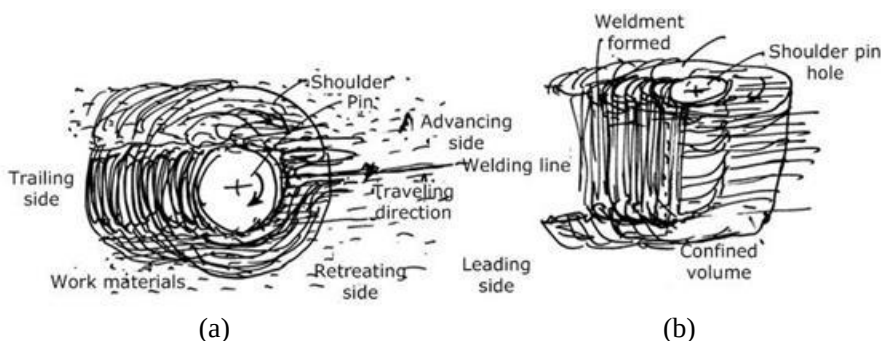


Figure 25. Typical generic flow pattern around the rotating pin in FSW: a) top view and b) side view.

However, for the success of the process, the basic requirement that should be fulfilled is the transportation of the soft metal in the confined volume. The transportation of the soft metal from the leading advancing side of the tool motion should be able to be extruded into the leading retreating side and passes through the boundary between the leading edge and the trailing edge in the manner of extrusion to reach the trailing retreating and trailing advancing sides. At the same time, the rotating tool is traveling forward leaving an empty volume behind in both trailing-advancing and trailing-retreating sides. This empty volume should be filled immediately by the materials just extruded from the leading retreating side into both retreating and advancing sides of the trailing edge. This is the process of the weldment formation.

The transportation of the soft metals by means of extrusion from leading advancing side to the leading retreating side as well as from the leading retreating into the trailing retreating and advancing sides should be made in balance since the volumetric ratio of those two transported soft metals determines the quality of the weldment. For two similar materials, the ratio of 1:1 is considered to be the best since it represents the volumetric balance of composition of the two materials in the weldment to construct layers of lamellae. This ratio could be achieved by positioning the tool pin on the welding line. However, when two dissimilar work materials with significant difference in melting temperature, the higher melting temperature material will reach its softening point when the lower melting temperature material has already approaching its solidus line. Thus, the tool pin should be positioned bias towards the work material with higher melting temperature with the expectation that the work material of higher melting temperature reaches its softening point before the other pair of lower melting temperature work material reaching its solidus line.

The continuous transportation and penetration of one soft metal to the other during the welding process build a flow pattern of those work materials (Figure 25). The ideal flow pattern of soft metal should reveal an orderly pattern of the two work materials in the weldment. This represents a balance and uniform transportation and penetration of the two work materials during welding. In contrary, the irregular random pattern of transportation and penetration of material flow represents the imbalance and different softening level of the two work materials.

7.0. WELDING METALLURGY

To produce a high integrity defect-free weldment, process variables, the tool rotational speed, traveling speed, the downward plunge force as well as tool pin design must be chosen carefully. Although FSW is a solid state process, it is also considered as a hot-working process in which a large amount of deformation is imparted to the workpiece through the rotating-traveling pin and shoulder. Such deformation gives rise to a weld nugget, which is comparable to the diameter of the pin, a thermomechanically-affected region (TMAZ), which is comparable to the diameter of tool shoulder, and a heat-affected zone (HAZ), which lays out of the weldment formation region. Frequently, the weld nugget appears to comprise equiaxed, fine, dynamically recrystallized grains whose size is substantially less than that in the work materials used. However, the evolution of microstructure in the dynamically recrystallized region and its relation to the deformation process variables such as strain, strain rate, and temperature should be well understood.

7.1. Materials Interactions

Although both friction stir welding (FSW) and friction welding (FW) are solid state welding processes without additional filler material, both processes have different working principle. In FSW, three surfaces of different materials, which are two work materials and one tool, are put in contact to generate heat, which is quantified by the expression (1) and (2).

Once the heat generated reaches the softening point of either one material, the pressure diminishes upon the deformation of that material. If both work materials have the same

melting temperature, both deform equally. And if they do not have the same melting temperatures, the work material with lower melting temperature deforms first. When the pressure, P is continuously applied, the deformation extends.

At the interface between the bottom surface of the rotating tool pin and upper surface of the work materials, heat is generated. At this stage, the plastic deformation of the work materials starts to take place due to the applied pressure through the axis of the tool on the surface of the work materials. Joining of the two work materials cannot be accomplished yet because the tool pin penetrates the weld material in a non-confined volume, which does not produce the intended weldment. This is because of the material escape from the welding area cannot be compensated since the principle of this welding process is the non-filler welding technique. A confined volume should be created using a shoulder, which is made of the same material with pin tool that has bigger diameter enough to cover up the escaping some area of the soft metal.

The pin is then penetrates further into the work materials until the shoulder makes an intimate contact with the surface of the work materials. The confined volume where the work materials are stirred is then created. Work material, which is in contact with the area of the shoulder, is the most affected part by the rotational action of the shoulder and the pin. Since the pin rotates at the same rotational velocity with the shoulder, the pin is not going to influence significantly the work material in the confined volume.

The temperature of work materials in the stirred confined volume increases due to heat generated through shearing during mechanical friction between shoulder and work materials. In addition, strain energy stored during plastic deformation also contributes to the increase of temperature [4]. As the results, diffusion, hot working and annealing that include recovery, recrystallization and grain growth as well as heat treatment occur. Once the level of thermo-mechanical in the soft stirred zone increases, the atomic diffusion at frictional interface starts to take place. However, because the temperature is too low, the rate of diffusion is low in order to promote bonding between the butted work materials. Due to high rotational speed and axial force of the tool applied on the surface of the butted work materials, the work materials reach its softening point. At this stage, the tool loses its pressure on the surface of the work materials because the work materials are softened. Thus, there should not be any significant temperature increases anymore since the frictional mechanism of the mechanical interface is lost.

In a confined volume, the tool shoulder in combination with the tool pin agitates the soft metal in circular manner. This is a stirring process that causes the soft metal close to shoulder surface is severely deformed plastically. Since the soft metal has high ductility, the soft metal is plastically deformed under the influence of stirring action. This plastic deformation process occurs due to the generation of dislocation in grains of soft metal and its movement at a certain slip planes and directions. This deformation diminishes away proportional to the distance from the surface of the shoulder to the part of the work materials that is down far below which is not under the influence of the rotational shoulder. The materials here remain static.

The pressure applied in confined volume causes localized deformation in the form of lamellae, which consists of alternate layers of the two stirred work materials. These layers of lamellae increase the contact area of the interface, which improves the intermixing bonding. However, at this point, the layers of lamellae contain isolated voids separated by areas of intimate contact. In the subsequent stirring of the materials, diffusion of atoms at the interface

of lamellae layers is transferred across their boundaries. As a result, the elimination of voids at interfaces occurs which gives the significant improvement of microstructures.

With the increase of temperature, atomic diffusion in the interface of two similar or dissimilar work materials in the stirred soft zone of the confined volume is accelerated. This is because the diffusion which involves a transfer process of atoms is activated by the heat generated during rotational friction process of the work materials. The formation of soft metal during stirring also facilitates atomic diffusion. This is because metal bonding in the soft metal is weaker compared to metal bonding in hard metal, where the weaker the bonding the easier the diffusion is. The diffusion process may occur in crystal lattice of the work piece as well the grain boundary. The grain boundary represents crystal imperfection due to mismatch in atomic arrangement which makes the diffusion of different atoms occurs easily. The stirring action in the soft confined volume promotes the increase of the number of lattice defects due to severe plastic deformation caused by mechanical friction. Thus, diffusion in lattice and grain boundaries improves the bonding between the lamellae layers of two similar or dissimilar materials, which consequently enhances the strength of the weldment.

However, when the temperature of softens metal reaches about $0.6T_m$, the plastic deformation process takes place at hot working stage [4]. At this stage, annealing which consists of recovery, recrystallization and grain growth becomes significant. In addition, in the case of friction stir welding of alloy either for two similar or dissimilar metals, the heat generated in the work materials provides a condition similar to metal heat treatment that might promote the formation of the solid solutions, secondary phase precipitates, brittle intermetallics and low melting point eutectics as well as phase transformation in the weldment microstructure.

7.2. Idealization of Weldment Formation

In FSW, the formation of the weldment of two work materials is achieved using friction as the heat source and stirring action to produce the weldment by mechanically mix the work materials without filler at temperatures below the solidus line of the work materials. The content of the weldment should consist of the two work materials in equal amount and distribution. In butted work materials configuration this could be achieved by carefully positioned the pin in the middle of the welding line. However, when it is desired to have a different proportion of materials, the pin can be positioned biased towards the desired dominant material.

Weldment produced should be consisted of work materials being welded mixed uniformly in a regular fine pattern. It means that in every unit of distance of tool travel, the tool pin should be able to perform large number of fine sweepings of work materials from advancing and retreating sides. These deformed fine slices are then accumulated in a regular fine pattern in the trailing side and when solidified, they become the weldment.

This is an ideal situation to produce good bonding between the materials from advancing side and retreating side because both materials will have maximum opportunity to be exposed to each other with the maximum possibility to be diffused to form metallic bonding. However, the brittle intermetallics phase may exist if the diffusion is excessive which is unfavorable for high performance weldment [5].

When the rotational speed is low and the traveling speed is high, every distance of tool travel will consist of small number of tool rotations. The work materials taken from the leading advancing side will be in a large amount, and transported to the leading retreating side where they are mixed and transported to the rear side of the rotating-traveling pin. The mixture will not have enough opportunity to be mixed uniformly even if they can produce a regular pattern. In this situation a good bonding is unlikely to be achieved.

Since this process does not involve solidification of fused metals, thus the microstructure of the weldment is almost similar to the base metal. Although the process involves the transportation of swept materials from advancing side to retreating side as well as from retreating side and transport both of the materials to the trailing side, those materials should not be heavily deformed otherwise it will have a totally different microstructure and mechanical properties compared to their base metals. This situation will not provide good weldment properties. Thus, weldment with good mechanical properties is expected if the mixed materials achieve a good bonding with the base metals both in advancing and retreating sides and also free from defect.

However, in real situation, a lot of other factors should be taken into consideration such as coefficient of thermal expansion, thermal diffusivity, residual stress of the work materials before the welding operation, tool pin and shoulder wear, difference in physical properties in dissimilar metals, time consumed during the operation that will change the materials set up and materials properties from the beginning until the end of the operation, the size of the work materials against the size of the tool and etc.

7.3. Weldment Microstructure Development

The characteristics of microstructures developed across the weldment can be used to identify zones affected by heat and severe deformation which determine the quality of the weldment. Based on the work carried out by the tool shoulder and the pin to soften and stir the butted work materials in order to form the weldment, the zones can be identified as stirred zone, which also known as nugget, thermo-mechanically affected zone (TMAZ), heat affected zone (HAZ), and the non-affected base metals. The identification of these zones is made on the purpose to correlate the characteristics of the microstructures to mechanical properties of the joint.

For small volume work materials, temperature generated by the mechanical friction of the welding tool and work materials will immediately dissipated across the volume of work materials. Thus, there is no gradient of temperature in the work materials because the heat already affects all the materials. In this case, there will be no non-heat-affected zone in the materials since all of the volume of the work materials is already affected by dissipated heat. Consequently, all the differences in microstructure developed are not the function of the differences in the exposure of high temperature instead they can be recognized based on the deformation induced by the tool shoulder and the pin during the welding process.

However, in the case of large volume work materials where the heat generated will be dissipated to large volume of work materials, and since the work materials are exposed to the atmosphere, the heat will also be dissipated to the atmosphere. Thus, the farther the materials from the heat source, the lower the temperature will be. In this case, the gradient of

temperature exists. Thus, the different in the microstructure could be described according to the length of exposure time toward high temperature and to the straining due to heavy deformation with the intensive shearing during welding.

To explain the development of the weldment microstructure during FSW process, the mechanism of weldment formation inside the confined volume in the trailing edge should be well understood. While traveling along the welding line, the rotating pin inside the confined volume makes a heavy friction with the material in advancing-leading and retreating-leading sides, which gives result in localized heat generation that soften the edges of both butted work materials. The heavy friction promotes sweeping of the base metal from the advancing-leading side, deforms and transports the materials to the retreating-leading side across the welding line (Figure 26).

During transportation, the soften metal is being extruded through a narrow slit formed by the leading front of the rotating-traveling tool pin and the materials in the advancing-leading and retreating-leading sides. Then, the swept-transported material from the advancing side enters the leading-retreating side and mix with the swept softened material. Both are then mixed and transported together in the leading-retreating (L-R).

The materials from both advancing-retreating leading side will then be further transported and accumulated into an empty volume in the trailing side left by the rear side of the traveling rotating pin. Then, the rotating-traveling shoulder performs the forging action on the soft accumulated materials in the trailing side (Figure 27). This is the beginning of the weldment formation.

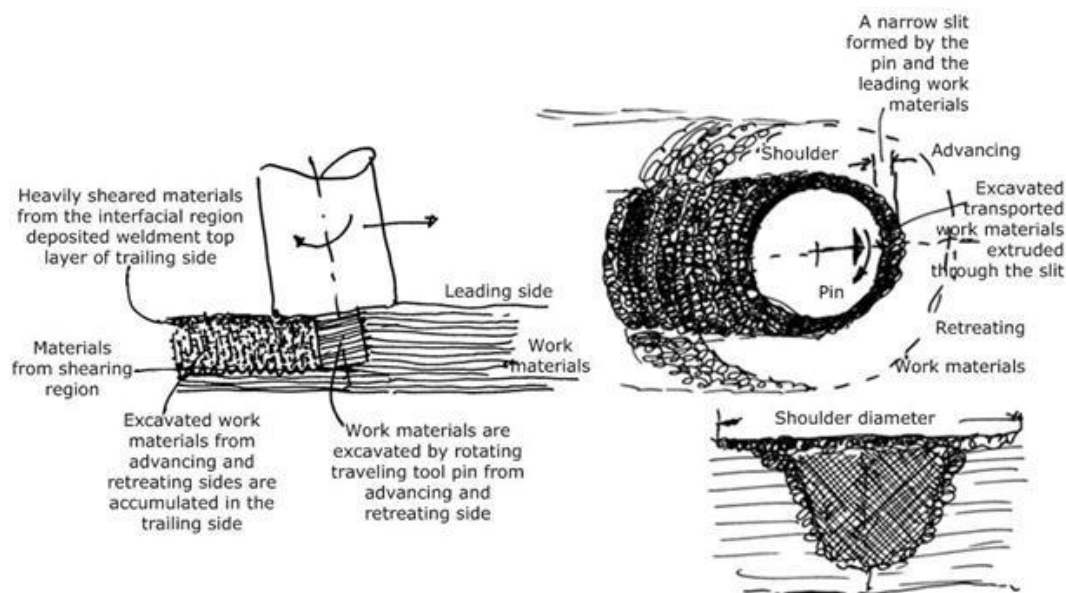


Figure 27. The generation of flow pattern and FSW weldment in the trailing edge.

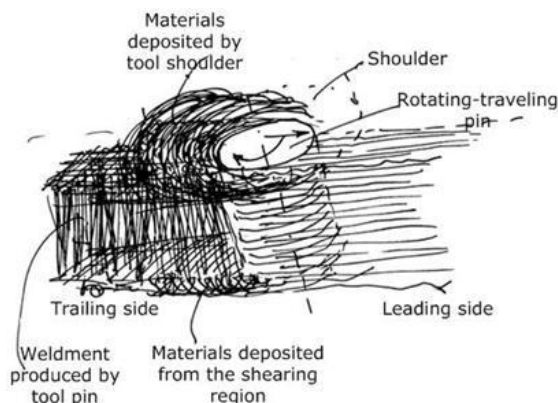


Figure 27. Transportation and accumulation of metal mixture to form a weldment in the confined volume.

The nugget is the accumulated advancing-retreating mixture of materials that experiences plastic deformation during the transportation of the materials through the extrusion and forging actions by rotating traveling tool pin and the trailing side of the shoulder. Since the plastic deformation severely occurs, critical recrystallization temperature of the material in the nugget becomes low. Thus, the temperature during the FSW process is sufficient to promote nucleation of new grains, which are fine and equiaxed.

In an ideal case, the mixture of materials should be constructed of regular series of lamellae of alternating layers of material advancing and retreating. This situation can be made when the rotating speed of the tool shoulder and the pin is in harmony with traveling speed of the tool along the welding line at a certain degree of softness of the material advancing and retreating. This is an ideal situation because the alternating layers of lamellae will diffuse one to the other. The weldment produced in this circumstance will have mechanical properties about the average of the two advancing and retreating material.

Ideally, upon the synchronization of the rotational speed and traveling speed, the mixture will be structured in such a way to form a series of layers called lamellae. However, in reality, since the materials inside the confined volume under the shoulder of the tool will be influenced by the rotational motion of the shoulder, whereas at the base the work materials are static, the upper part of the weldment will be formed by the mixture of the advancing and retreating materials in random fashion, the middle part with layers of lamellae, and the bottom part with poor mixture of work materials from advancing and retreating sides.

Since the way of the materials being transported along the welding zones is different, there will be differences in the rate of grain growth in the weldment at advancing and retreating sides.

Besides the nugget, there is another affected zone along the welding line, which is recognized as TMAZ. TMAZ has larger and elongated grains, compared to the grains in the nugget since it experiences less heating and deformation. This zone is created as a result of heavy shearing between the rotating pin under the shoulder and the static neighboring material at the base metal. In the advancing-leading zone, the rotating pin will push the material into the retreating-leading zone. If the material in the advancing-leading and retreating-leading zones is already soft enough, the extrusion process is not going to happen because the material in the advancing-leading and retreating-leading zones under the rotating

shoulder will follow the rotation of the pin and the shoulder. But since the rotational speed of the shoulder is not transferrable to the lower part of the bottom of the pin, where the materials there can be considered static, thus, there will be a heavy shearing of those materials underneath the shoulder with the base metal in the advancing and retreating sides. Extrusion process will occur between the confined volume of materials, which is rotating under the shoulder with the base metal. The tool pin, which is made of hard metal with low coefficient of heat conductivity and a determined cylindrical shape and surface, the shearing caused by the pin leaves insignificant markings of TMAZ. In this condition, the rotating shoulder is responsible to generate heat to keep the materials inside the confined volume remains soft. The rotating pin takes the responsibility to transport the materials from the advancing - leading side to retreating - leading side and upon the traveling motion of the pin, the volume left behind the pin will be the volume for the transported materials to be accumulated and becomes the weldment. The materials taken from the advancing - leading side transported to the retreating - leading zone will cause shearing with the base metal. Since the materials in the confined volume are not under the influence of rotating shoulder i.e. the material is not rotating together with the shoulder, thus the development of TMAZ becomes less significant. This condition can occur if the heat generated by the pin is not sufficiently high to cause the excessive heat that will soften a larger amount of material in the confined volume. The size of the cross section of the nugget produced is approximately the size of the cross section of the pin with the grain size much finer than in its peripheral materials.

However, under a certain circumstances when using tool material with high coefficient of conductivity, excessive heat produced will cause larger amount of work materials stick to the tool shoulder and the pin during tool rotational-traveling motion. In this situation, the nature of the weldment formation changes since the sticking work materials have modified the geometrical shape of the pin. The sticking materials will act as a pin with modified shape. In this particular case, the outer circumferential part of the rotating-traveling modified pin will take the role to transport the material from advancing side to retreating side and transport the mixture of those materials into the vacant volume left by the modified pin in the trailing side. When the outer diameter of the modified pin becomes larger than the outer diameter of the shoulder, the system lost the confined volume. The system fails to perform its FSW function.

If the surrounding metals close to the heat source are being affected thermally by the propagated heat but not affected mechanically by shearing of mechanical friction, these areas are usually known as heat affected zone (HAZ) which are located next to the TMAZ. Materials in these areas will experience the changing of microstructure and properties due to the exposure to high temperature. Normally this zone has coarser and equiaxed grain due to annealing effect during FSW.

7.4. Weldment Properties

Temperature gradient is not the main factor in the properties of weldment since the heat generated by mechanical friction by the stirring tool is immediately dissipated to soften the work materials inside the confined volume. The heat is also dissipated throughout the work materials. If the work material has a small volume which mean it has a smaller capacity to contain the heat. Thus, the heat dissipated from the confined volume will be dissipated and

the work materials temperature increase to the level that closed to the temperature of confined volume of the nugget. Since the temperature gradient is no longer the issue, the evolution of the microstructure could be described according to the length of exposure toward high thermal and the straining due to heavy deformation with the intensive shearing during welding.

In the case of work materials are butted, the top surface of those two materials should be positioned flashed to each other. This is to ensure that the shoulder of the tool, when makes an intimate contact with the work materials, will be able to produce a confined volume underneath where the weldment is formed. The intimate contact should be maintained throughout the welding process because if the shoulder lost its intimate contact, the materials inside the confined volume may escape through, and as the result, since the process is non-filler, cavity might be existed in the weldment. Thus, physically, the weldment made in FSW is flat, flushed with the work materials. Microstructurally, the weldment consists of nugget, TMAZ and HAZ (see Figure 28). Each one of these has its own characteristics as follows:

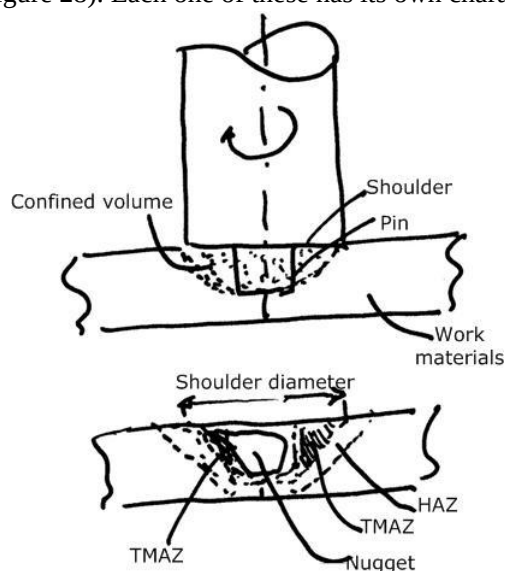


Figure 28. Locations of three different zones in the FSW weldment.

Nugget: This is the weldment that consists materials from both advancing and retreating sides accumulated in trailing side of the tool pin. During the transportation from the advancing-leading side to retreating-leading side and then accumulated in trailing edge, the materials undergo severe plastic deformation followed by recrystallization with limited grain growth. Thus, the nugget has fine and equiax grains, which provide highest mechanical properties such as hardness and tensile strength. However, there is a possibility that the nugget has lower hardness and during mechanical testing, the failure occurs in that zone. This situation happens if the friction is too high that generates excessive heat, which causes the lowering in dislocation density relative to the other zones.

Thermo-mechanically affected zone (TMAZ): The materials, which are taken from both sides of work materials by the rotational and traveling motion of the pin and the shoulder along the welding line, are the materials used to form the weldment. When the materials are

swept from both sides of work materials by the rotating pin and the rotating shoulder in confined volume, the regions of separation are developed under heavy metal deformation and exposure to high temperature. This area has its own microstructure signature and is known as thermo-mechanical affected zone. The grains in this zone are deformed and elongated with the size coarser compared to grains in nugget because recrystallization does not occur. The grain size is smaller compared to grains of HAZ, thus TMAZ should have higher mechanical properties compared to HAZ. TMAZ is not necessarily presence in all weldment. Typically, it is presence in copper alloys and steel but not in aluminium alloys. In some cases, it is difficult to make a distinction between TMAZ and HAZ.

Heat Affected Zone (HAZ): In general, the heat affected zone experiences microstructural changes because grain growth and or dissolution of precipitation particles may occur. If cold working have previously been used to harden the work materials before FSW, recrystallization followed by grain growth will result in softening. During the grain growth, final grain size of work materials in HAZ will depend on peak temperature and time for cooling. The longer the cooling time, the more grain growth will take place. Similarly, alloys that are hardened by precipitation will usually be softened by FSW due to dissolution of precipitated particles as it has been exposed to high temperature.

8.0. JOINING OF TWO SIMILAR AND DISSIMILAR MATERIALS

Although techniques used to join two similar or dissimilar metals are the same, the results obtained will depend on the metallurgical characteristic of the two work materials. The joining of similar and dissimilar work material will be elaborated below.

8.1. Joining Two Similar Materials

If the process is performed below T_m , the joint will be made in solid state where diffusion is more likely to take place. However, the diffusion itself is not enough if higher strength of joint is sought. Thus, a better way of joining should be found. One way to achieve a higher strength of joining is to intermix the material in such a way until elements of materials can be self-locking and diffused. These requirements can be achieved by FSW.

When the work materials are of similar materials and upon the softening of the material underneath the shoulder, these materials are mechanically mixed together and form a hollow cone. This is the beginning of the formation of the weldment. When the work materials are of two similar materials, the behavior of the materials will be similar such as they will be softened in about the same time. However, the flow behavior of the softened work materials will not be similar because the rotational and traveling directions are different for the materials in advancing and retreating sides. This will be clearly seen in the formation of weldment where the materials are accumulated in the trailing side of the pin will be mostly pushed to the advancing side by the rotational direction of the tool. Moreover, the backwardly tilt angle of the tool will reduce the volume of confined volume in the trailing edge. Thus, work materials will be compacted in this region and since the flow of materials in this region is in the direction from retreating side to advancing side, the accumulation of materials will be

started from the retreating side. Hence during the accumulation of materials into the trailing the weldment will have a finer microstructure.

8.2. Joining Two Dissimilar Materials

When joining dissimilar metals, the two materials have different physical properties such as T_m and hardness which will affect their ability to be intermixed because of the difference in flowability. Moreover, the diffusivity of the two different materials is more difficult and complex compared to two similar.

Diffusion is likely to occur when two materials are in contact under high pressure and high temperature. The higher the pressure and the higher temperature are, however, during the diffusion, formation of new phase as a results of phase transformation, intermetallic formation on eutectics alloy, might come up as products that influence the characteristic of the intended joint especially in the presence of intermetallic compound (IMC) which is brittle, that will deteriorate the mechanical strength of the weldment [6].

When the materials are made of two dissimilar metals, the difference in hardness or T_m influences the flow behavior of the soften materials inside the confined volume under the shoulder. As has been described previously, the flow behavior of the work materials inside the confined volume shows that the sweeping of work material from leading-advancing side by the tool pin in its cooperation with tool shoulder and transport the material to the leading–retreating side can be performed successfully if the material from the advancing side is made of higher T_m or higher hardness (which means higher viscosity) compared to the material occupies the leading-retreating side because of the less resistance made by the materials in retreating side to the incoming materials.

When the higher hardness or higher T_m work material is placed on the advancing side with the lower hardness or lower T_m work material is in the retreating side, during FSW process, the lower T_m work material will be softer than the higher T_m work material. Thus, the material which is transported from advancing side will enter the region of less viscous metal and the mechanical mixing between these material will take place and subsequently, the ‘room’ left by the tool pin in the trailing-advancing side will be filled up easily by this mixture. As a consequence, when the material in advancing side will be transported in a big amount because the viscosity is less in the retreating side, it will generate a lot of shearing action between the materials under the shoulder (in the confined volume) with the base metal in front of the confined volume while the tool is travelling along the welding line. The effect of this situation can be seen in the microstructure built-up in the cross section of the weldment where it can be seen a significant amount of TMAZ on the advancing side and a slight TMAZ developed in the retreating side. In contrary, if the lower hardness or T_m material is positioned in the advancing side, it will not be easy for the softer material from the advancing side to enter the harder material in the retreating side. In this case, since the amount of the material transferred from advancing to retreating side is not much, the room that empty in the trailing side will be a small one to be filled up with mixture of materials from advancing and retreating sides.

SUMMARY

Operating under the solidus line, friction stir welding (FSW) opens up varieties of new application where two similar or two dissimilar materials are required. Careful attention should be given to the material properties of the work materials as well as the tool material. Tool material should have a low heat conductivity to prevent sticking of confined volume on the tool. Once the work materials are sticking, the weldment process could be considered as a failure. This is because for sure, cavities will occupy the weldment since a lot of materials escaped from below the rotating-traveling tool shoulder once the confined volume is lost.

To produce a good weldment, the rotating speed of the tool should be combined carefully with the traveling speed. This combination will influence the way work materials are being swept, extruded and accumulated in the trailing side in order to form the weldment. However, once process parameters are set for the welding operation, process variables should be monitored. In FSW, process variables that are important to monitor under given rotational and traveling speeds are the pressure built up in advancing, leading, and retreating sides. Obviously, the pressure will be higher in the retreating side compared to other sides because in the retreating side, swept work materials being transported and inserted into the retreating side will cause a pressure increase. As a consequence, temperature will increase and the work materials inside the confined volume will be too soft. This will not be favorable because it will not give a good quality solid weldment. Thus, careful attention should be given during the welding process to ensure that the pressure and temperature in confined volume is sufficient to produce good weldment.

ACKNOWLEDGMENT

The authors are pleased to acknowledge financial support from Universiti Sains Malaysia under Research University Grant account 1001/PMekanik/814084. Special thanks are extended to Normariah Che Maideen and Emee Marina Salleh for their help in preparing the manuscript.

REFERENCES

- [1] Liu, G; Murr, LE; Niou, CS; McClure, JC; Vega, FR. *Scr. Mater.*, 1997, 37, 355-361.
- [2] Peel, M; Steuwer, A; Preuss, M; Withers, PJ. *Acta Mater.*, 2003, 51, 4791-4801.
- [3] Padmanaban, G; Balasubramanian, V. *Mater. Des.*, 2009, 30, 2647-2656.
- [4] Zettler, R. WTSH. In *Friction Stir Welding: From basis to applications.*; Lohwasser D; Chen, Z; Ed.; CRC Press LLC and Woodhead Publishing Limited: Boca Raton, FL, 2010; pp42-68
- [5] Nandan, R; DebRoy, T; Bhadeshia, HKDH, *Prog. Mater. Sci.*, 53, 2008, 980-1023.
- [6] Abdollah-Zadeh, A; Saeid, T; Sazgari, B. *J. Alloys & Comp.*, 2008, 460, 535-538.

Chapter 5

PLASTIC LIMIT LOAD SOLUTIONS FOR HIGHLY UNDERMATCHED WELDED JOINTS

Sergei Alexandrov

A.Yu. Ishlinsky Institute for Problems in Mechanics, Russian Academy of Sciences,
Moscow, Russia

ABSTRACT

Limit load is an essential input parameter in many engineering applications. In the case of welded structures with cracks, a number of parameters on which the limit load depends, such as those with the units of length, makes it difficult to present the results of numerical solutions in a form convenient for direct engineering applications, such as flaw assessment procedures. Therefore, the development of sufficiently accurate analytical and semi-analytical approaches is of interest for applications. The present paper deals with limit load solutions for highly undermatched welded joints (the yield stress of the base material is much higher than the yield stress of the weld material). Such a combination of material properties is typical for some aluminum alloys used in structural applications.

1. INTRODUCTION

Limit load is an essential input parameter in many engineering applications such as metal forming analysis (Avitzur, 1980) and flaw assessment procedures (Zerbst *et.al.*, 2000). The upper bound theorem is a convenient tool for finding an approximate value of limit loads. A review of limit load solutions for cracked structures made of homogeneous material has been given in Miller (1988). Welded joints can be treated as piece-wise homogeneous structures. Let $\sigma_0^{(B)}$ be the yield stress in tension of base material and $\sigma_0^{(W)}$ be the yield stress in tension of weld material. The ratio $M = \sigma_0^{(W)} / \sigma_0^{(B)}$ is called the mis-matched ratio. The welded joints can be conveniently divided into two groups; namely, undermatched joints for which

$M < 1$ and overmatched joints for which $M > 1$. It is also advantageous to separately consider highly undermatched joints. A distinguished feature of highly undermatched joints is that plastic deformation in such joints is wholly confined within the weld material, whereas the base material remains elastic. In general, the value of M has a great effect on the magnitude of the limit load. Moreover, the development of efficient analytical or semi-analytical methods of solution significantly depends on the type of joint (undermatched or overmatched). However, the present chapter reviews upper bound limit load solutions for highly undermatched welded joints only. Therefore, the value of M has no effect on the solution. However, it should be low enough to ensure that plastic deformation is wholly confined within the weld material. This condition cannot be verified by solutions in which this condition is included as an assumption. However, it is always possible to specify such a low value for M that the condition in question is satisfied. In engineering applications, it is in general necessary to find the limit load solutions with and without the assumption that plastic deformation is wholly confined within the weld material. Then, the upper bound theorem allows one to choose one of these solutions. The value of $\sigma_0^{(B)}$ is not involved in the solutions for highly undermatched welded joints because there is no plastic deformation in the base material. Therefore, to simplify writing, σ_0 stands for $\sigma_0^{(W)}$ throughout this chapter.

Special attention is devoted to efficient non-standard methods for constructing kinematically admissible velocity fields that account for some features of real velocity fields in highly undermatched welded joints.

A number of upper bound limit load solutions for the configurations considered in this chapter, but at $M > 1$, have been proposed in Joch *et.al.* (1993), Alexandrov and Goldstein (1999), Alexandrov *et.al.* (1999^a).

Plastic anisotropy has a great effect of the magnitude of limit loads for both undermatched and overmatched welded joints (Capsoni *et.al.*, 2001^{a,b}, Alexandrov and Gracio, 2003, Alexandrov and Kontchakova, 2004, Alexandrov and Kontchakova, 2005, Alexandrov *et.al.*, 2007, Alexandrov and Tzou, 2007, Alexandrov *et.al.*, 2008, Alexandrov, 2010). Nevertheless, this material property has not yet been accounted for in flaw assessment procedures. On the other hand, kinematically admissible velocity fields used for structures made of isotropic materials are also applicable for those made of anisotropic materials. Therefore, anisotropic limit load solutions are not discussed in the present chapter.

2. PRELIMINARY REMARKS

The upper bound theorem for rigid perfectly plastic materials can be found in many textbooks and monographs on plasticity theory, for example Hill (1950) and Kachanov (1956). Its generalization on quite a general rigid plastic material model is given in Hill (1956). It is worth noting that the upper bound solutions found by means of the theorem for rigid plastic solids are applicable for the corresponding elastic-plastic solids (Drucker *et.al.*, 1952). For rigid perfectly plastic solids the functional for minimization that follows from the upper bound theorem depends on the yield criterion. In the present chapter Mises yield criterion is adopted. The upper bound theorem allows one to evaluate one scalar quantity. If a single load is unknown, rigid perfectly plastic solutions provide upper bounds on this load. If

several loads are unknown, rigid perfectly plastic solutions provide upper bounds on a combination of these loads. The magnitude of velocity is immaterial in the case of rigid perfectly plastic solutions. It is worth noting that in the case of other material models it is not always possible to extract an upper bound on the load applied from the scalar quantity that can be evaluated from the upper bound theorem (Alexandrov, 2000; Alexandrov and Goldstein, 2005; Tzou and Alexandrov, 2006). For such models, the magnitude of velocity has an effect on the load required to deform material. The present chapter solely deals with rigid perfectly plastic solids. Therefore, $\sigma_0 = \text{constant}$.

Let (σ_{ij}, ξ_{ij}) be the stress and strain rate in a rigid plastic mass of volume V which is loaded by prescribed external stresses F_i over a part S_f of its surface, and by prescribed velocities over the remainder S_v . In the case of Mises rigid perfectly plastic material the upper bound theorem can be written in the form

$$\iint_{S_v} (F_i v_i) dS \leq \sigma_0 \iiint_V \zeta_{eq} dV - \iint_{S_f} (F_i u_i) dS + \frac{\sigma_0}{\sqrt{3}} \iint_{S_d} \llbracket u_\tau \rrbracket dS \quad (1)$$

Here and in what follows the summation convention, according to which a recurring letter suffix indicates that the sum must be formed of all terms obtainable by assigning to the suffix the values 1, 2, and 3, is adopted. Similarly, in a quantity containing two repeated suffixes, say i and j , the summation must be carried out for all values 1, 2, 3 of both i and j . In equation (1), v_i is the real velocity field, u_i is any kinematically admissible velocity field, S_d is the area of velocity discontinuity surfaces, $\llbracket u_\tau \rrbracket$ is the amount of jump of the tangential velocity across the velocity discontinuity surface found from the kinematically admissible velocity field. Note that the normal velocity must be continuous across any velocity discontinuity surface. The equivalent strain rate is defined by

$$\xi_{eq} = \sqrt{\frac{2}{3} \xi_{ij} \xi_{ij}} \quad (2)$$

where the components of the strain rate tensor, ξ_{ij} , are calculated by means of the real velocity field according to

$$\xi_{ij} = \frac{1}{2} (v_{i,j} + v_{j,i}) \quad (3)$$

In the case of kinematically admissible velocity fields equations (2) and (3) transform to

$$\zeta_{eq} = \sqrt{\frac{2}{3} \zeta_{ij} \zeta_{ij}}, \quad \zeta_{ij} = \frac{1}{2} (u_{i,j} + u_{j,i}) \quad (4)$$

The kinematically admissible velocity field is defined as any velocity field that satisfies the incompressibility equation and the velocity boundary conditions. The incompressibility equation in the case of kinematically admissible velocity fields can be written in the form

$$\zeta_{ii} = 0 \quad (5)$$

Having any kinematically admissible velocity field the right hand side of equation (1) can be calculated since F_i is prescribed over S_f . In general, the velocity field v_i is unknown. However, the value of its components involved in the integrand on the left hand side of (1) is known from the boundary conditions. Therefore, a combination of unknown components of F_i involved in the integrand on the left hand side of (1) can be evaluated with the use of any kinematically admissible velocity field. For finding analytical or semi-analytical solutions the kinematically admissible velocity field is usually chosen in the form of a function that contains one or several undetermined parameters. Substituting this function into (1) transforms the functional on its right hand side into a function. Then, this function should be minimized with respect of the undetermined parameters to find the best upper bound based on the kinematically admissible velocity field chosen.

In all boundary value problems considered in this chapter S_f is traction free. Therefore,

$$\iint_{S_f} (F_i u_i) dS = 0 \quad (6)$$

Moreover, external load is represented by a combination of concentrated forces and couples applied to rigid blocks. Let ω be the angular velocity of a generic rigid block to which a couple G is applied and U be the velocity of a point of this block at which a force F is applied (Figure 1). It is also assumed the vectors $\mathbf{G}$ and $\boldsymbol{\omega}$ as well as $\mathbf{F}$ and $\mathbf{U}$ are collinear and have the same direction. Then,

$$\iint_{S_v} (F_i v_i) dS = FU + G\omega \quad (7)$$

Substituting (6) and (7) into (1) leads to

$$\sum_{j=1}^n (F^j U_j + G^j \omega_j) \leq \sigma_0 \iiint_V \zeta_{eq} dV + \frac{\sigma_0}{\sqrt{3}} \iint_{S_d} [u_\tau] dS \quad (8)$$

where n is the number of rigid blocks to which the force, F^j , or couple, G^j , or force and couple is applied. It is convenient to rewrite equation (8) in the form

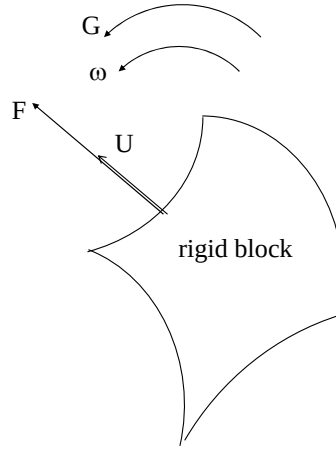


Figure 1. Force and couple applied to a rigid block

$$\sum_{j=1}^n (F_u^j U_j + G_u^j \omega_j) = \sigma_0 \iiint_V \zeta_{eq} dV + \frac{\sigma_0}{\sqrt{3}} \iint_{S_d} [u_\tau] dS \quad (9)$$

where F_u^j and G_u^j are upper bounds on F^j and G^j , respectively. It is assumed here that the right hand side of (9) should be minimized with respect to free parameters involved in the kinematically admissible velocity field.

The assumed function (or functions) involved in the kinematically admissible velocity field can have a large effect on the accuracy of the result. It is especially important to take into account the behavior of the real functions that must exist near singular surfaces. This is true even when finite element methods, such as UBET (Bramley, 2001), are used. It has been shown in Alexandrov and Richmond (2001) that the equivalent strain rate follows an inverse square root rule in the vicinity of surfaces on which the shear stress is equal to the shear yield stress (there is an exception to this rule and it is discussed in Alexandrov and Richmond, 2001). In particular, the shear stress is equal to the shear yield stress on velocity discontinuity surfaces. Therefore, it is reasonable to choose kinematically admissible velocity fields such that

$$\zeta_{eq} = O\left(\frac{1}{\sqrt{s}}\right), \quad s \rightarrow 0 \quad (10)$$

where s is the normal distance from the velocity discontinuity surface. Substituting (10) into (1) leads to the improper volume integral. Even though it is easy to show convergence, one needs to take this into account in numerical calculation. Note that if a kinematically admissible velocity field is chosen such that equation (10) is satisfied, the stress boundary condition over the velocity discontinuity surface is automatically satisfied, though it is not a requirement of the upper bound theorem.

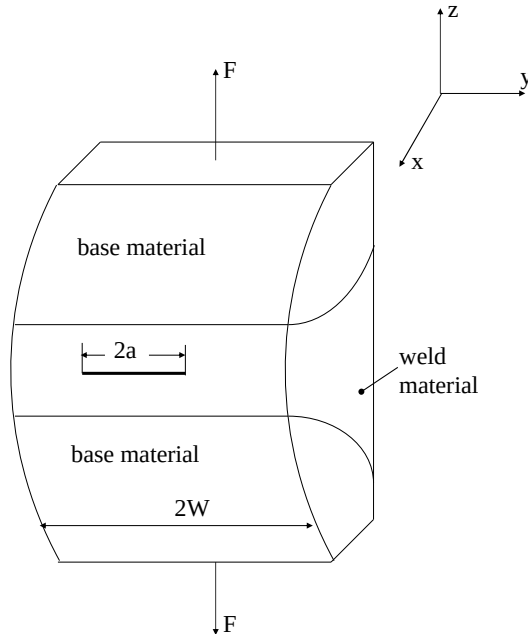


Figure 2. Geometry of structure under consideration – notation

Assume that the structure has a plane of symmetry, $z = 0$. It is advantageous to choose kinematically admissible velocity fields such that the shear strain rate vanishes at $z = 0$. For, as follows from the associated flow rule, the shear stress resulting from such velocity fields vanishes at $z = 0$ as well and this is the stress boundary condition at the plane of symmetry.

In many cases it is important to find the limit load for structures with a crack. A difficulty here is that there are a great number of geometric parameters of interest. Therefore, any method that allows one to reduce the number of parameters is very useful. For a class of structures such a method has been developed in Alexandrov and Kocak (2008). A structure with a through crack of length $2a$ and the orientation of the axes of a Cartesian coordinate system xyz are shown in Figure 2. This model is selected to consider the complexity of the weld thickness and the shape of the joined region. A particular case of this structure is the structure with no crack, $a = 0$. This structure is of special importance for the approach proposed. The class of structures under consideration is restricted by the assumptions that there is a plane of symmetry $z = \text{constant}$ (obviously, the crack must lie within this plane) and that all cross-sections $y = \text{constant}$ of the structure with no crack are identical. The latter, in particular, means that the structure with no crack has a plane of symmetry, $y = \text{constant}$, and that two boundaries of the structure are determined by the equations $y = \text{constant}$. It is possible to choose the origin of the Cartesian coordinate system such that the planes of symmetry are given by the equations $y = 0$ and $z = 0$. In this coordinate system, the aforementioned two boundaries are determined by the equations $y = \pm W$, where $2W$ is the width of the specimen. In the general case of the structure with a crack, the crack is located in the plane $z = 0$ and its tips in this plane are determined by the equations $y = \pm a$.

Thus, $y = 0$ and $z = 0$ are also planes of symmetry for the structure with the crack. Plastic properties of the material may vary continuously or piece-wise continuously throughout the volume of the material but their distribution should be symmetric relative to the plane $z = 0$ and should be identical in all cross-sections $y = \text{constant}$. A typical example of such structures is shown in Figure 2. It is a weld specimen whose plastic properties are defined by the tensile yield stress of the base material and the tensile yield stress of the weld material. To complete the description of the problem under consideration, it is necessary to specify that a tensile load, F , is applied in a direction parallel to the z -axis (Figure 2). Introduce a reference length L . Then, it is always possible to write the upper bound limit load for the structure with no crack, $F_u^{(0)}$, as

$$\frac{F_u^{(0)}}{4\sigma_0 BW} = \Omega(w) \quad (11)$$

where $2B$ is the thickness of the specimen at $z = 0$, $w = W/L$ and $\Omega(w)$ is the function of w that has been calculated for the structure with no crack with the use of the upper bound theorem. The notation for $\Omega(w)$ emphasizes that $\Omega(w)$ depends on w , although it may also depend on other parameters. In contrast, the structure may contain no parameter with units of length other than W . In such cases, $\Omega(w)$ is a constant. It has been shown in Alexandrov and Kocak (2008) that the upper bound limit load for the structure with a crack is

$$\frac{F_u}{4\sigma_0 BW} = \left(1 - \frac{a}{W}\right) \Omega(w_1) \quad (12)$$

where $w_1 = (W - a)/L$. Thus, once the function $\Omega(w)$ for the structure with no crack involved in (11) has been determined, the upper bound limit load for the structure with a crack is given by the simple formula (12). Note that there is no restriction on the method used to find $\Omega(w)$. In particular, a finite element method can be used to determine $\Omega(w)$ with a high accuracy. Then, equation (12) gives the limit load for the structure with a crack with the same accuracy.

3. PLANE STRAIN SOLUTIONS

In the case of plane strain solutions it is always possible to choose an orthogonal coordinate system $\alpha\beta z$ whose z -axis is orthogonal to the plane of flow. In such a coordinate system $\zeta_{\alpha z} = \zeta_{\beta z} = \zeta_{zz} = 0$. Therefore, the equation of incompressibility (5) and equation (4) transform to

$$\zeta_{\alpha\alpha} + \zeta_{\beta\beta} = 0, \quad \zeta_{eq} = \frac{2}{\sqrt{3}} \sqrt{\zeta_{\alpha\alpha}^2 + \zeta_{\alpha\beta}^2} = \frac{2}{\sqrt{3}} \sqrt{\zeta_{\beta\beta}^2 + \zeta_{\alpha\beta}^2} \quad (13)$$

The thickness of specimens in plane strain solutions is denoted $2B$ and the value of B has no effect on the solutions, though it is of course involved in dimensionless representations of the final result.

3.1. Middle Crack Tension Plates

Geometry of the specimen with a through-thickness crack, the system of loading, the direction of velocity of rigid blocks of base material U and the orientation of the axes of the Cartesian coordinate system xy are shown in Figure 3 where $2H$ is the thickness of the weld and $2W$ is the width of the specimen. A slip-line solution for such a specimen has been given in Hao *et.al.* (1997) and a finite element solution in Kim and Schwalbe (2001^a). It is obvious that the configuration shown in Figure 3 is a particular case of that in Figure 2. Therefore, equation (12) can be adopted. In particular, it is possible to assume that $L \equiv H$ in the definitions for w and w_1 . If the ratio H/W is small enough, the velocity discontinuity line occurs at the bi-material interface. On the basis of the approach developed in Alexandrov and Kocak (2008), an upper bound solution can be immediately derived from the solution of the very well-known Prandtl's problem for compression of a layer between two rough, parallel plates where the friction stress is assumed to be equal to the shear yield stress. The latter condition is of importance because the same magnitude of the shear stress occurs at the velocity discontinuity line. Therefore, the mathematical formulations of the problems for compression and tension of a layer with no crack are the same (the difference in sign is not essential). The solution of the Prandtl's problem is given, for example, in Hill (1950). In our nomenclature, the solution is represented as

$$\Omega(w) \equiv \frac{F_u^{(Pr)}}{4\sigma_0 BW} = \frac{(3+w)}{2\sqrt{3}} \quad (14)$$

for the range

$$w \geq 1 \quad (15)$$

Combining equations (12) and (14), the upper bound on the limit load for the specimen under consideration can be written as

$$f_u = \frac{F_u}{4\sigma_0 BW} = \frac{1}{2\sqrt{3}} \left(1 - \frac{a}{W} \right) \left(3 + \frac{W-a}{H} \right) \quad (16)$$

where f_u is the dimensionless upper bound limit load. The inequality (15) transforms to

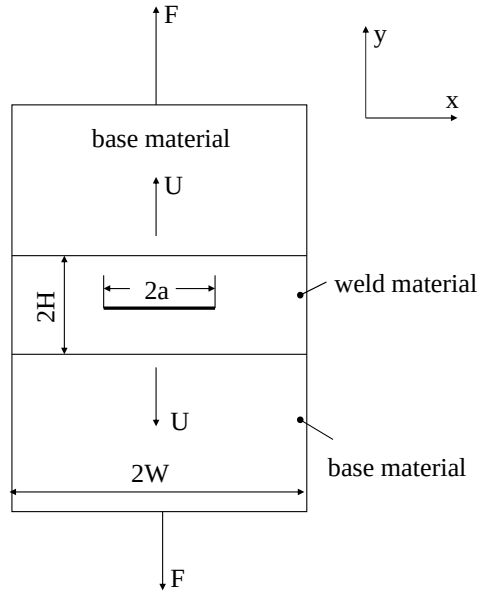


Figure 3. Geometry of structure under consideration – notation

$$\frac{W - a}{H} \geq 1 \quad (17)$$

3.2. Tensile Plates with a Crack Located at Some Distance from the Mid-Plane of the Weld

Geometry of the specimen with a through-thickness crack, the system of loading, the direction of velocity of rigid blocks of base material U and the orientation of the axes of the Cartesian coordinate system xy are shown in Figure 4 where ε is the distance to the crack from the mid-plane of the weld, $2a$ is the length of the crack, $2W$ is the width of the specimen and $2H$ is the thickness of the weld. It is assumed that $0 \leq \varepsilon \leq H$. The coordinate axes coincide with the intersection of the axes of symmetry of the specimen with no crack. The coordinates of the crack tips are $x = x_d$ and $y = y_d = \varepsilon$ for tip d , and $x = x_c$ and $y = y_c = \varepsilon$ for tip c . It is obvious that $x_d - x_c = 2a$. By assumption, $x_d \geq 0$ and $x_c \leq 0$. The previous configuration is obtained if $\varepsilon = 0$ and $x_d = |x_c|$. Numerical solutions for the special case of interface cracks (in this case $\varepsilon = H$) symmetric relative to the y -axis have been proposed in Kim and Schwalbe (2001^{b,c}) in the form of interpolating functions. A possible effect of the location of the crack is briefly discussed in Kim and Schwalbe (2001^b). A trivial modification of previously published solutions based on 4 isolated velocity discontinuity lines has been given in Kotousov and Jaffar (2006). In this chapter, a new analytic solution is obtained with the use of the solution (14).

The general structure of the chosen kinematically admissible velocity field in the weld is illustrated in Figure 5. It consists of two plastic zones and two rigid zones. The rigid zone 1

whose boundary is $mecdgk$ moves along with the base material located above the weld along the positive direction of the y -axis with velocity U . The rigid zone 2 whose boundary is m_1ecdgk_1 moves along with the base material located under the weld along the negative direction of the y -axis with the same velocity U . The plastic zones are separated from the rigid zones by the velocity discontinuity lines me , m_1e , kg , and k_1g . Also, there are four velocity discontinuity lines between the plastic zones and the base material. Those are qm , q_1m_1 , kp , and k_1p_1 . Moreover, there are two velocity discontinuity lines separating the rigid zones. Those are ec and dg . It follows from the virtual work rate principle of a continuum that

$$2FU = \dot{E}_1 + \dot{E}_2 + \dot{E}_d \quad (18)$$

where $\dot{E}_1$ is the energy dissipation rate in plastic zone 1 including the energy dissipation rate at the velocity discontinuity lines kp , k_1p_1 , kg , and k_1g , $\dot{E}_2$ is the energy dissipation rate in plastic zone 2 including the energy dissipation rate at the velocity discontinuity lines qm , q_1m_1 , me , and m_1e , and $\dot{E}_d$ is the energy dissipation rate at the velocity discontinuity lines dg and ce (Figure 5). The amount of velocity jump across each of these lines, dg and ce , is $2U$. The length of each line is ε . Therefore,

$$\dot{E}_d = \frac{8U\sigma_0 B\varepsilon}{\sqrt{3}} \quad (19)$$

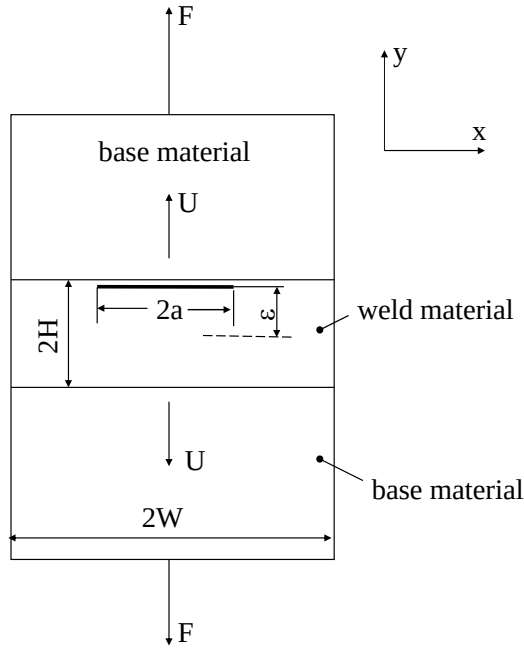
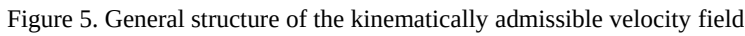


Figure 4. Geometry of structure under consideration – notation


$$F^{(P)} = \frac{2\sigma_0 BT}{\sqrt{3}} \left(3 + \frac{T}{H} \right) \quad (20)$$
$$T/H \geq 1 \quad (21)$$
$$\dot{E}_p = \frac{2U\sigma_0 BT}{\sqrt{3}} \left(3 + \frac{T}{H} \right) \quad (22)$$

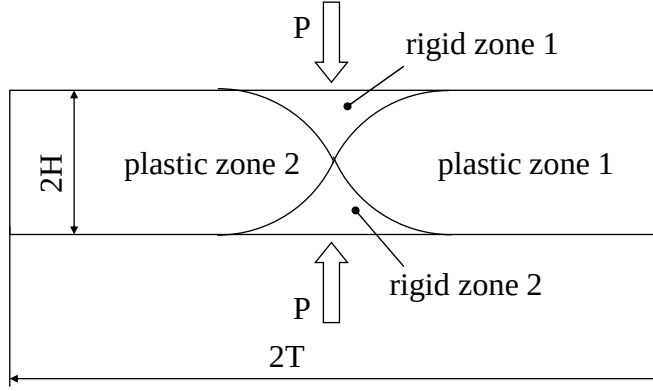


Figure 6. Plastic and rigid zones in compression of a plastic layer

The velocity field that appears in plastic zone 1 (Figure 6) can be used as the kinematically admissible velocity field in plastic zone 1 (Figure 5). Note that the energy dissipation rate at velocity discontinuity lines is the same as at friction surfaces at sliding where the friction stress is equal to the shear yield stress. Therefore, replacing the velocity discontinuity lines kp and k_1p_1 (Figure 5) with the friction surfaces is not essential. Thus, replacing T with $W_d = W - x_d$ and $\dot{E}_p$ with $\dot{E}_1$ in (22) leads to

$$\dot{E}_1 = \frac{2U\sigma_0 B(W - x_d)}{\sqrt{3}} \left[3 + \frac{(W - x_d)}{H} \right] \quad (23)$$

Analogously, comparing plastic zones 2 in Figures 5 and 6 and taking into account that $W_c = W + x_c$ results in

$$\dot{E}_2 = \frac{2U\sigma_0 B(W + x_c)}{\sqrt{3}} \left[3 + \frac{(W + x_c)}{H} \right] \quad (24)$$

Substituting (19), (23) and (24) into (18) results in

$$\frac{F_u}{4\sigma_0 BW} = f_u = \frac{1}{4\sqrt{3}} \left(1 - \frac{x_d}{W} \right) \left[3 + \frac{(W - x_d)}{H} \right] + \frac{1}{4\sqrt{3}} \left(1 + \frac{x_c}{W} \right) \left[3 + \frac{(W + x_c)}{H} \right] + \frac{\varepsilon}{\sqrt{3}W} \quad (25)$$

As follows from (21), the range of validity of this solution is

$$\frac{W - x_d}{H} \geq 1, \quad \frac{W + x_c}{H} \geq 1 \quad (26)$$

If the crack is symmetric relative to the y -axis then $x_d = -x_c = a$ and equation (25) simplifies to

$$\frac{F_u}{4\sigma_0 BW} = f_u = \frac{1}{2\sqrt{3}} \left(1 - \frac{a}{W} \right) \left[3 + \frac{(W-a)}{H} \right] + \frac{\varepsilon}{\sqrt{3}W} \quad (27)$$

It follows from (26) that this solution is valid for

$$\frac{W-a}{H} \geq 1 \quad (28)$$

It has been assumed that $x_d \geq 0$ and $x_c \leq 0$. Nevertheless, the solution (25) is formally valid even if $x_d < 0$ or $x_c > 0$. However, the larger $|x_d|$ (or x_c) in the case of $x_d < 0$ (or $x_c > 0$), the less accurate the solution is. Because the present analysis does not allow one to evaluate the loss of accuracy when $x_d < 0$ (or $x_c > 0$), it is recommended to use the solution (25) for specimens with $x_d \geq 0$ and $x_c \leq 0$.

3.3. Scarf-Joint Specimens with No Crack

Geometry of the specimen, the system of loading and the Cartesian coordinate system xy are shown in Figure 7 where $2H$ is the thickness of the weld, $2W$ is the width of the specimen and $\pi/2 - \beta$ is the orientation of the weld relative to the line of action of force F . It is supposed that the base material moves with velocity U along the line of action of force F , though it is not dictated by symmetry in the case under consideration. The general structure of the chosen kinematically admissible velocity field in the weld is shown in Figure 8 where $U_n = U \cos \beta$, and $U_\tau = U \sin \beta$. It consists of two plastic zones and two rigid zones. Because of symmetry, it is sufficient to get the solution in the domain $x \geq 0$. Note that U_n and U_τ are the velocity components of the rigid zones (base material). The normal velocity, U_n , must be continuous at the rigid/plastic interfaces whereas the tangential component, U_τ , may be discontinuous. In order to propose the kinematically admissible velocity field in the plastic zone, it is reasonable to modify the velocity field from the Prandtl-Nadai solution for compression of a plastic layer between two rough, parallel plates in which $U_\tau = 0$ (Figure 6). The modified velocity field has been proposed in Aleksandrov and Konchakova (2007) and has the following form in plastic zone 1 (Figure 8)

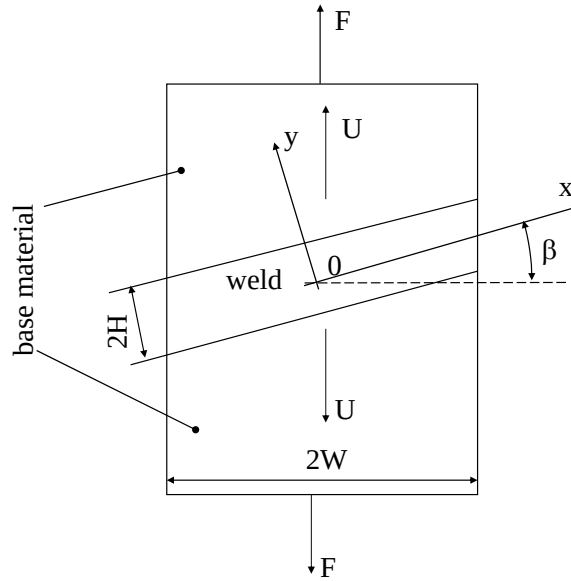


Figure 7. Geometry of structure under consideration – notation

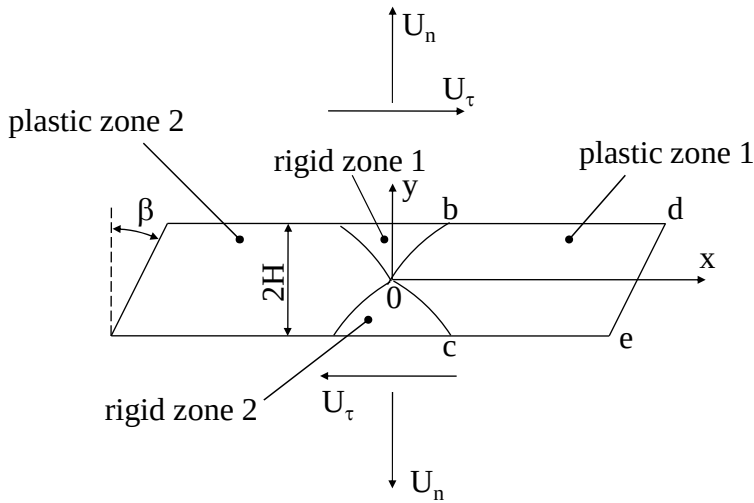


Figure 8. General structure of the kinematically admissible velocity field

$$\frac{u_x}{U_n} = C - \frac{x}{H} - 2\sqrt{1 - \left(\frac{y}{H}\right)^2} + 2C_1 \frac{y}{H}, \quad \frac{u_y}{U_n} = \frac{y}{H} \quad (29)$$

where u_x and u_y are the velocity components with respect to the xy coordinate system, and C and C_1 are undetermined constants. In the case of $C_1 = 0$ the classical velocity field from the Prandtl-Nadai solution is obtained (Hill, 1950). It is convenient to introduce the following new dimensionless variables

$$\mu = \frac{x}{H} \quad \text{and} \quad \sin \phi = \frac{y}{H} \quad (30)$$

Then,

$$\frac{d\phi}{dy} = \frac{1}{H \cos \phi} \quad (31)$$

and the velocity field (29) transforms to

$$\frac{u_x}{U_n} = C - \mu - 2 \cos \phi + 2C_1 \sin \phi, \quad \frac{u_y}{U_n} = \sin \phi \quad (32)$$

The non-zero strain rate components in the Cartesian coordinate system xy are

$$\zeta_{xx} = \frac{\partial u_x}{\partial x}, \quad \zeta_{yy} = \frac{\partial u_y}{\partial y}, \quad \zeta_{xy} = \frac{1}{2} \left(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x} \right) \quad (33)$$

Substituting (32) into (33) and using (31) result in

$$\zeta_{xx} = -\frac{U_n}{H}, \quad \zeta_{yy} = \frac{U_n}{H}, \quad \zeta_{xy} = \frac{U_n}{H} (\tan \phi + C_1) \quad (34)$$

Substituting (34) into (13) shows that the incompressibility equation is satisfied. The equivalent strain rate is determined from (13) and (34) as

$$\zeta_{eq} = \frac{2}{\sqrt{3}} \frac{U_n}{H} \sqrt{1 + (\tan \phi + C_1)^2} \quad (35)$$

Consider the velocity discontinuity line $0b$ (Figure 8). Let γ be the orientation of the tangent to this line relative to the x -axis. Then, the unit normal vector to line $0b$ is determined as (Figure 9)

$$\mathbf{n} = -\sin \gamma \mathbf{i} + \cos \gamma \mathbf{j} \quad (36)$$

where $\mathbf{i}$ and $\mathbf{j}$ are the base vectors of the Cartesian coordinate system. By definition, $\tan \gamma = dy/dx$. Then, it follows from (30) and (31) that

$$\tan \gamma = \frac{\cos \phi d\phi}{d\mu} \quad (37)$$

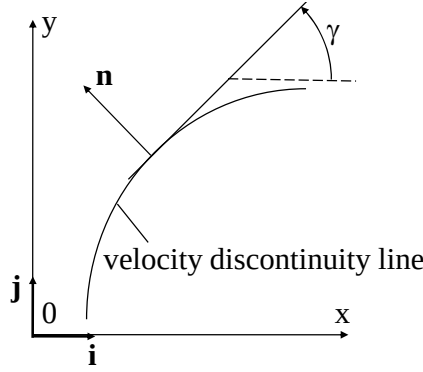


Figure 9. Geometry of generic velocity discontinuity line.

The normal velocity must be continuous across the velocity discontinuity line. This condition can be written in the form

$$\mathbf{u}_R \cdot \mathbf{n} = \mathbf{u}_P \cdot \mathbf{n} \quad (38)$$

where $\mathbf{u}_R$ is the velocity vector in the rigid zone 1 and $\mathbf{u}_P$ is the velocity vector in the plastic zone 1 (Figure 8). These vectors can be expressed in terms of $\mathbf{i}$ and $\mathbf{j}$ as

$$\mathbf{u}_R = U_r \mathbf{i} + U_n \mathbf{j}, \quad \mathbf{u}_P = u_x \mathbf{i} + u_y \mathbf{j} \quad (39)$$

As follows from (36), $\mathbf{n} \cdot \mathbf{i} = -\sin \gamma$ and $\mathbf{n} \cdot \mathbf{j} = \cos \gamma$. Therefore, substituting (39) into (38) gives

$$-U_r \sin \gamma + U_n \cos \gamma = -u_x \sin \gamma + u_y \cos \gamma \quad (40)$$

Using (32) and (37) and taking into account that $U_n = U \cos \beta$ and $U_r = U \sin \beta$ equation (40) can be transformed to

$$\frac{d\mu}{d\phi}(1 - \sin \phi) = \cos \phi (\mu + 2 \cos \phi - 2C_1 \sin \phi - C + \tan \beta) \quad (41)$$

This is a linear ordinary differential equation. Therefore, its general solution can be found with no difficulty. In order to formulate the boundary condition to equation (41), it is necessary to mention that the velocity field (32) is kinematically admissible if and only if the area of contact of the rigid zones (Figure 8) reduces to a point. Therefore, the velocity discontinuity line must pass through the origin of the coordinate system and the boundary condition to equation (41) is, as follows from (30), $\mu = 0$ at $\phi = 0$. The solution of equation (41) satisfying this condition is

$$\mu_{0b}(\phi) = \frac{\phi + \sin \phi \cos \phi + (\tan \beta - C) \sin \phi - C_1 \sin^2 \phi}{1 - \sin \phi} \quad (42)$$

The notation for $\mu_{0b}(\phi)$ emphasizes that equation (42) gives the dependence of μ on ϕ along the line $0b$. It follows from (30) that $\phi = \pi/2$ at $y = H$. At this value of ϕ the denominator of the right hand side of (42) vanishes. Therefore, the velocity discontinuity line $0b$ can have a common point with the line $y = H$ if and only if the numerator of the right hand side of (42) vanishes at $\phi = \pi/2$. This requires

$$C + C_1 = \tan \beta + \frac{\pi}{2} \quad (43)$$

The equation for the velocity discontinuity line $0c$ (Figure 8) can be obtained in a similar manner and it is

$$\mu_{0c}(\phi) = \frac{-\phi - \sin \phi \cos \phi + (\tan \beta + C) \sin \phi + C_1 \sin^2 \phi}{1 + \sin \phi} \quad (44)$$

The condition analogous to (43) is

$$C_1 - C = \tan \beta - \frac{\pi}{2} \quad (45)$$

Combining (43) and (45) leads to

$$C = \frac{\pi}{2}, \quad C_1 = \tan \beta \quad (46)$$

Substituting (46) into (42) and (44) gives

$$\begin{aligned} \mu_{0b}(\phi) &= \frac{\phi + (\cos \phi + \tan \beta - \pi/2) \sin \phi - \tan \beta \sin^2 \phi}{1 - \sin \phi}, \\ \mu_{0c}(\phi) &= \frac{-\phi + (\tan \beta + \pi/2 - \cos \phi) \sin \phi + \tan \beta \sin^2 \phi}{1 + \sin \phi} \end{aligned} \quad (47)$$

The right hand side of the first and the second of these equations reduces to the expression $0/0$ at $\phi = \pi/2$ and $\phi = -\pi/2$, respectively. Applying l'Hospital's rule to these equations results in

$$\mu_b = \frac{\pi}{2} + \tan \beta, \quad \mu_c = \frac{\pi}{2} - \tan \beta \quad (48)$$

where μ_b and μ_c are the value of μ at points b and c , respectively (Figure 8).

The energy dissipation rate at the velocity discontinuity line $0b$ is

$$\dot{E}_{0b} = \frac{2B\sigma_0}{\sqrt{3}} \int [u_\tau]_{0b} dl \quad (49)$$

where $[u_\tau]_{0b}$ is the amount of velocity jump across the velocity discontinuity line $0b$ and dl is the infinitesimal length element. By definition, $dl = \sqrt{(dx)^2 + (dy)^2}$. Therefore, it follows from (30), (41) and (46) that

$$dl = \frac{H \cos \phi}{(1 - \sin \phi)} \sqrt{\left[2 \tan \beta \sin \phi + \frac{\pi}{2} - \tan \beta - 2 \cos \phi - \mu_{0b}(\phi)\right]^2 + (1 - \sin \phi)^2} d\phi \quad (50)$$

The amount of velocity jump can be found from the following equation

$$[u_\tau]_{0b} = |\mathbf{u}_R - \mathbf{u}_P| \quad (51)$$

where the velocity vectors should be calculated at the velocity discontinuity line $0b$. Using (32), (40) and (46) and taking into account that $U_n = U \cos \beta$ and $U_\tau = U \sin \beta$ equation (51) can be transformed to

$$[u_\tau]_{0b} = U \cos \beta \sqrt{\left[2 \tan \beta \sin \phi + \frac{\pi}{2} - \tan \beta - 2 \cos \phi - \mu_{0b}(\phi)\right]^2 + (1 - \sin \phi)^2} \quad (52)$$

Substituting (50) and (52) into (49) gives

$$\begin{aligned} \dot{E}_{0b} = & \frac{2U\sigma_0 B H \cos \beta}{\sqrt{3}} \times \\ & \int_0^{\pi/2} \frac{\cos \phi}{(1 - \sin \phi)} \left\{ \left[2 \tan \beta \sin \phi + \frac{\pi}{2} - \tan \beta - 2 \cos \phi - \mu_{0b}(\phi) \right]^2 + (1 - \sin \phi)^2 \right\} d\phi \end{aligned} \quad (53)$$

The energy dissipation rate at the velocity discontinuity line $0c$ (Figure 8) can be found in a similar manner. As a result,

$$\dot{E}_{0c} = \frac{2U\sigma_0 BH \cos \beta}{\sqrt{3}} \times \int_{-\pi/2}^0 \frac{\cos \phi}{(1 + \sin \phi)} \left\{ \left[2 \tan \beta \sin \phi + \frac{\pi}{2} + \tan \beta - 2 \cos \phi - \mu_{0c}(\phi) \right]^2 + (1 + \sin \phi)^2 \right\} d\phi \quad (54)$$

Equations (47) should be used to exclude $\mu_{0b}(\phi)$ and $\mu_{0c}(\phi)$ in the integrands in (53) and (54). The integrals in (53) and (54) are improper. Even though it is easy to show convergence, one needs to take this into account in a numerical code.

There are two more velocity discontinuity lines, bd and ce (Figure 8). The values of μ at points d and e are determined from geometric relations and (30) as

$$\mu_d = \tan \beta + \frac{W}{H \cos \beta}, \quad \mu_e = -\tan \beta + \frac{W}{H \cos \beta} \quad (55)$$

The amount of velocity jump across the line bd is $[[u_\tau]]_{bd} = U_\tau - u_x = U \sin \beta - u_x$ where u_x should be calculated at $\phi = \pi/2$ by means of (32) and (46). Then, the energy dissipation rate at the velocity discontinuity line bd is

$$\dot{E}_{bd} = \frac{2\sigma_0 BH}{\sqrt{3}} \int_{\mu_b}^{\mu_d} [[u]]_{bd} d\mu = \frac{2U\sigma_0 BH \cos \beta}{\sqrt{3}} \int_{\mu_b}^{\mu_d} \left(\mu - \frac{\pi}{2} - \tan \beta \right) d\mu \quad (56)$$

Analogously, for the velocity discontinuity line ce (Figure 8)

$$\dot{E}_{ce} = \frac{2U\sigma_0 BH \cos \beta}{\sqrt{3}} \int_{\mu_c}^{\mu_e} \left(\mu - \frac{\pi}{2} + \tan \beta \right) d\mu \quad (57)$$

Integration in (56) and (57) can be carried out analytically to give, with the use of (48) and (55),

$$\dot{E}_{bd} = \dot{E}_{ce} = \frac{U\sigma_0 BH \cos \beta}{\sqrt{3}} \left(\frac{W}{H \cos \beta} - \frac{\pi}{2} \right)^2 \quad (58)$$

The energy dissipation rate in the plastic zone is

$$\dot{E}_{pl} = 2\sigma_0 B \iint \zeta_{eq} dx dy \quad (59)$$

where integration should be completed over the area of plastic zone 1 (Figure 8). Substituting (30), (31), (35), and (46) into (59) and taking into account that $U_n = U \cos \beta$ leads to

$$\dot{E}_{pl} = \frac{4U\sigma_0 BH \cos \beta}{\sqrt{3}} \iint \sqrt{1 + (\tan \phi + \tan \beta)^2} \cos \phi d\mu d\phi \quad (60)$$

Since the integrand is independent of μ , integration with respect to this argument can be carried out analytically to give

$$\dot{E}_{pl} = \frac{4U\sigma_0 BH \cos \beta}{\sqrt{3}} \left[\int_0^{\pi/2} \sqrt{1 + (\tan \phi + \tan \beta)^2} \cos \phi (\mu_{de}(\phi) - \mu_{ob}(\phi)) d\phi + \int_{-\pi/2}^0 \sqrt{1 + (\tan \phi + \tan \beta)^2} \cos \phi (\mu_{de}(\phi) - \mu_{oc}(\phi)) d\phi \right] \quad (61)$$

Here $\mu_{de}(\phi)$ is the dependence of μ on ϕ along the line de (Figure 8). The dependence of x on y along this line can be found from geometric consideration with no difficulty. Then, it follows from (30) that

$$\mu_{de}(\phi) = \sin \phi \tan \beta + \frac{W}{H \cos \beta} \quad (62)$$

In the case under consideration,

$$\sigma_0 \iiint_V \zeta_{eq} dV = \dot{E}_{pl}, \quad \frac{\sigma_0}{\sqrt{3}} \iint_{S_d} [u_r] dS = \dot{E}_{ob} + \dot{E}_{oc} + \dot{E}_{ce} + \dot{E}_{bd}$$

and, then, equation (9) becomes

$$\frac{F_u}{4\sigma_0 BW} = f_u = \frac{\dot{E}_{ob} + \dot{E}_{oc} + \dot{E}_{ce} + \dot{E}_{bd} + \dot{E}_{pl}}{4U\sigma_0 BW} \quad (63)$$

where F_u is the upper bound of the actual force F and f_u is its dimensionless representation. It is seen from Figure 8 and equation (30) that the kinematically admissible velocity field chosen is applicable if $\mu_d \geq \mu_b$ and $\mu_e \geq \mu_c$. Using (48) and (55) these inequalities can be transformed to

$$\frac{W}{H \cos \beta} - \frac{\pi}{2} \geq 0 \quad (64)$$

The right hand side of (63) can be calculated by means of (53), (54), (58), and (61) with the use of (47) and (62). The variation of f_u with H/W for several values of β is depicted in Figure 10 for the range of H/W satisfying (64). Note that a particular case of the

configuration shown in Figure 7 at $\beta = 0$ coincides with a particular case of the configuration shown in Figure 3 at $a = 0$. Since the solution (16) has been based on the numerical solution (14), the former at $a = 0$ can be used to verify the accuracy of the solution (63) at $\beta = 0$. The corresponding values of f_u are shown in Figure 11 where the dashed line corresponds to the solution (16) and the solid line to the solution (63). It is seen that the difference is very small.

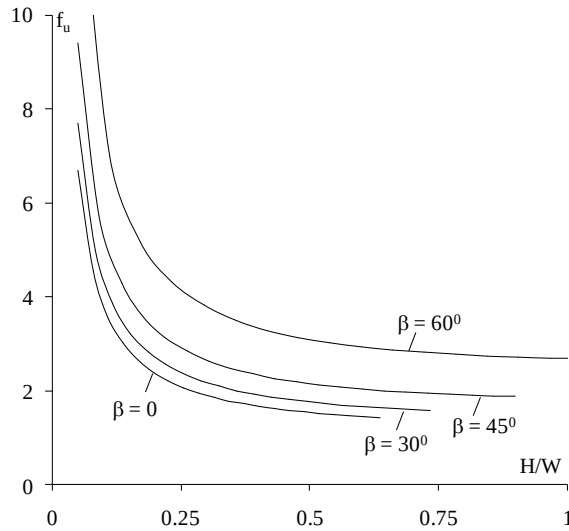


Figure 10. Variation of dimensionless limit load with H/W for several values of β

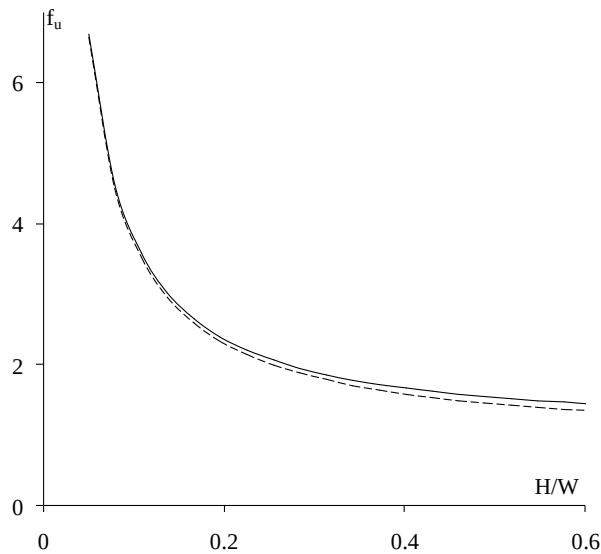


Figure 11. Comparison between solutions (16) and (63)

3.4. Scarf-Joint Specimens with a Crack

Consider the previous configuration (Figure 7) assuming that there is a crack within the weld parallel to the x -axis (Figure 12). The position and size of the crack are completely determined by the coordinates of its tips, namely x_s and $y_s = \varepsilon$ for tip s and x_t and $y_t = \varepsilon$ for tip t . By assumption, $x_s \geq 0$ and $x_t \leq 0$. The value of ε varies in the range $0 \leq \varepsilon \leq H$. The general structure of the chosen kinematically admissible velocity field within the weld is shown in Figure 13. It consists of two plastic zones and two rigid zones. The rigid zone 1 whose boundary is $mktsbe$ moves along with the base material located above the weld. The rigid zone 2 whose boundary is $pktsbc$ moves along with the base material located under the weld. The plastic zones are separated from the rigid zones by the velocity discontinuity lines eb , bc , mk , and kp . Also, there are 4 velocity discontinuity lines between the plastic zones and the base material. Those are ed , cf , mn , and qp . Moreover, there are 2 velocity discontinuity lines separating the rigid zones. Those are sb and kt . The amount of velocity jump across each of these velocity lines, sb and kt , is $2U$. Therefore, it follows from the virtual work rate principle of a continuum that

$$2FU = \dot{E}_1 + \dot{E}_2 + \frac{4UB\sigma_0}{\sqrt{3}}(L_{sb} + L_{kt}) \quad (65)$$

where $\dot{E}_1$ is the energy dissipation rate in plastic zone 1 including the energy dissipation rate at the velocity discontinuity lines be , ed , bc , and cf and $\dot{E}_2$ is the energy dissipation rate in plastic zone 2 including the energy dissipation rate at the velocity discontinuity lines km , mn , kp , and pq . Also, L_{sb} is the length of line sb and L_{kt} is the length of line kt (Figure 13). It follows from geometric consideration (Figure 14) that

$$L_{sb} = \frac{\varepsilon}{\cos \beta} \quad \text{and} \quad W_1 = W - \Delta W_1 \quad (66)$$

where

$$\Delta W_1 = r_s \cos(\gamma + \beta), \quad r_s = \sqrt{x_s^2 + \varepsilon^2}, \quad \tan \gamma = \frac{\varepsilon}{x_s} \quad (67)$$

Analogously,

$$L_{kt} = \frac{\varepsilon}{\cos \beta}, \quad W_2 = W - \Delta W_2, \quad (68)$$

$$\Delta W_2 = r_t \cos(\phi + \beta), \quad r_t = \sqrt{x_t^2 + \varepsilon^2}, \quad \tan \phi = \frac{\varepsilon}{x_t}$$

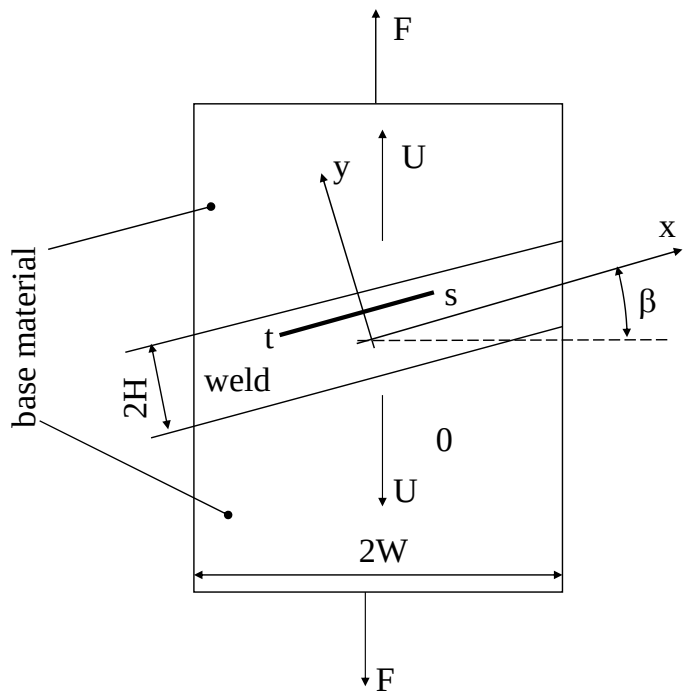


Figure 12. Geometry of structure under consideration – notation

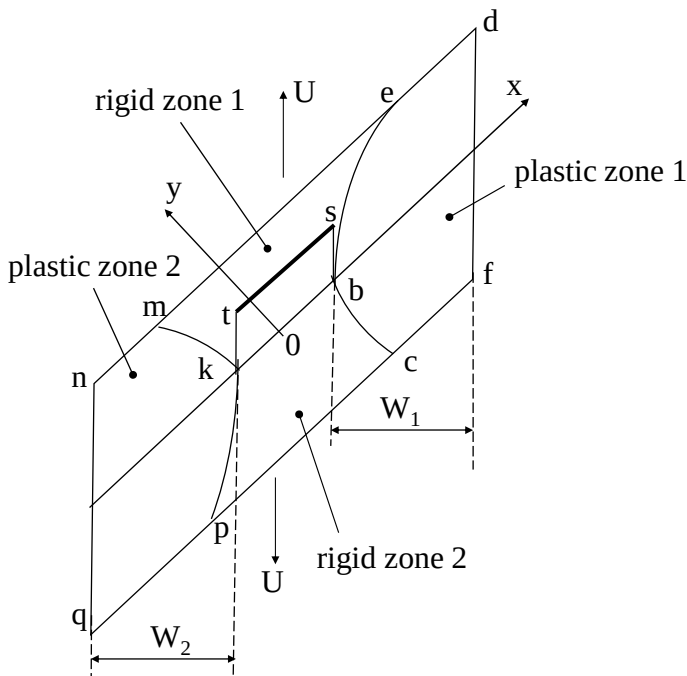


Figure 13. General structure of the kinematically admissible velocity field

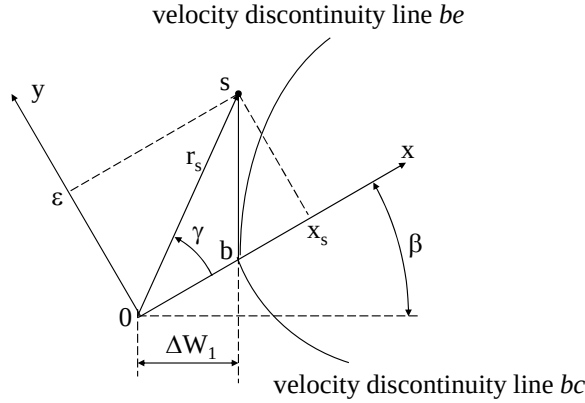


Figure 14. Illustration of geometric relations for determining W_1 and the length of velocity discontinuity line sb

In order to find the values of $\dot{E}_1$ and $\dot{E}_2$, it is possible to adopt the solution given in the previous section. For the specimen with no crack $W_1 = W_2 = W$ because of symmetry and the solution has the form of (63). Using the virtual work rate principle of a continuum and taking into account that the two plastic zones shown in Figure 8 are identical it is possible to find that the energy dissipation rate in each plastic zone, including the velocity discontinuity lines, is $\dot{E}_0 = F_u^{(0)}U$ where $F_u^{(0)}$ is equal to F_u from (63). Thus

$$\dot{E}_0 \left(\frac{H}{W}, \beta \right) = 4U\sigma_0 B W f_u^{(0)} \left(\frac{H}{W}, \beta \right) \quad (69)$$

where $f_u^{(0)}$ is equal to f_u from (63). It is emphasized in (69) that $\dot{E}_0$ and $f_u^{(0)}$ depend on H/W and β . The velocity field that appears in plastic zone 1 (Figure 8) can be used as the kinematically admissible velocity field in plastic zone 1 (Figure 13). Thus, replacing W with W_1 and $\dot{E}_0$ with $\dot{E}_1$ in (69) leads to

$$\dot{E}_1 = 4U\sigma_0 B W_1 f_u^{(0)} \left(\frac{H}{W_1}, \beta \right) \quad (70)$$

Analogously, comparing the velocity fields in plastic zones 2 in Figures 8 and 13 results in

$$\dot{E}_2 = 4U\sigma_0 B W_2 f_u^{(0)} \left(\frac{H}{W_2}, \beta \right) \quad (71)$$

The inequality (64) transforms to

$$\frac{W_1}{H \cos \beta} - \frac{\pi}{2} \geq 0, \quad \frac{W_2}{H \cos \beta} - \frac{\pi}{2} \geq 0 \quad (72)$$

Substituting (66), (68), (70), and (71) into (65) gives

$$\frac{F_u}{4\sigma_0 BW} = f_u = \frac{1}{2} \frac{W_1}{W} f_u^{(0)} \left(\frac{H}{W_1}, \beta \right) + \frac{1}{2} \frac{W_2}{W} f_u^{(0)} \left(\frac{H}{W_2}, \beta \right) + \frac{\varepsilon}{\sqrt{3}W \cos \beta} \quad (73)$$

Since the value of $f_u^{(0)}$ has been already found (see Figure 10), equation (73) immediately provides the solution for the configuration under consideration. The range of validity of the solution is given in (72). As in the case considered in Section 3.2, the restrictions $x_s \geq 0$ and $x_t \leq 0$ may or may not be important. It depends on specific applications.

3.5. Additional Comments on the Limit Load Solutions for Tensile Plates

Several upper bound limit load solutions for tensile plates with a crack have been proposed in Sections 3.1, 3.2 and 3.4. In the present section, the corresponding dimensionless limit loads will be denoted by $f_u^{(1)}$. The kinematically admissible velocity fields adopted to find $f_u^{(1)}$ contain no free parameters for minimization in (9). Another way to use the upper bound theorem is to adopt a qualitatively different kinematically admissible velocity field. Using such a field it is possible to find another value of the upper bound limit load, say $f_u^{(2)}$. Then, according to the upper bound theorem, the solution based on the two kinematically admissible velocity fields is

$$f_u = \min \{ f_u^{(1)}, f_u^{(2)} \} \quad (74)$$

When the crack is large enough, a better prediction, as compared to $f_u^{(1)}$, can be obtained with the use of kinematically admissible velocity fields consisting of isolated velocity discontinuity lines. Since the configurations shown in Figures 3 and 4 are particular cases of the configuration shown in Figure 12, the latter will be considered first. The general structure of the chosen kinematically admissible velocity field within the weld is shown in Figure 15. The velocity field consists of four rigid blocks separated by the velocity discontinuity lines sc , sb , td , and te . Rigid blocks 1 and 2 move along with the base material with velocity U in the opposite directions. Rigid blocks 3 and 4 move with velocities U_3 and U_4 , respectively. The magnitude and direction of these velocities are unknown. Represent these vectors in the form

$$\mathbf{U}_3 = U_{3x}\mathbf{i} + U_{3y}\mathbf{j}, \quad \mathbf{U}_4 = U_{4x}\mathbf{i} + U_{4y}\mathbf{j} \quad (75)$$

Let $\mathbf{n}$ be the unit normal to line sc . Then (Figure 15),

$$\mathbf{n} = -\sin \varphi_1 \mathbf{i} + \cos \varphi_1 \mathbf{j} \quad (76)$$

The velocity vector of rigid block 1 is represented as (Figure 15)

$$\mathbf{U} = U \sin \beta \mathbf{i} + U \cos \beta \mathbf{j} \quad (77)$$

Since the normal velocity must be continuous across the velocity discontinuity line, $\mathbf{U}_3 \cdot \mathbf{n} = \mathbf{U} \cdot \mathbf{n}$. Substituting (75), (76) and (77) into this equation gives

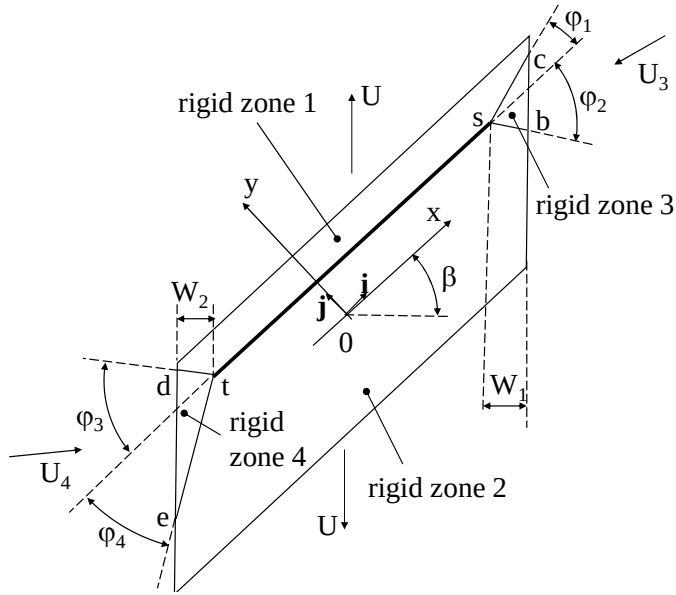


Figure 15. General structure of the kinematically admissible velocity field

$$-U_{3x} \sin \varphi_1 + U_{3y} \cos \varphi_1 = U (\cos \beta \cos \varphi_1 - \sin \beta \sin \varphi_1) \quad (78)$$

The velocity discontinuity lines sb , td and te can be treated in a similar manner to result in

$$\begin{aligned} -U_{3x} \sin \varphi_2 - U_{3y} \cos \varphi_2 &= U (\cos \beta \cos \varphi_2 + \sin \beta \sin \varphi_2), \\ U_{4x} \sin \varphi_3 + U_{4y} \cos \varphi_3 &= U (\cos \beta \cos \varphi_3 + \sin \beta \sin \varphi_3), \\ U_{4x} \sin \varphi_4 - U_{4y} \cos \varphi_4 &= U (\cos \beta \cos \varphi_4 - \sin \beta \sin \varphi_4) \end{aligned} \quad (79)$$

Solving equations (78) and (79) for U_{3x} , U_{3y} , U_{4x} , and U_{4y} gives

$$\begin{aligned} U_{3x} &= -U \frac{[2 \cos \beta \cos \varphi_1 \cos \varphi_2 - \sin \beta \sin(\varphi_1 - \varphi_2)]}{\sin(\varphi_1 + \varphi_2)}, \\ U_{3y} &= -U \frac{[2 \sin \beta \sin \varphi_1 \sin \varphi_2 + \cos \beta \sin(\varphi_1 - \varphi_2)]}{\sin(\varphi_1 + \varphi_2)}, \\ U_{4x} &= U \frac{[2 \cos \beta \cos \varphi_3 \cos \varphi_4 + \sin \beta \sin(\varphi_3 - \varphi_4)]}{\sin(\varphi_3 + \varphi_4)}, \\ U_{4y} &= U \frac{[2 \sin \beta \sin \varphi_3 \sin \varphi_4 - \cos \beta \sin(\varphi_3 - \varphi_4)]}{\sin(\varphi_3 + \varphi_4)} \end{aligned} \quad (80)$$

Let $\boldsymbol{\tau}$ be the unit vector parallel to line sc . Then (Figure 15),

$$\boldsymbol{\tau} = \cos \varphi_1 \mathbf{i} + \sin \varphi_1 \mathbf{j} \quad (81)$$

The amount of velocity jump across this velocity discontinuity line is determined by

$$[[u_\tau]]_{sc} = |(\mathbf{U} - \mathbf{U}_3) \cdot \boldsymbol{\tau}| \quad (82)$$

Substituting (75), (77) and (81) into (82) leads to

$$[[u_\tau]]_{sc} = |(U \sin \beta - U_{3x}) \cos \varphi_1 + (U \cos \beta - U_{3y}) \sin \varphi_1| \quad (83)$$

Excluding here U_{3x} and U_{3y} by means of (80) gives

$$\frac{[[u_\tau]]_{sc}}{U} = \frac{2 \cos(\beta - \varphi_2)}{\sin(\varphi_1 + \varphi_2)} \quad (84)$$

Analogously,

$$\frac{[[u_\tau]]_{sb}}{U} = \frac{2 \cos(\beta + \varphi_1)}{\sin(\varphi_1 + \varphi_2)}, \quad \frac{[[u_\tau]]_{td}}{U} = \frac{2 \cos(\beta + \varphi_4)}{\sin(\varphi_3 + \varphi_4)}, \quad \frac{[[u_\tau]]_{te}}{U} = \frac{2 \cos(\beta - \varphi_3)}{\sin(\varphi_3 + \varphi_4)} \quad (85)$$

for lines sb , td and te , respectively. It follows from geometric consideration (Figure 15) that

$$L_{sc} = \frac{W_1}{\cos(\varphi_1 + \beta)}, \quad L_{sb} = \frac{W_1}{\cos(\varphi_2 - \beta)}, \quad L_{td} = \frac{W_2}{\cos(\varphi_3 - \beta)}, \quad L_{te} = \frac{W_2}{\cos(\varphi_4 + \beta)} \quad (86)$$

where L_{sc} , L_{sb} , L_{td} , and L_{te} are the lengths of the velocity discontinuity lines sc , sb , td and te , respectively. W_1 and W_2 in (86) should be excluded by means of (66), (67) and (68). The energy dissipation rate at the velocity discontinuity line sc is determined by

$$\dot{E}_{sc} = \frac{2\sigma_0 B}{\sqrt{3}} L_{sc} |[u_\tau]_{sc}| \quad (87)$$

Substituting (83) and (86) into (87) gives

$$\dot{E}_{sc} = \frac{4U\sigma_0 BW_1}{\sqrt{3}} \frac{\cos(\beta - \varphi_2)}{\cos(\varphi_1 + \beta)\sin(\varphi_1 + \varphi_2)} \quad (88)$$

Analogously,

$$\begin{aligned} \dot{E}_{sb} &= \frac{4U\sigma_0 BW_1}{\sqrt{3}} \frac{\cos(\beta + \varphi_1)}{\cos(\varphi_2 - \beta)\sin(\varphi_1 + \varphi_2)}, \\ \dot{E}_{td} &= \frac{4U\sigma_0 BW_2}{\sqrt{3}} \frac{\cos(\beta + \varphi_4)}{\cos(\varphi_3 - \beta)\sin(\varphi_3 + \varphi_4)}, \\ \dot{E}_{te} &= \frac{4U\sigma_0 BW_2}{\sqrt{3}} \frac{\cos(\beta - \varphi_3)}{\cos(\varphi_4 + \beta)\sin(\varphi_3 + \varphi_4)} \end{aligned} \quad (89)$$

for lines sb , td and te , respectively. Since there is no plastic domain of a finite size,

$$\sigma_0 \iiint_V \zeta_{eq} dV = 0, \quad \frac{\sigma_0}{\sqrt{3}} \iint_{S_d} |[u_\tau]| dS = \dot{E}_{sc} + \dot{E}_{sb} + \dot{E}_{td} + \dot{E}_{te} \quad (90)$$

Therefore, (9) transforms to

$$2F_u U = \dot{E}_{sc} + \dot{E}_{sb} + \dot{E}_{td} + \dot{E}_{te} \quad (91)$$

Substituting (88) and (89) into (91) leads to

$$\frac{F_u}{4\sigma_0 BW} = f_u^{(2)} = \frac{W_1}{2\sqrt{3}W \sin(\varphi_1 + \varphi_2)} \left[\frac{\cos(\beta - \varphi_2)}{\cos(\varphi_1 + \beta)} + \frac{\cos(\beta + \varphi_1)}{\cos(\varphi_2 - \beta)} \right] + \frac{W_2}{2\sqrt{3}W \sin(\varphi_3 + \varphi_4)} \left[\frac{\cos(\beta + \varphi_4)}{\cos(\varphi_3 - \beta)} + \frac{\cos(\beta - \varphi_3)}{\cos(\varphi_4 + \beta)} \right] \quad (92)$$

The value of $f_u^{(2)}$ depends on four free parameters, namely φ_1 , φ_2 , φ_3 , and φ_4 . According to the upper bound theorem, the right hand side of (92) should be minimized with respect to these parameters. It is however necessary to take into account geometric restrictions imposed on these parameters. In particular, since it has been assumed that plastic deformation is wholly confined within the weld material, the maximum possible value of φ_1 is obtained when point c (Figure 15) coincides with the intersection of the boundary between the base and weld materials. Therefore, $\varphi_1^{\max} = \angle scg$ (Figure 16). Moreover, it follows from geometric relations that

$$\varphi_1^{\max} + \angle scg = \frac{\pi}{2} - \beta \quad (93)$$

The law of sines results in

$$\frac{\sin \varphi_1^{\max}}{H - \varepsilon} = \frac{\sin(\angle scg)}{W_1} \quad (94)$$

Excluding $\angle scg$ in (94) by means of (93) gives

$$\frac{\sin \varphi_1^{\max}}{H - \varepsilon} = \frac{\cos(\beta + \varphi_1^{\max})}{W_1} \quad (95)$$

or, with the use of trigonometric relations,

$$\tan \varphi_1^{\max} = \frac{(H - \varepsilon) \cos \beta}{W_1 + (H - \varepsilon) \sin \beta} \quad (96)$$

Analogously,

$$\tan \varphi_2^{\max} = \frac{(H + \varepsilon) \cos \beta}{W_1 - (H + \varepsilon) \sin \beta}, \quad \tan \varphi_3^{\max} = \frac{(H - \varepsilon) \cos \beta}{W_2 - (H - \varepsilon) \sin \beta}, \quad \tan \varphi_4^{\max} = \frac{(H + \varepsilon) \cos \beta}{W_2 + (H + \varepsilon) \sin \beta} \quad (97)$$

for φ_2 , φ_3 , and φ_4 , respectively. A necessary condition for a minimum of $f_u^{(2)}$ is

$$\frac{\partial f_u^{(2)}}{\partial \varphi_1} = 0, \quad \frac{\partial f_u^{(2)}}{\partial \varphi_2} = 0, \quad \frac{\partial f_u^{(2)}}{\partial \varphi_3} = 0, \quad \frac{\partial f_u^{(2)}}{\partial \varphi_4} = 0 \quad (98)$$

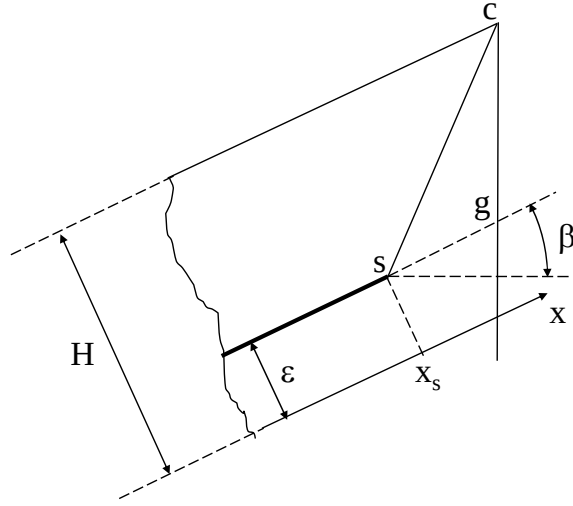


Figure 16. Illustration of geometric relations for determining the maximum possible value of φ_1

From here four equations for φ_1 , φ_2 , φ_3 , and φ_4 are obtained. The solution to this system of equations will be denoted by $\varphi_1^{(m)}$, $\varphi_2^{(m)}$, $\varphi_3^{(m)}$, and $\varphi_4^{(m)}$. It follows from the structure of equation (92) that the equations for φ_1 and φ_2 are independent of the equations for φ_3 and φ_4 . In particular, substituting (92) into (98) and taking into account that W_1 and W_2 are independent of the free parameters result in

$$2 \cos^2(\beta + \varphi_1^{(m)}) = \sin^2(\varphi_1^{(m)} + \varphi_2^{(m)}), \quad 2 \cos^2(\beta - \varphi_2^{(m)}) = \sin^2(\varphi_1^{(m)} + \varphi_2^{(m)}) \quad (99)$$

and

$$2 \cos^2(\beta - \varphi_3^{(m)}) = \sin^2(\varphi_3^{(m)} + \varphi_4^{(m)}), \quad 2 \cos^2(\beta + \varphi_4^{(m)}) = \sin^2(\varphi_3^{(m)} + \varphi_4^{(m)}) \quad (100)$$

A consequence of equations (99) is

$$\cos^2(\beta + \varphi_1^{(m)}) = \cos^2(\beta - \varphi_2^{(m)}) \quad (101)$$

The solution of this equation $\varphi_1^{(m)} = -\varphi_2^{(m)}$ should be excluded because both $\varphi_1 > 0$ and $\varphi_2 > 0$. Therefore, it follows from (101) that

$$\varphi_2^{(m)} = \varphi_1^{(m)} + 2\beta \quad (102)$$

Combining (102) and any of equations (99) gives

$$\varphi_1^{(m)} = \frac{\pi}{4} - \beta, \quad \varphi_2^{(m)} = \frac{\pi}{4} + \beta \quad (103)$$

Analogously, from equations (100)

$$\varphi_3^{(m)} = \frac{\pi}{4} + \beta, \quad \varphi_4^{(m)} = \frac{\pi}{4} - \beta \quad (104)$$

In order to find $f_u^{(2)}$, it is first necessary to determine $\varphi_1^{(2)}$, $\varphi_2^{(2)}$, $\varphi_3^{(2)}$, and $\varphi_4^{(2)}$ with the use of equations (96), (97), (103), and (104) according to $\varphi_1^{(2)} = \min\{\varphi_1^{(m)}, \varphi_1^{\max}\}$, $\varphi_2^{(2)} = \min\{\varphi_2^{(m)}, \varphi_2^{\max}\}$, $\varphi_3^{(2)} = \min\{\varphi_3^{(m)}, \varphi_3^{\max}\}$, $\varphi_4^{(2)} = \min\{\varphi_4^{(m)}, \varphi_4^{\max}\}$. Having these values of $\varphi_1^{(2)}$, $\varphi_2^{(2)}$, $\varphi_3^{(2)}$, and $\varphi_4^{(2)}$ the magnitude of $f_u^{(2)}$ can be immediately found from (92) replacing φ_1 , φ_2 , φ_3 , and φ_4 with $\varphi_1^{(2)}$, $\varphi_2^{(2)}$, $\varphi_3^{(2)}$, and $\varphi_4^{(2)}$, respectively. In particular, if $\varphi_1^{(2)} = \varphi_1^{(m)}$, $\varphi_2^{(2)} = \varphi_2^{(m)}$, $\varphi_3^{(2)} = \varphi_3^{(m)}$, and $\varphi_4^{(2)} = \varphi_4^{(m)}$, it follows from (92) that

$$f_u^{(2)} = \frac{W_1 + W_2}{\sqrt{3}W} \quad (105)$$

In the case of the configuration considered in Section 3.1, it follows from (66), (67) and (68) that $W_1 = W_2 = W - a$. Therefore, equation (105) simplifies to

$$f_u^{(2)} = \frac{2(W - a)}{\sqrt{3}W} \quad (106)$$

Substituting (16) and (106) into (74) gives

$$f_u = \min\left\{\frac{1}{2\sqrt{3}}\left(1 - \frac{a}{W}\right)\left(3 + \frac{W - a}{H}\right), \frac{2}{\sqrt{3}}\left(1 - \frac{a}{W}\right)\right\} \quad (107)$$

Solving the equation

$$\frac{1}{2\sqrt{3}}\left(1-\frac{a}{W}\right)\left(3+\frac{W-a}{H}\right)=\frac{2}{\sqrt{3}}\left(1-\frac{a}{W}\right) \quad (108)$$

it is possible to rewrite the solution (107) in the form

$$f_u = \begin{cases} \frac{1}{2\sqrt{3}}\left(1-\frac{a}{W}\right)\left(3+\frac{W-a}{H}\right), & \text{for } \frac{W-a}{H} \geq 1 \\ \frac{2}{\sqrt{3}}\left(1-\frac{a}{W}\right), & \text{for } \frac{W-a}{H} \leq 1 \end{cases} \quad (109)$$

It is seen from (109) that the condition (17) is satisfied. Also, the assumption of $\varphi_1^{(2)} = \varphi_1^{(m)}$, $\varphi_2^{(2)} = \varphi_2^{(m)}$, $\varphi_3^{(2)} = \varphi_3^{(m)}$, and $\varphi_4^{(2)} = \varphi_4^{(m)}$ is confirmed by the condition $(W-a)/H \leq 1$ when $f_u = f_u^{(2)}$ in (109). Note that $\beta = 0$ for the configuration under consideration.

In the case of the configurations considered in Sections 3.2 and 3.4 the energy dissipation rate at the velocity discontinuity lines gd and ec (Figure 5) or sb and kt (Figure 13) can be too large for small cracks. Note that the kinematically admissible velocity fields for the specimens with no crack (Figures 6 and 8) are also kinematically admissible for the corresponding specimens with cracks. Therefore, f_u from (16) at $a = 0$ is $f_u^{(3)}$ for the specimen shown in Figure 4 and f_u from (63) is $f_u^{(3)}$ for the specimen shown in Figure 12. For such specimens the final expression for the limit load based on the three kinematically admissible velocity fields proposed is

$$f_u = \min\{f_u^{(1)}, f_u^{(2)}, f_u^{(3)}\} \quad (110)$$

As an example, consider the special case of the configuration shown in Figure 5 for which $f_u^{(1)}$ is given by (27). The value of $f_u^{(3)}$ is obtained from (27) at $a = 0$ and $\varepsilon = 0$. As a result,

$$f_u^{(3)} = \frac{1}{2\sqrt{3}}\left(3+\frac{W}{H}\right) \quad (111)$$

In order to choose between $f_u^{(1)}$ and $f_u^{(3)}$, it is necessary to solve the equation $f_u^{(1)} = f_u^{(3)}$. Using (27) and (111) this equation transforms to

$$\left(1 - \frac{a}{W}\right) \left[3 + \frac{(W-a)}{H}\right] + \frac{2\varepsilon}{W} = 3 + \frac{W}{H} \quad (112)$$

The critical value of ε can be found from (112) in the form

$$\frac{2\varepsilon_{cr}}{W} = \frac{a}{H} + \frac{a}{W} \left[3 + \frac{(W-a)}{H}\right] \quad (113)$$

Then, it is necessary to choose $f_u^{(1)}$ if $\varepsilon \leq \varepsilon_{cr}$ and $f_u^{(3)}$ if $\varepsilon \geq \varepsilon_{cr}$. The final result of this calculation should be compared to the value of $f_u^{(2)}$ found by means of (92) where $\beta = 0$. Then, equation (110) should be used.

3.6. Pure Bending

Geometry of the specimen, the system of loading and the axes of the Cartesian coordinate system xy are shown in Figure 17 where $2H$ is the thickness of the weld and $2W$ is the width of the specimen. The rigid zones of base material rotate with an angular velocity ω . The axes x and y coincide with the axes of symmetry of the specimen. Because of symmetry it is sufficient to get the solution in the domain $x \geq 0$ and $y \geq 0$.

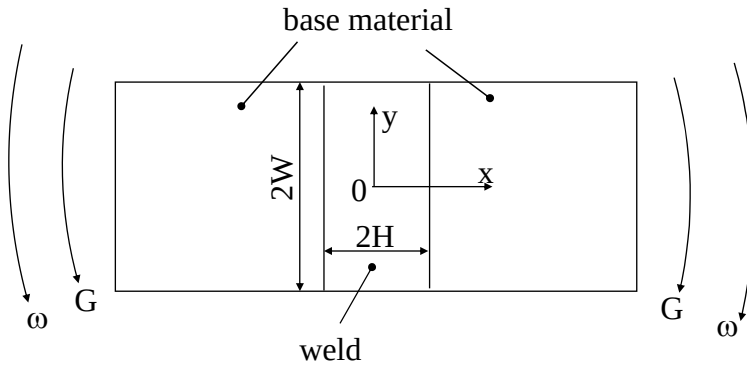


Figure 17. Geometry of structure under consideration – notation

Two different solutions have been proposed in Alexandrov and Kocak (2007) and Alexandrov (2008). The former is based on the kinematically admissible velocity field whose general structure within one quarter of the weld is shown in Figure 18. The rigid zone rotates along with the base material. The straight rigid plastic boundary Ob is a velocity discontinuity line. In order to find the kinematically admissible velocity field in the plastic zone, it is possible to adopt the exact solution to the complete system of equations of plane-strain plasticity in the domain $0 \leq r < \infty$ and $\alpha \leq \theta \leq \pi/2$ where the plane polar coordinate system $r\theta$ is defined by the following transformation equations

$$x = r \cos \theta, \quad y = r \sin \theta \quad (114)$$

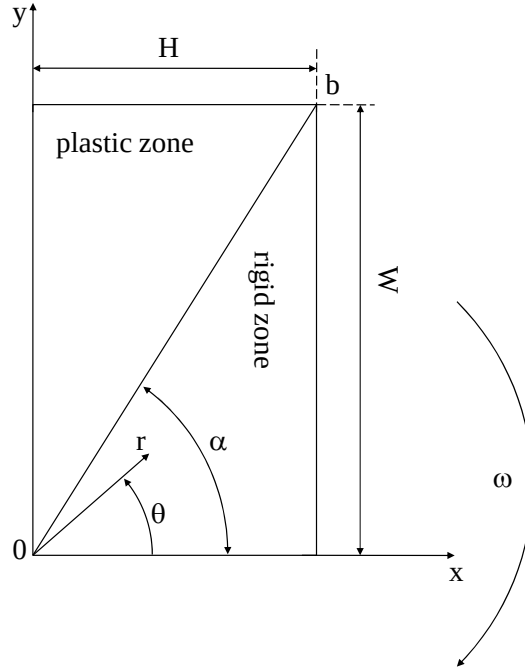


Figure 18. General structure of the kinematically admissible velocity field

The orientation of the rigid plastic boundary is determined from geometry of the specimen as

$$\tan \alpha = \frac{W}{H} \quad (115)$$

The equilibrium equations in the polar coordinate system have the form

$$\frac{\partial \sigma_{rr}}{\partial r} + \frac{1}{r} \frac{\partial \sigma_{r\theta}}{\partial \theta} + \frac{\sigma_{rr} - \sigma_{\theta\theta}}{r} = 0, \quad \frac{\partial \sigma_{r\theta}}{\partial r} + \frac{1}{r} \frac{\partial \sigma_{\theta\theta}}{\partial \theta} + \frac{2\sigma_{r\theta}}{r} = 0 \quad (116)$$

where σ_{rr} , $\sigma_{\theta\theta}$ and $\sigma_{r\theta}$ are the components of the stress tensor in the polar coordinate system. The plane-strain yield criterion is satisfied by the standard substitution (Hill, 1950)

$$\sigma_{rr} = \sigma + \frac{\sigma_0}{\sqrt{3}} \cos 2\psi, \quad \sigma_{\theta\theta} = \sigma - \frac{\sigma_0}{\sqrt{3}} \cos 2\psi, \quad \sigma_{r\theta} = \frac{\sigma_0}{\sqrt{3}} \sin 2\psi \quad (117)$$

where σ is the hydrostatic stress and ψ is the orientation of the major principal stress relative to the r -axis. The main assumption is that ψ is independent of r . Then, substituting (117) into (116) leads to

$$r \frac{\partial \sigma}{\partial r} + \frac{2\sigma_0}{\sqrt{3}} \cos 2\psi \left(\frac{d\psi}{d\theta} + 1 \right) = 0, \quad \frac{\partial \sigma}{\partial \theta} + \frac{2\sigma_0}{\sqrt{3}} \sin 2\psi \left(\frac{d\psi}{d\theta} + 1 \right) = 0 \quad (118)$$

These equations are compatible if and only if

$$\frac{\sqrt{3}\sigma}{\sigma_0} = A \ln r + p(\theta) \quad (119)$$

where A is constant and $p(\theta)$ is an arbitrary function of θ . Substituting (119) into the first equation of the system (118) gives

$$\frac{d\psi}{d\theta} = -\frac{(A + 2 \cos 2\psi)}{2 \cos 2\psi} \quad (120)$$

Since $0b$ is the velocity discontinuity line (Figure 18), $|\sigma_{r\theta}| = \sigma_0/\sqrt{3}$ at $\theta = \alpha$. Moreover, the direction of plastic flow in the vicinity of the velocity discontinuity line (towards the origin of the coordinate system) requires that $\sigma_{r\theta} < 0$. Therefore, one of the boundary conditions for equation (120) is determined from (117) as

$$\psi = -\frac{\pi}{4} \quad (121)$$

for $\theta = \alpha$. The other boundary condition follows from the condition at the axis of symmetry $x = 0$ where $\sigma_{r\theta} = 0$. In addition, it is necessary to mention that material fibers perpendicular to the axis of symmetry are subject to tension. Therefore, $\sigma_{\theta\theta} - \sigma_{rr} > 0$. Taking into account this inequality and the condition $\sigma_{r\theta} = 0$, it can be found from (117) that

$$\psi = -\frac{\pi}{2} \quad (122)$$

for $\theta = \pi/2$. Even though there are the two boundary conditions for the first order differential equation (120), there is no contradiction because its right-hand side involves an arbitrary constant A . The solution to equation (120) determines ψ as a function of θ .

Substituting this solution and (119) into the second equation of the system (118) and integrating determine the function $p(\theta)$. However, this function is not essential for the limit load in question and, therefore, is not determined here. The general solution to equation (120) can be written in an analytic form. However, the final expression is cumbersome and, therefore, it is more convenient to use the solution to (120) in the following form

$$\theta = -2 \int_{-\pi/2}^{\psi} \frac{\cos 2\chi d\chi}{(A + 2\cos 2\chi)} + \frac{\pi}{2} \quad (123)$$

where χ is a dummy variable of integration. The solution in the form of (123) satisfies the boundary condition (122). Combining the solution (123) and the boundary condition (121) results in the following equation for A

$$\alpha = -2 \int_{-\pi/2}^{-\pi/4} \frac{\cos 2\chi d\chi}{(A + 2\cos 2\chi)} + \frac{\pi}{2} \quad (124)$$

This equation determines A as a function of α or, taking into account (115), as a function of H/W . This function is illustrated in Figure 19. Thus, the stress field found satisfies the equilibrium equations in the plastic zone and the stress boundary conditions at $\theta = \alpha$ and $\theta = \pi/2$. Even though these equations and conditions are not involved in the upper bound theorem, it is advantageous to use the stress solution for constructing the kinematically admissible velocity field.

The circumferential velocity can be assumed in the form

$$u_\theta = -\omega r u_0(\theta) \quad (125)$$

where $u_0(\theta)$ is an arbitrary function of θ or ψ because ψ is the function of θ due to (123). The function $u_0(\theta)$ must satisfy the following boundary conditions

$$u_0 = 0 \quad (126)$$

for $\theta = \pi/2$ (or $\psi = -\pi/2$) and

$$u_0 = 1 \quad (127)$$

for $\theta = \alpha$ (or $\psi = -\pi/4$). The condition (126) is a symmetry condition and the condition (127) follows from the continuity of the normal velocity across the velocity discontinuity line

$\theta = \alpha$. Using (125) the equation of incompressibility in the polar coordinate can be written in the form

$$\frac{\partial u_r}{\partial r} + \frac{u_r}{r} - \omega \frac{du_0}{d\theta} = 0 \quad (128)$$

where u_r is the radial velocity. The general solution of equation (128) is

$$u_r = \frac{\omega r}{2} \frac{du_0}{d\theta} + \frac{C}{r} \quad (129)$$

where C is a constant of integration. It is necessary to put $C = 0$, otherwise $|u_r| \rightarrow \infty$ as $r \rightarrow 0$. Thus, equation (129) becomes

$$u_r = \frac{\omega r}{2} \frac{du_0}{d\theta} \quad (130)$$

For the problem under consideration, the associated flow rule reduces to

$$\frac{\sigma_{rr} - \sigma_{\theta\theta}}{\sigma_{r\theta}} = \frac{\xi_{rr} - \xi_{\theta\theta}}{\xi_{r\theta}} \quad (131)$$

where ξ_{rr} , $\xi_{\theta\theta}$ and $\xi_{r\theta}$ are the components of the strain rate tensor in the polar coordinates. Substituting (117), (125) and (126) into (131) gives

$$\frac{d^2 u_0}{d\theta^2} = 2 \tan 2\psi \frac{du_0}{d\theta} \quad (132)$$

Using (120) differentiation with respect to θ in this equation can be replaced with differentiation with respect to ψ to arrive at

$$\cos 2\psi \frac{d^2 u_0}{d\psi^2} + 2 \sin 2\psi \frac{du_0}{d\psi} = 0 \quad (133)$$

The general solution of this equation is

$$u_0 = C_1 \sin 2\psi + C_2 \quad (134)$$

where C_1 and C_2 are constants of integration. Using the boundary conditions (126) and (127) these constants are expressed as $C_1 = -1$ and $C_2 = 0$. Then, the solution (134) becomes

$$u_0 = -\sin 2\psi \quad (135)$$

Substituting (135) into (125) and (130) and using (120) give the velocity field in the form

$$u_\theta = \omega r \sin 2\psi, \quad u_r = \frac{\omega r}{2} (A + 2 \cos 2\psi) \quad (136)$$

This velocity field is taken as the kinematically admissible velocity field in the plastic zone (Figure 18). The corresponding strain rate components are calculated from (136), with the use of (120), as

$$\zeta_{rr} = \frac{\omega}{2} (A + 2 \cos 2\psi), \quad \zeta_{\theta\theta} = -\frac{\omega}{2} (A + 2 \cos 2\psi), \quad \zeta_{r\theta} = \frac{\omega}{2} \tan 2\psi (A + 2 \cos 2\psi) \quad (137)$$

The solution to equation (124) illustrated in Figure 19 shows that $A \leq 0$ and $A + 2 \cos 2\psi < 0$ for any ψ of the interval $-\pi/4 \geq \psi \geq -\pi/2$. Moreover, $\cos 2\psi \leq 0$ within this interval. Therefore, the equivalent strain rate is determined from (137) as

$$\zeta_{eq} = \frac{\omega (A + 2 \cos 2\psi)}{\sqrt{3} \cos 2\psi} \quad (138)$$

It is seen from this equation that the equivalent strain rate approaches infinity near the velocity discontinuity surface $\theta = \alpha$ where $\psi = -\pi/4$. This result is in agreement with (10). In fact, it is possible to show that the asymptotic behavior of the equivalent strain rate given by (138) exactly follows the rule (10).

The amount of velocity jump across the velocity discontinuity surface $\theta = \alpha$ is equal to the radial velocity in the plastic zone at $\psi = -\pi/4$. Therefore, it follows from (136) and the condition $A < 0$ (Figure 19) that

$$[[u_r]] = -\frac{\omega r A}{2} \quad (139)$$

Equation (9) in the case under consideration becomes

$$\frac{G_u \omega}{2} = \sigma_0 \iiint_V \zeta_{eq} dV + \frac{\sigma_0}{\sqrt{3}} \iint_{S_d} [u_\tau] dS \quad (140)$$

It has been taken into account here that integration should be carried out over a quarter of the specimen. It follows from (120) that

$$r d\theta dr = -\frac{2r \cos 2\psi}{(A + 2 \cos 2\psi)} d\psi dr \quad (141)$$

Substituting (138), (139) and (141) in (140) and integrating with respect to r from 0 to $y = W$ (or, as follows from (114), to $W/\sin \theta$) give

$$\frac{G_u}{2\sigma_0 B W^2} = g_u = -\frac{2}{\sqrt{3}} \int_{-\pi/4}^{-\pi/2} \frac{d\psi}{\sin^2 \theta} - \frac{A}{2\sqrt{3} \sin^2 \alpha} \quad (142)$$

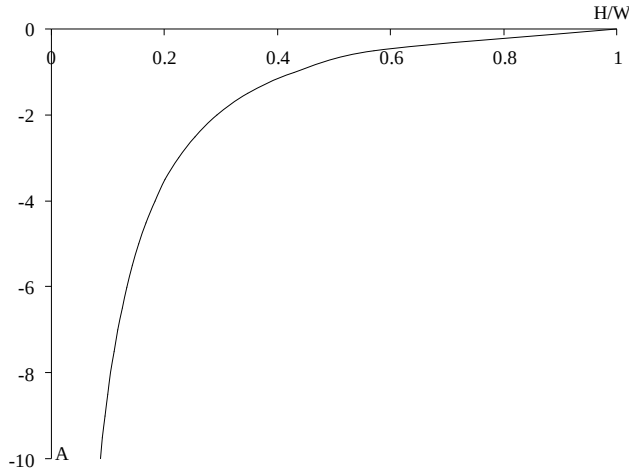


Figure 19. Variation of A with H/W

Integration here can be completed numerically with no difficulty because θ is a function ψ due to (123). In particular, the dependence of the dimensionless bending moment g_u on H/W is illustrated in Figure 20.

Another solution for the configuration shown in Figure 17 has been proposed in Alexandrov (2008). The general structure of the kinematically admissible velocity field within one quarter of the weld is shown in Figure 21. The rigid zone rotates along with the base material about the origin of the Cartesian coordinate system with an angular velocity ω . Therefore, the velocity vector in this zone can be represented as

$$\mathbf{u}^{(r)} = \omega y \mathbf{i} - \omega x \mathbf{j} \quad (143)$$

where $\mathbf{i}$ and $\mathbf{j}$ are the base vectors of the Cartesian coordinate system. The boundary conditions for the velocity u_x in the plastic zone are

$$u_x = \omega y \quad (144)$$

at $x = H$ and

$$u_x = 0 \quad (145)$$

at $x = 0$. The condition (145) is a symmetry condition similar to (126) and the condition (144) follows from the continuity of the normal velocity across the velocity discontinuity line $x = H$ coinciding with the interface between the weld and base materials. The simplest representation for the velocity u_x satisfying the boundary conditions (144) and (145) is

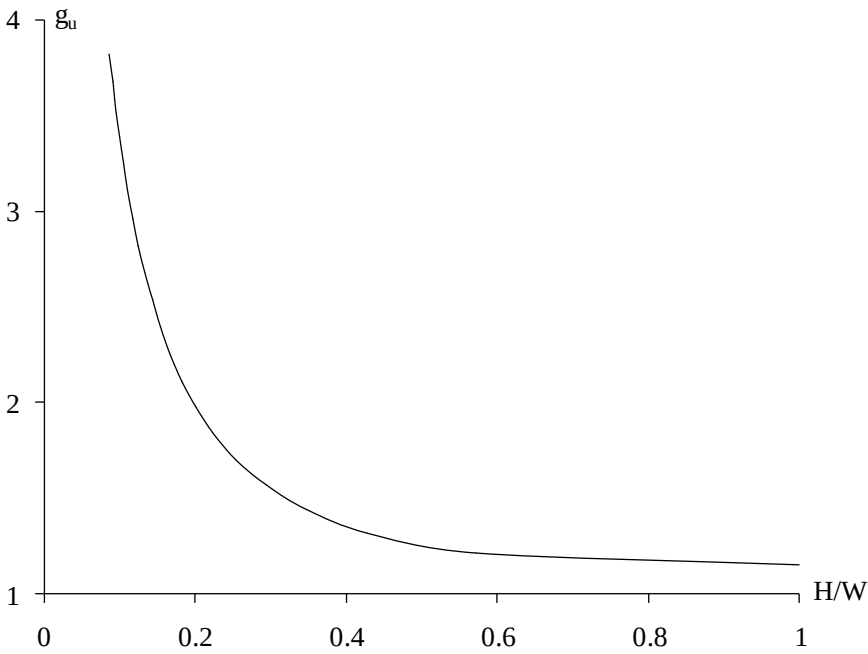


Figure 20. Variation of the dimensionless bending moment with H/W

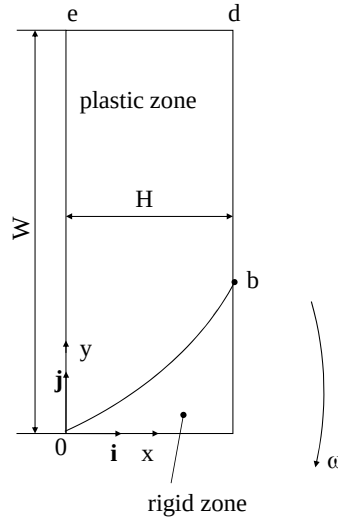


Figure 21. General structure of the kinematically admissible velocity field

$$u_x = \frac{\omega y x}{H} \quad (146)$$

Substituting (146) into the incompressibility equation $\partial u_x / \partial x + \partial u_y / \partial y = 0$ and integrating give the velocity u_y in the form

$$u_y = -\frac{\omega y^2}{2H} + \omega H \eta(x) \quad (147)$$

where $\eta(x)$ is an arbitrary function of x . In order to propose the specific function $\eta(x)$, it is advantageous to account for (10). One of the simplest representations of $\eta(x)$ satisfying this condition is

$$\eta(x) = \eta_0 + \eta_1 \sqrt{1 - \left(\frac{x}{H}\right)^2} \quad (148)$$

where η_0 and η_1 are arbitrary constants. It is convenient to introduce the following dimensionless quantities

$$\chi = \frac{y}{W}, \quad \sin \gamma = \frac{x}{H} \quad (149)$$

Taking into account (148) and (149) the kinematically admissible velocity field in the plastic zone given by (146) and (147) can be written in the form

$$\frac{\mathbf{u}^{(p)}}{\omega W} = \chi \sin \gamma \mathbf{i} + \left[\frac{H}{W} (\eta_0 + \eta_1 \cos \gamma) - \frac{\chi^2 W}{2H} \right] \mathbf{j} \quad (150)$$

Also, equation (143) transforms to

$$\frac{\mathbf{u}^{(r)}}{\omega W} = \chi \mathbf{i} - \frac{H}{W} \sin \gamma \mathbf{j} \quad (151)$$

Let φ be the angle between the tangent to the velocity discontinuity line Ob (Figure 21) and the x -axis, measured from the axis anti-clockwise (Figure 22). Then, the unit normal is represented by

$$\mathbf{n} = -\sin \varphi \mathbf{i} + \cos \varphi \mathbf{j} \quad (152)$$

Since, by definition, $\tan \varphi = dy/dx$, it follows from (149) that

$$\tan \varphi = \frac{W}{H \cos \gamma} \frac{d\chi}{d\gamma} \quad (153)$$

The normal velocity must be continuous across the velocity discontinuity line Ob (Figure 21). Therefore, $\mathbf{u}^{(r)} \cdot \mathbf{n} = \mathbf{u}^{(p)} \cdot \mathbf{n}$. Substituting (150), (151) and (152) into this equation and using (153) result in

$$\frac{2\chi(1-\sin \gamma)}{\cos \gamma} \frac{d\chi}{d\gamma} = \chi^2 - 2(\eta_0 + \eta_1 \cos \gamma + \sin \gamma) \left(\frac{H}{W} \right)^2 \quad (154)$$

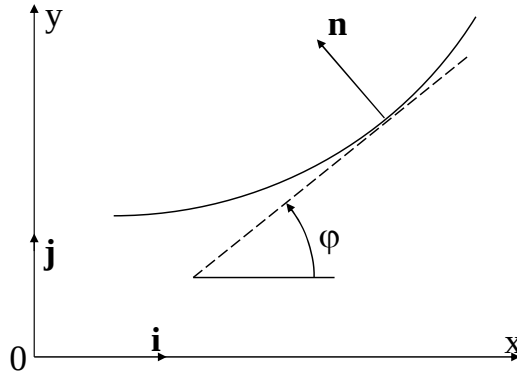


Figure 22. Geometry of velocity discontinuity line

The solution to this equation determines the shape of the line Ob . The velocity fields $\mathbf{u}^{(r)}$ and $\mathbf{u}^{(p)}$ can be kinematically admissible if and only if this line passes through the origin of the coordinate system. Therefore, the boundary condition for equation (154) is

$$\chi = 0 \quad (155)$$

at $\gamma = 0$. Equation (154) is reduced to a linear ordinary differential equation of first order by substitution $\mathcal{G} = \chi^2$. Therefore, its general solution can be found with no difficulty. The particular solution satisfying the boundary condition (155) has the form

$$\mathcal{G}_{0b}(\gamma) = -\left(\frac{H}{W}\right)^2 \left[\frac{\sin^2 \gamma + \eta_1(\gamma + \sin \gamma \cos \gamma) + 2\eta_0 \sin \gamma}{1 - \sin \gamma} \right] \quad (156)$$

The notation for $\mathcal{G}_{0b}(\gamma)$ emphasizes that equation (156) gives the dependence of $\mathcal{G}$ on γ along the line Ob . It follows from (156) that, in general, $\mathcal{G}_{0b} \rightarrow \infty$ as $\gamma \rightarrow \pi/2$ (or $x \rightarrow H$). In order to obtain a finite value of $\mathcal{G}_{0b}$ as $\gamma \rightarrow \pi/2$, it is necessary to put

$$2\eta_0 = -(1 + \pi\eta_1/2) \quad (157)$$

In this case equation (156) transforms to

$$\mathcal{G}_{0b}(\gamma) = \left(\frac{H}{W}\right)^2 \left[\frac{2\sin \gamma(1 - \sin \gamma) - \eta_1(2\gamma + \sin 2\gamma - \pi \sin \gamma)}{2(1 - \sin \gamma)} \right] \quad (158)$$

The value of $\mathcal{G}_{0b}$ at $\gamma = \pi/2$ corresponding to point b (Figure 21) is determined from (158) by applying l'Hospital's rule

$$\mathcal{G}_b = \left(\frac{H}{W}\right)^2 \left(1 - \frac{\eta_1 \pi}{2}\right), \quad \chi_b = \left(\frac{H}{W}\right) \sqrt{1 - \frac{\eta_1 \pi}{2}} \quad (159)$$

The value of χ_b has been calculated with the use of the definition for $\mathcal{G}$. Using (149) and (150) it is possible to find the components of the strain rate tensor in the plastic zone in the form

$$\zeta_{xx} = \frac{\omega W \chi}{H}, \quad \zeta_{yy} = -\frac{\omega W \chi}{H}, \quad \zeta_{xy} = \frac{\omega \sin \gamma}{2} \left(1 - \frac{\eta_1}{\cos \gamma}\right) \quad (160)$$

This expression for ξ_{xy} and (149) show that the kinematically admissible velocity field (150) and the associated flow rule result in a stress field that satisfies the stress boundary condition at the axis of symmetry $x = 0$ where the shear stress vanishes. It is an additional advantage of the kinematically admissible velocity field chosen. Combining (13) and (160) gives the equivalent strain rate in the form

$$\zeta_{eq} = \frac{\omega W}{\sqrt{3}H \cos \gamma} \sqrt{4\chi^2 \cos^2 \gamma + \left(\frac{H}{W}\right)^2 \sin^2 \gamma (\cos \gamma - \eta_1)^2} \quad (161)$$

The amount of velocity jump across the velocity discontinuity line $0b$ is determined as

$$\|u_\tau\|_{0b} = \sqrt{\left(u_x^{(p)} - u_x^{(r)}\right)^2 + \left(u_y^{(p)} - u_y^{(r)}\right)^2} \quad (162)$$

Here the velocity components should be taken at $\mathcal{G} = \mathcal{G}_{0b}(\gamma)$ where the function $\mathcal{G}_{0b}(\gamma)$ is given by (158). Then, using (150), (151) and (157) equation (162) can be transformed to

$$\|u_\tau\|_{0b} = \omega W \sqrt{(1 - \sin \gamma)^2 \mathcal{G}_{0b}(\gamma) + \left[\left(\eta_1 \cos \gamma + \sin \gamma - \frac{1}{2} - \frac{\pi \eta_1}{4} \right) \frac{H}{W} - \frac{\mathcal{G}_{0b}(\gamma)W}{2H} \right]^2} \quad (163)$$

The infinitesimal length element of the velocity discontinuity line is determined by

$$dl_{0b} = \sqrt{(dx)^2 + (dy)^2} \quad (164)$$

where dx and dy should be replaced with $d\gamma$ and $d\chi$ by means of (149), χ^2 with $\mathcal{G}_{0b}(\gamma)$ and $d\chi/d\gamma$ should be excluded with the use of equation (154). Then, equation (164) becomes, with the use of (157),

$$dl_{0b} = \frac{H \cos \gamma}{\sqrt{\mathcal{G}_{0b}(\gamma)(1 - \sin \gamma)}} \sqrt{(1 - \sin \gamma)^2 \mathcal{G}_{0b}(\gamma) + \left[\frac{\mathcal{G}_{0b}(\gamma)W}{2H} - \left(\eta_1 \cos \gamma + \sin \gamma - \frac{1}{2} - \frac{\pi \eta_1}{4} \right) \frac{H}{W} \right]^2} d\gamma \quad (165)$$

The other velocity discontinuity line occurs at $x = H$ in the range $\chi_b \leq \chi \leq 1$ (between points b and d in Figure 21) where χ_b is given by (159). The amount of velocity

jump is $[[u_\tau]]_{bd} = |u_y^{(p)} - u_y^{(r)}|$ where the velocity components should be taken at $x = H$ (or $\gamma = \pi/2$). Therefore, it follows from (150), (151) and (157) that

$$[[u_\tau]]_{bd} = \frac{\omega H}{2} \left(\frac{\chi^2 W^2}{H^2} + \frac{\pi \eta_1}{2} - 1 \right) \quad (166)$$

It has been assumed here that

$$\frac{\chi^2 W^2}{H^2} + \frac{\pi \eta_1}{2} - 1 \geq 0 \quad (167)$$

This inequality should be verified *a posteriori*.

The first term on the right hand side of (9) is determined from (149) and (161) in the form

$$\begin{aligned} \sigma_0 \iiint_V \zeta_{eq} dV &= \frac{2\omega B W^2 \sigma_0}{\sqrt{3}} \Omega_V(\eta_1), \\ \Omega_V(\eta_1) &= \int_0^{\pi/2} \int_{\sqrt{\zeta_{ob}(\gamma)}}^1 \sqrt{4\chi^2 \cos^2 \gamma + \left(\frac{H}{W}\right)^2 \sin^2 \gamma (\cos \gamma - \eta_1)^2} d\chi d\gamma \end{aligned} \quad (168)$$

The second term in the case under consideration becomes

$$\frac{\sigma_0}{\sqrt{3}} \iint_{S_d} [[u_\tau]] dS = \frac{2B\sigma_0}{\sqrt{3}} \int_l [[u_\tau]]_{ob} dl_{ob} + \frac{2B\sigma_0}{\sqrt{3}} \int_{y_b}^{y_d} [[u_\tau]]_{bd} dy \quad (169)$$

where y_b and y_d are the y -coordinates of points b and d , respectively (Figure 21). It follows from (163) and (165) that

$$\begin{aligned} \frac{2B\sigma_0}{\sqrt{3}} \int_l [[u_\tau]]_{ob} dl_{ob} &= \frac{2\omega B W H \sigma_0}{\sqrt{3}} \Omega_{ob}(\eta_1), \\ \Omega_{ob}(\eta_1) &= \int_0^{\pi/2} \frac{\cos \gamma}{\sqrt{\mathcal{G}_{ob}(\gamma)(1 - \sin \gamma)}} \left\{ (1 - \sin \gamma)^2 \mathcal{G}_{ob}(\gamma) + \left[\frac{\mathcal{G}_{ob}(\gamma)W}{2H} - \left(\eta_1 \cos \gamma + \sin \gamma - \frac{1}{2} - \frac{\pi \eta_1}{4} \right) \frac{H}{W} \right]^2 \right\} d\gamma \end{aligned} \quad (170)$$

Also, from (166) with the use of (149)

$$\frac{2B\sigma_0}{\sqrt{3}} \int_{y_b}^{y_d} \left[u_\tau \right]_{bd} dy = \frac{\omega BWH\sigma_0}{\sqrt{3}} \int_{\chi_b}^1 \left(\frac{\chi^2 W^2}{H^2} + \frac{\pi\eta_1}{2} - 1 \right) d\chi = \frac{\omega BWH\sigma_0}{\sqrt{3}} \Omega_{bd}(\eta_1), \quad (171)$$

$$\Omega_{bd}(\eta_1) = \left[\frac{1}{3} \left(\frac{W}{H} \right)^2 (1 - \chi_b^3) + \left(\frac{\pi\eta_1}{2} - 1 \right) (1 - \chi_b) \right]$$

The left hand side of (9) reduces to $G_u \omega / 2$. Therefore, equation (9) becomes

$$\frac{G_u}{2BW^2\sigma_0} = g_u = \frac{2}{\sqrt{3}} \Omega_v(\eta_1) + \frac{2H}{\sqrt{3}W} \Omega_{ob}(\eta_1) + \frac{H}{\sqrt{3}W} \Omega_{bd}(\eta_1) \quad (172)$$

Here $\Omega_v(\eta_1)$, $\Omega_{ob}(\eta_1)$ and $\Omega_{bd}(\eta_1)$ can be found from (168), (170) and (171). The notation for these quantities emphasizes that they depend on η_1 . Therefore, according to the upper bound theorem, the right hand side of (172) should be minimized with respect to η_1 . As a result of numerical minimization the values of η_1 and g_u have been obtained. The inequality (167) was checked in course of calculation. The variation of g_u with H/W is depicted in Figure 23.

Let $g_u^{(1)}$ be the value of g_u given in equation (142) and $g_u^{(2)}$ be the value of g_u given in equation (172). According to the upper bound theorem, the final result based on the two kinematically admissible velocity fields proposed is

$$g_u = \min \{ g_u^{(1)}, g_u^{(2)} \} \quad (173)$$

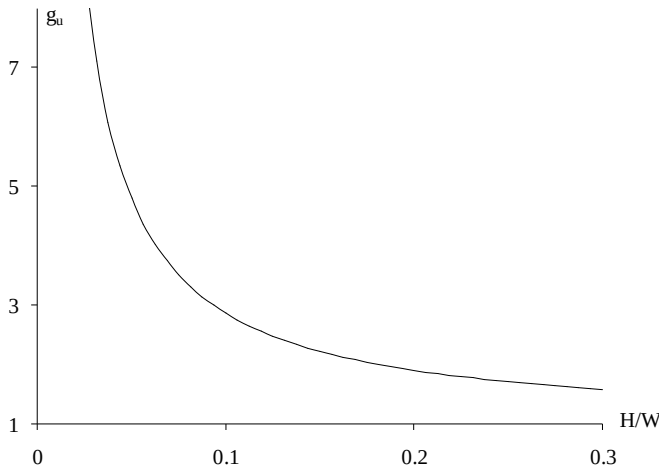


Figure 23. Variation of dimensionless bending moment with H/W

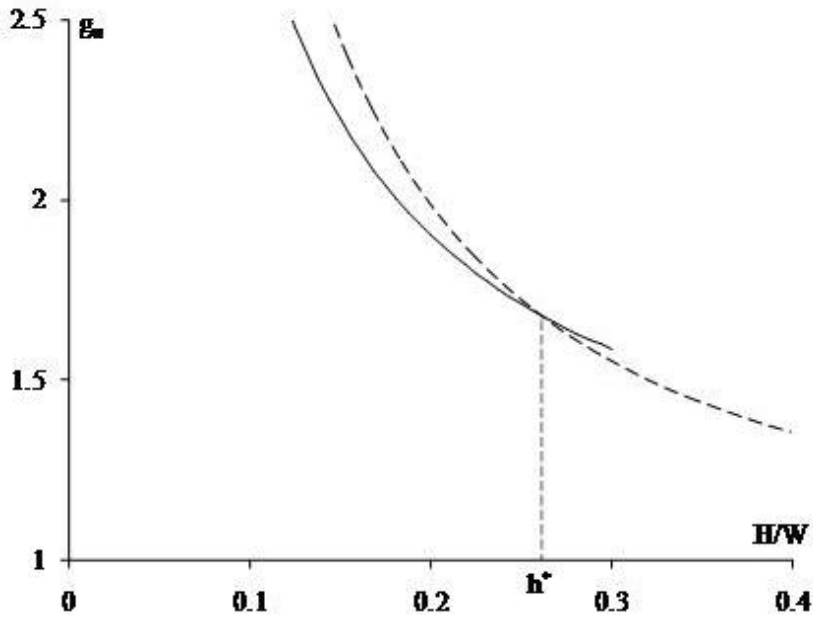


Figure 24. Comparison between the bending moments from equations (142) and (172)

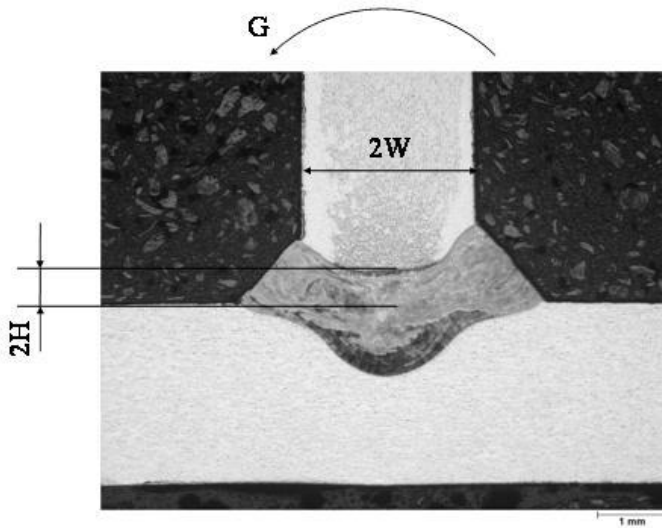


Figure 25. Welded T-joint under bending

The variation of both $g_u^{(1)}$ and $g_u^{(2)}$ with H/W is depicted in Figure 24 where the dashed line corresponds to $g_u^{(1)}$ and the solid line to $g_u^{(2)}$. It is seen from this figure that the curves intersect at $H/W = h^*$ such that $g_u^{(1)} > g_u^{(2)}$ in the range $H/W < h^*$ and $g_u^{(1)} < g_u^{(2)}$ in the range $H/W > h^*$. Thus (173) can be rewritten in the form

$$g_u = \begin{cases} g_u^{(1)} & \text{for } H/W > h^* \\ g_u^{(2)} & \text{for } H/W < h^* \end{cases} \quad (174)$$

It follows from the numerical solution that $h^* \approx 0.26$.

The solution (174) is also applicable for welded T-joints under bending (Figure 25). This kind of joints is widely used in thin-walled aerospace structures (Alexandrov and Kocak, 2007).

4. AXISYMMETRIC SOLUTIONS

In the case of axisymmetric problems it is often convenient to adopt a cylindrical coordinate system $r\theta z$. In such a coordinate system $\zeta_{r\theta} = \zeta_{z\theta} = 0$ and $\zeta_{r\theta} = u_r/r$ where u_r is the radial velocity (u_z will stand for the axial velocity). Therefore, the equation of incompressibility (5) becomes

$$\frac{\partial u_r}{\partial r} + \frac{\partial u_z}{\partial z} + \frac{u_r}{r} = 0 \quad (175)$$

and equation (4) transforms to

$$\zeta_{eq} = \sqrt{\frac{2}{3}} \sqrt{\left(\frac{\partial u_r}{\partial r}\right)^2 + \left(\frac{\partial u_z}{\partial z}\right)^2 + \left(\frac{u_r}{r}\right)^2 + \frac{1}{2} \left(\frac{\partial u_r}{\partial z} + \frac{\partial u_z}{\partial r}\right)^2} \quad (176)$$

4.1. Round Bar with an Axisymmetric Crack at the Plane of Symmetry under Tension

Geometry of the specimen, the system of loading, the direction of velocity of the rigid blocks of base material U and the cylindrical coordinate system are shown in Figure 26 where $2H$ is the thickness of the weld, R is the radius of the specimen, and a is the radius of the crack. A limit load solution for this configuration has been proposed in Alexandrov *et.al.* (1999^b). An improved solution is provided in this section.

Because of symmetry it is sufficient to consider the domain $z \geq 0$. The general structure of the kinematically admissible velocity field within the weld is shown in Figure 27 where bc is the rigid plastic boundary and is also a velocity discontinuity line. Another velocity discontinuity line coincides with the interface between the weld and base materials between point c and d . The rigid zone moves along the z -axis with velocity U .

The velocity boundary conditions are

$$u_z = 0 \quad (177)$$

for $z = 0$ and

$$u_z = U \quad (178)$$

for $z = H$. The normal velocity must be continuous across the line bc . In order to construct the kinematically admissible velocity in the plastic zone, it is natural to start with the assumption

$$u_z = \frac{U}{H} z \quad (179)$$

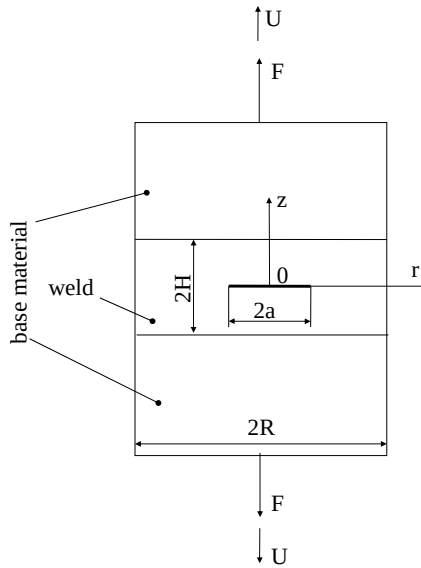


Figure 26. Geometry of structure under consideration – notation

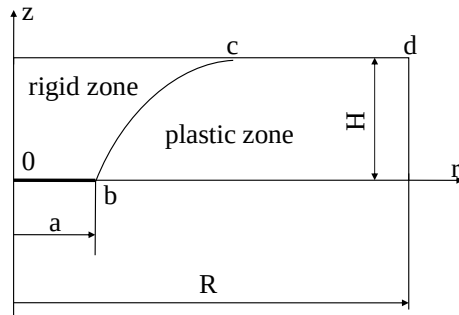


Figure 27. General structure of the kinematically admissible velocity field

Then, the boundary conditions (177) and (178) are satisfied. Substituting (179) into (175) and integrating give

$$\frac{u_r}{U} = \frac{C(z)R}{r} - \frac{r}{2H} \quad (180)$$

where $C(z)$ is an arbitrary function of z . Since the normal strain rates are bounded, the condition (10) is equivalent to

$$\zeta_{rz} = O\left(\frac{1}{\sqrt{H-z}}\right), \quad z \rightarrow H \quad (181)$$

near the velocity discontinuity line cd . It follows from (179) and (180) that one of the simplest functions $C(z)$ that satisfies (181) is

$$C(z) = C_0 + C_1 \sqrt{1 - \left(\frac{z}{H}\right)^2} \quad (182)$$

where C_0 and C_1 are constants. Since $dC/dz = 0$ at $z = 0$, an advantage of the representation (182) is that $\zeta_{rz} = 0$ at $z = 0$ as in the exact solution. Substituting (182) into (180) and, then, (179) and (180) into (176) lead to

$$\zeta_{eq} = \frac{U}{H} \sqrt{1 + \frac{4(C_0 + C_1 \cos \mu)^2 h^2}{3\rho^4} + \frac{C_1^2 \tan^2 \mu}{3\rho^2}} \quad (183)$$

where

$$h = \frac{H}{R}, \quad \rho = \frac{r}{R}, \quad \sin \mu = \frac{z}{H} \quad (184)$$

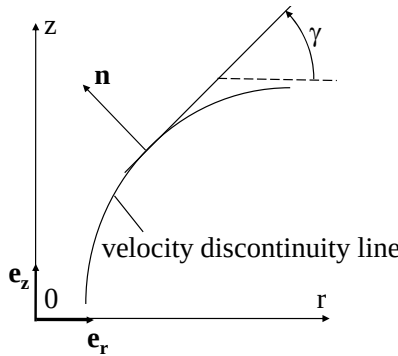


Figure 28. Geometry of velocity discontinuity line bc

The unit normal to the velocity discontinuity line bc is represented by (Figure 28)

$$\mathbf{n} = -\sin \gamma \mathbf{e}_r + \cos \gamma \mathbf{e}_z \quad (185)$$

where $\mathbf{e}_r$ and $\mathbf{e}_z$ are the base vectors of the cylindrical coordinate system and γ is the orientation of the tangent to the velocity discontinuity line bc . The continuity of the normal velocity across the velocity discontinuity line bc requires $\mathbf{u}_R \cdot \mathbf{n} = \mathbf{u}_P \cdot \mathbf{n}$ where $\mathbf{u}_R = U\mathbf{e}_z$ is the velocity vector in the rigid zone and $\mathbf{u}_P$ is the velocity vector in the plastic zone whose components are given by (179) and (180). Using (179), (180), (182), (184), and (185) this equation transforms to

$$\cos \gamma = \sin \mu \cos \gamma - \left(\frac{C_0 + C_1 \cos \mu}{\rho} - \frac{\rho}{2h} \right) \sin \gamma \quad (186)$$

By definition, $\tan \gamma = dz/dr$ or, taking into account (184), $\tan \gamma = h \cos \mu d\mu/d\rho$. Therefore, equation (186) becomes

$$\left[\frac{\rho}{2} - \frac{h(C_0 + C_1 \cos \mu)}{\rho} \right] \frac{d\mu}{d\rho} = \frac{1 - \sin \mu}{\cos \mu} \quad (187)$$

This equation can be reduced to a linear ordinary differential equation by the substitution $\chi = \rho^2$

$$\frac{d\chi}{d\mu} = \frac{\cos \mu}{(1 - \sin \mu)} [\chi - 2h(C_0 + C_1 \cos \mu)] \quad (188)$$

The general solution of this equation can be found with no difficulty and has the form, with the use of the definition for χ ,

$$\rho^2 = -\frac{h[4C_0 \sin \mu + C_1(2\mu + \sin 2\mu) + C_2]}{2(1 - \sin \mu)} \quad (189)$$

where C_2 is a constant of integration. The curve (189) should pass through the crack tip (Figure 27). Therefore, with the use of (184), $\rho = a/R = a_0$ for $\mu = 0$. Substituting this condition into (189) gives $C_2 = -2a_0^2/h$. Then, equation (189) becomes

$$\rho^2 = -\frac{h[4C_0 \sin \mu + C_1(2\mu + \sin 2\mu)] - 2a_0^2}{2(1 - \sin \mu)} \quad (190)$$

It is seen from this equation that $\rho \rightarrow \infty$ as $\mu \rightarrow \pi/2$ (or $z \rightarrow H$), unless

$$4C_0 = \frac{2a_0^2}{h} - C_1\pi \quad (191)$$

Excluding C_0 in (190) by means of (191) results in

$$\rho_{bc}^2(\mu) = a_0^2 - \frac{hC_1}{2} \frac{(2\mu + \sin 2\mu - \pi \sin \mu)}{(1 - \sin \mu)} \quad (192)$$

Here the subscript bc emphasizes that equation (192) determines the velocity discontinuity line bc . Using l'Hospital's rule the radial coordinate of point c (Figure 27) is determined from (192) as

$$\rho_c = \sqrt{a_0^2 - \frac{\pi h C_1}{2}} \quad (193)$$

Since $\rho = 1$ at point d , the kinematically admissible velocity field proposed is valid if and only if $\rho_c \leq 1$ or

$$a_0^2 - \frac{\pi h C_1}{2} \leq 1 \quad (194)$$

It is also obvious that $\rho_c > 0$. Therefore, it follows from (193) and (194) that

$$\frac{2(a_0^2 - 1)}{\pi h} \leq C_1 < \frac{2a_0^2}{\pi h} \quad (195)$$

The amount of velocity jump across the velocity discontinuity line bc is determined as

$$[[u_\tau]]_{bc} = \sqrt{(U - u_z)^2 + u_r^2} \quad (196)$$

Substituting (179), (180), (182), (184), and (191) into (196) gives

$$[u_r]_{bc} = U \sqrt{(1 - \sin \mu)^2 + \frac{1}{\rho_{bc}^2(\mu)} \left(C_1 \cos \mu + \frac{a_0^2}{2h} - \frac{\pi C_1}{4} - \frac{\rho_{bc}^2(\mu)}{2h} \right)^2} \quad (197)$$

The infinitesimal length element of the velocity discontinuity line bc is

$$dl_{bc} = \sqrt{(dr)^2 + (dz)^2} \quad (198)$$

Substituting (184), (188) and (191) into (198) results in

$$dl_{bc} = \frac{Rh \cos \mu}{(1 - \sin \mu)} \sqrt{(1 - \sin \mu)^2 + \frac{1}{\rho_{bc}^2(\mu)} \left(C_1 \cos \mu + \frac{a_0^2}{2h} - \frac{\pi C_1}{4} - \frac{\rho_{bc}^2(\mu)}{2h} \right)^2} d\mu \quad (199)$$

It follows from (197) and (199) that the energy dissipation rate at the velocity discontinuity line bc is, after integration with respect to θ ,

$$\begin{aligned} \frac{\sigma_0}{\sqrt{3}} \iint_{S_d} [u_r]_{bc} r d\theta dl_{bc} &= \frac{2\pi\sigma_0 UR^2 h}{\sqrt{3}} \Omega_{bc} \\ \Omega_{bc} &= \int_0^{\pi/2} \frac{\cos \mu}{(1 - \sin \mu)} \left[(1 - \sin \mu)^2 + \frac{1}{\rho_{bc}^2(\mu)} \left(C_1 \cos \mu + \frac{a_0^2}{2h} - \frac{\pi C_1}{4} - \frac{\rho_{bc}^2(\mu)}{2h} \right)^2 \right] \rho_{bc}(\mu) d\mu \end{aligned} \quad (200)$$

The amount of velocity jump across the velocity discontinuity line cd (Figure 27) is equal to $|u_r|$ at $\mu = \pi/2$. Therefore, the energy dissipation rate at this line is, with the use of (180), (182), (184), and (191) and after integration with respect to θ ,

$$\frac{\sigma_0}{\sqrt{3}} \iint_{S_d} [u_r]_{cd} r d\theta dr = \frac{\pi\sigma_0 UR^2}{\sqrt{3}} \int_{\rho_c}^1 \left| \frac{1}{2} \left(\frac{2a_0^2}{h} - C_1\pi \right) - \frac{\rho^2}{h} \right| d\rho \quad (201)$$

Assuming that

$$\frac{2\rho^2}{h} - \frac{2a_0^2}{h} + C_1\pi \geq 0 \quad (202)$$

integration in (201) can be carried out analytically to give

$$\frac{\sigma_0}{\sqrt{3}} \iint_{S_d} \llbracket u_\tau \rrbracket_{cd} r d\theta dr = \frac{\pi \sigma_0 U R^2}{\sqrt{3}} \Omega_{cd},$$

$$\Omega_{cd} = \left[\frac{(1 - \rho_c^3)}{3h} - \frac{1}{2} \left(\frac{2a_0^2}{h} - C_1 \pi \right) (1 - \rho_c) \right] \quad (203)$$

Since ρ varies in the range $1 \geq \rho \geq \rho_c$ and the left hand side of (202) is an increasing function of ρ , it is only necessary to verify the inequality (202) at $\rho = \rho_c$. Then, it follows from (193) that (202) is always satisfied.

The energy dissipation rate in the plastic volume is, with the use of (183), (184), and (191) and after integration with respect to θ ,

$$\sigma_0 \iiint_V \zeta_{eq} r dr dz d\theta = 2\pi \sigma_0 U R^2 \Omega_v,$$

$$\Omega_v = \int_0^{\pi/2} \int_{\rho_{bc}(\mu)}^1 \sqrt{\rho^2 + \frac{(2a_0^2/h - C_1 \pi + 4C_1 \cos \mu)^2 h^2}{12\rho^2} + \frac{C_1^2 \tan^2 \mu}{3} \cos \mu d\rho d\mu} \quad (204)$$

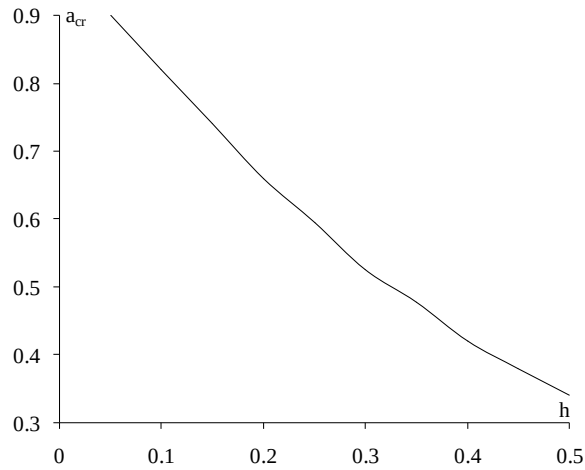
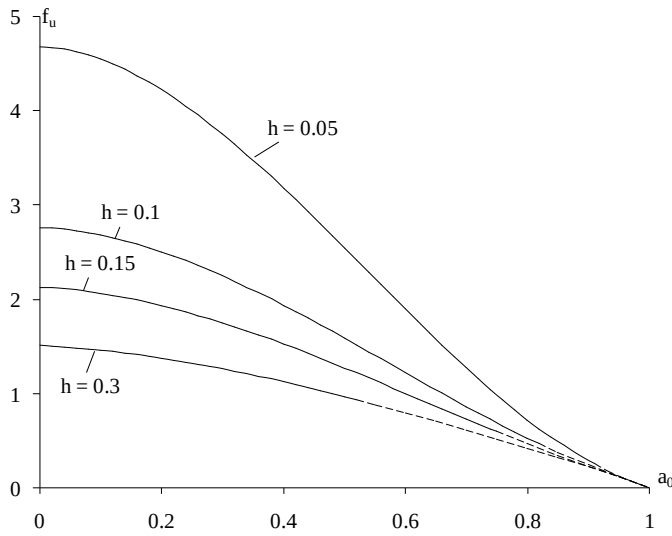
Substituting (200), (203) and (204) into (9) gives

$$\frac{F_u}{\pi R^2 \sigma_0} = f_u = 2\Omega_v + \frac{\Omega_{cd}}{\sqrt{3}} + \frac{2h}{\sqrt{3}} \Omega_{bc} \quad (205)$$

where Ω_v , Ω_{cd} and Ω_{bc} should be found by numerical integration and are functions of C_1 . In order to find the best upper bound based on the kinematically admissible velocity field chosen, it is necessary to minimize the right hand side of (205) with respect to C_1 . Having the value of C_1 the inequality (194) can be solved to determine the critical value of a_0 such that the range of validity of the solution is $a_0 \leq a_{cr}$. It is obvious that a_{cr} depends on h . This dependence is depicted in Figure 29. The variation of the dimensionless upper bound limit load with a_0 in the range $0 \leq a_0 \leq a_{cr}$ is shown for several h -values in Figure 30 (solid lines).

In the range $a_{cr} \leq a_0 \leq 1$, the velocity discontinuity line bc (Figure 27) intersects the stress free surface (Figure 31). Let μ_c be the value of μ at point c . Since $\rho = 1$ at this point, it follows from (192) that

$$C_1 = -\frac{2(1-a_0^2)(1-\sin \mu_c)}{h(2\mu_c + \sin 2\mu_c - \pi \sin \mu_c)} \quad (206)$$

Figure 29. Variation of a_{cr} with h Figure 30. Variation of dimensionless upper bound limit load with crack size for several h -values

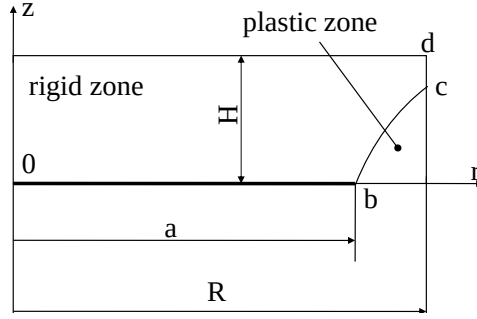


Figure 31. General structure of the kinematically admissible velocity field

In order to find $\bar{f}_u$ in this case, it is possible to use (205) where it is necessary to put $\Omega_{cd} = 0$ and to calculate Ω_V and Ω_{bc} according to

$$\Omega_V = \int_0^{\mu_c} \int_{\rho_{bc}(\mu)}^1 \sqrt{\rho^2 + \frac{(2a_0^2/h - C_1\pi + 4C_1\cos\mu)^2 h^2}{12\rho^2} + \frac{C_1^2 \tan^2 \mu}{3} \cos\mu d\rho d\mu}, \quad (207)$$

$$\Omega_{bc} = \int_0^{\mu_c} \frac{\cos\mu}{(1-\sin\mu)} \left[(1-\sin\mu)^2 + \frac{1}{\rho_{bc}^2(\mu)} \left(C_1 \cos\mu + \frac{a_0^2}{2h} - \frac{\pi C_1}{4} - \frac{\rho_{bc}^2(\mu)}{2h} \right)^2 \right] \rho_{bc}(\mu) d\mu$$

where C_1 should be excluded by means of (206). Then, the right hand side of (205) should be minimized with respect to μ_c to find the best upper bound based on the kinematically admissible velocity field chosen. The variation of the dimensionless upper bound limit load with a_0 in the range $a_{cr} \leq a_0 \leq 1$ is shown for several h -values in Figure 30 (broken lines).

4.2. Round Bar with an Axisymmetric Crack at Some Distance from the Mid-Plane of the Weld

Geometry of the specimen, the system of loading, the direction of velocity of the rigid blocks of base material U and the cylindrical coordinate system are shown in Figure 32 where $2H$ is the thickness of the weld, R is the radius of the specimen, and a is the radius of the crack. The crack is located at some distance ε from the mid-plane of the weld. The value of ε varies in the range $0 \leq \varepsilon \leq H$. The configuration considered in the previous section is obtained at $\varepsilon = 0$. This solution can be adopted to find an upper limit load for the structure under consideration. The general structure of the chosen kinematically admissible velocity field within the weld is shown in Figure 33 where the kinematically admissible velocity field in the plastic zone is taken in the form of (179) and (180). Let $\bar{f}_u^{(0)}$ be the upper bound limit

load for the specimen with $\varepsilon = 0$. The further analysis in this section is restricted to the class of specimens for which $f_u^{(0)}$ is given by (205). The energy dissipation rate in the plastic zone, including the energy dissipation rate at the velocity discontinuity surfaces bc and cd , is

$$\dot{E}_0 = \pi UR^2 \sigma_0 f_u^{(0)} \quad (208)$$

The amount of velocity jump across the velocity discontinuity line bt (Figure 33) is $2U$. Therefore, the energy dissipation rate at this line is

$$\dot{E}_1 = \frac{4\pi}{\sqrt{3}} U \sigma_0 a \varepsilon \quad (209)$$

Substituting (208) and (209) into (9) gives

$$2F_u U = 2\dot{E}_0 + \dot{E}_1 = 2\pi UR^2 \sigma_0 f_u^{(0)} + \frac{4\pi}{\sqrt{3}} U \sigma_0 a \varepsilon \quad (210)$$

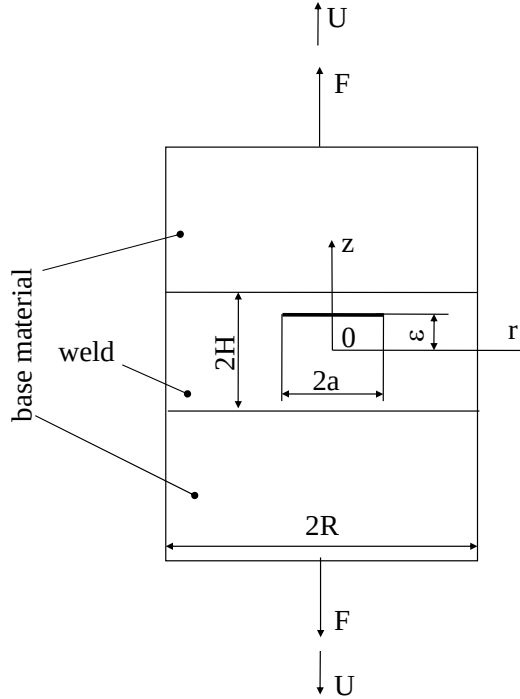


Figure 32. Geometry of structure under consideration – notation

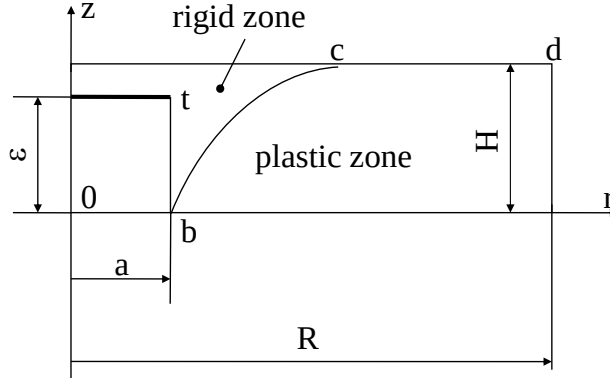


Figure 33. General structure of the kinematically admissible velocity field

It has been taken into account here that two identical forces are applied to the specimen (Figure 32) and $\dot{E}_0$ in (208) is for one half of the plastic zone. Equation (210) can be rewritten in the following dimensionless form

$$\frac{F_u}{\pi R^2 \sigma_0} = f_u = f_u^{(0)} + \frac{2a_0}{\sqrt{3}} \frac{\varepsilon}{R} \quad (211)$$

Since $f_u^{(0)}$ has been already found (Figure 30), the upper bound limit load for the structure under consideration can be immediately found from (211).

The last term on the right hand side of (211) can make a too large contribution for structures with small cracks. In order to precisely determine the range of applicability of the solution (211), it is necessary to compare f_u from (211) and f_u from (205) at $a = 0$. The smallest value should be adopted as the limit load.

5. THREE-DIMENSIONAL SOLUTION

Geometry of the M(T) specimen and the system of loading are shown in Figure 34. This specimen is a special case of the specimen shown in Figure 2 and, therefore, equation (12) is applicable. The solution given in this section has been proposed in Alexandrov and Kocak (2008). The thickness of the specimen is constant and equal to $2B$, and the thickness of the weld is $2H$. The specimen has three planes of symmetry, and the axes of the Cartesian coordinate system can be chosen along three straight lines of intersection of these planes. Therefore, it is sufficient to get the solution in the domain $x \geq 0$, $y \geq 0$ and $z \geq 0$. It is possible to put $L \equiv H$ in the definition for w and W_1 used in equations (11) and (12).

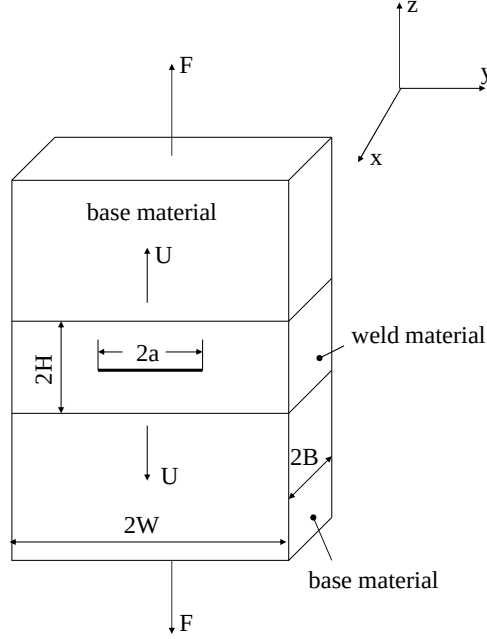


Figure 34. Geometry of structure under consideration – notation

To apply the approach developed in Alexandrov and Kocak (2008) and briefly discussed in Section 2, the specimen with no crack is considered first. Assume a kinematically admissible velocity field of the form

$$\frac{u_z}{U} = \eta, \quad \frac{u_x}{U} = \alpha \nu + k(\eta), \quad \frac{u_y}{U} = -(1 + \alpha)\mu + g(\eta) \quad (212)$$

where $\eta = z/H$, $\nu = x/H$, and $\mu = y/H$ are the dimensionless coordinates, $k(\eta)$ and $g(\eta)$ are arbitrary functions of η , and α is an arbitrary constant. The velocity field (212) satisfies the equation of incompressibility at any $k(\eta)$, $g(\eta)$ and α . It is natural to assume that $z = H$ is a velocity discontinuity surface. Therefore, in the vicinity of this surface the actual velocity field is singular and the distribution of the shear strain rates ζ_{xz} and ζ_{yz} should lead to (10). It is possible if

$$\zeta_{xz} = O(1/\sqrt{1-\eta}), \quad \zeta_{yz} = O(1/\sqrt{1-\eta}) \quad \text{as } \eta \rightarrow 1 \quad (213)$$

In addition, because of symmetry the actual velocity field satisfies the conditions

$$\zeta_{xz} = 0 \quad \text{and} \quad \zeta_{yz} = 0 \quad \text{at } \eta = 0 \quad (214)$$

Possible and one of most simple functions $k(\eta)$ and $g(\eta)$ satisfying equations (213) and (214) are $k(\eta) = k_0 + k_1\sqrt{1-\eta^2}$ and $g(\eta) = g_0 + g_1\sqrt{1-\eta^2}$ where k_0 , k_1 , g_0 and g_1 are arbitrary constants. Then, the velocity field (212) takes the form

$$\frac{u_z}{U} = \eta, \quad \frac{u_x}{U} = \alpha\nu + k_0 + k_1\sqrt{1-\eta^2}, \quad \frac{u_y}{U} = -(1+\alpha)\mu + g_0 + g_1\sqrt{1-\eta^2} \quad (215)$$

To satisfy the natural velocity boundary conditions $u_x = 0$ at $x = 0$ and $u_y = 0$ at $y = 0$, it is necessary to introduce a rigid zone in the vicinity of the planes of symmetry $x = 0$ and $y = 0$. This zone moves up with velocity U . The boundary of the rigid zone and the plastic zone is a piece-wise smooth surface consisting of two smooth parts. The structure of the velocity field (215) and the position of the axes of symmetry $x = 0$ and $y = 0$ require that the unit normal vectors to these smooth parts are represented by the following equations

$$\mathbf{n} = -\sin \varphi \mathbf{j} + \cos \varphi \mathbf{k} \quad \text{and} \quad \mathbf{m} = -\sin \phi \mathbf{i} + \cos \phi \mathbf{k} \quad (216)$$

where $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are the unit vectors parallel to the axes x , y , and z , respectively, φ is the orientation of the line (in planes $x = \text{constant}$) tangent to the velocity discontinuity surface (curve in planes $x = \text{constant}$) relative to the axis y , and ϕ is the orientation of the line (in planes $y = \text{constant}$) tangent to the velocity discontinuity surface (curve in planes $y = \text{constant}$) relative to the axis x . The cross-section of the velocity discontinuity surface corresponding to the unit vector $\mathbf{n}$ by a plane $x = \text{constant}$ and angle φ are shown in Figure 35. It follows from this figure that

$$\tan \varphi = \frac{dz}{dy} = \frac{d\eta}{d\mu} \quad (217)$$

The velocity vector in the rigid zone can be written as

$$\mathbf{u}_r = U\mathbf{k} \quad (218)$$

The velocity vector in the plastic zone is represented by

$$\mathbf{u}_{pl} = u_x \mathbf{i} + u_y \mathbf{j} + u_z \mathbf{k} \quad (219)$$

where u_x , u_y and u_z are given by equation (215). The normal velocity must be continuous across the velocity discontinuity surface. Therefore, $\mathbf{u}_r \cdot \mathbf{n} = \mathbf{u}_{pl} \cdot \mathbf{n}$. Using equations (215) - (219) this equation can be transformed to

$$\frac{d\mu}{d\psi} = \frac{[2(1+\alpha)\mu - g_0 - g_1 \cos 2\psi] \cos 2\psi}{1 - \sin 2\psi} \quad (220)$$

where

$$\sin 2\psi = \eta \quad \text{and} \quad d\eta = 2 \cos 2\psi d\psi \quad (221)$$

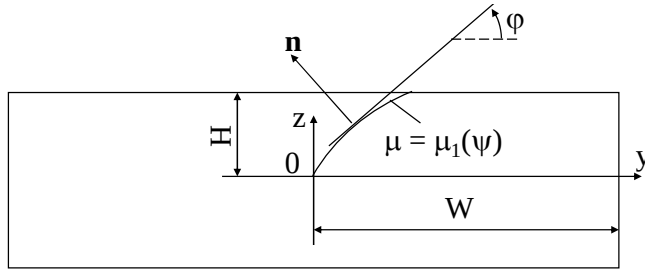


Figure 35. Shape of velocity discontinuity surface at $x = \text{constant}$

Equation (220) determines the same curve in each yz plane. This curve must contain the point $z = 0$ and $y = 0$. Therefore, the boundary condition to equation (220) is

$$\mu = 0 \quad \text{at} \quad \psi = 0 \quad (222)$$

A natural additional requirement is that the curve has a common point with the line $\eta = 1$ (or $\psi = \pi/4$). Since the denominator of the right hand side of equation (220) is zero at $\psi = \pi/4$, a necessary condition is that its nominator is also zero at $\psi = \pi/4$. The latter condition is satisfied at the point

$$\mu = \mu_0 = \frac{g_0}{1+\alpha} \quad (223)$$

Expanding the nominator and denominator of the right hand side of equation (220) in series in the vicinity of $\psi = \pi/4$ and $\mu = \mu_0$ gives

$$\frac{d(\mu - \mu_0)}{d(\psi - \pi/4)} = -2(1+\alpha) \frac{(\mu - \mu_0)}{(\psi - \pi/4)} - 4g_1$$

At $-1-\alpha \neq 1/2$ the solution to this equation is

$$\mu = \mu_0 - \frac{4g_1}{(3+2\alpha)} \left(\psi - \frac{\pi}{4} \right) + C_1 \left(\frac{\pi}{4} - \psi \right)^{-2(1+\alpha)} \quad (224)$$

where C_1 is a constant of integration. It is expected to assume that the normal strain rates ζ_{xx} and ζ_{yy} are compressible, $\zeta_{xx} < 0$ and $\zeta_{yy} < 0$. Then, it follows from the velocity field (215) that $0 > \alpha > -1$. In this case the condition $\mu = \mu_0$ at $\psi = \pi/4$ is satisfied if and only if $C_1 = 0$. Finally, the solution to equation (220) in the vicinity of $\psi = \pi/4$ is determined from equations (223) and (224) in the form

$$\mu_1(\psi) = \frac{g_0}{1+\alpha} - \frac{4g_1}{(3+2\alpha)} \left(\psi - \frac{\pi}{4} \right) \quad (225)$$

It is assumed that equation (225) is valid in the range

$$\frac{\pi}{4} \geq \psi \geq \psi_1 = \frac{\pi}{4} - \delta, \quad \delta \ll 1 \quad (226)$$

Equation (220) is a linear differential equation with respect to μ . Therefore, its solution satisfying the condition (222) can be written in the form

$$\begin{aligned} \mu_1(\psi) &= -\frac{g_0}{(1+\alpha)} \left[1 - (1 - \sin 2\psi)^{-(1+\alpha)} \right] - 2g_1 \Phi_1(\psi) (1 - \sin 2\psi)^{-(1+\alpha)} \\ \Phi_1(\psi) &= \int_0^\psi \cos^2 2\gamma (1 - \sin 2\gamma)^\alpha d\gamma \end{aligned} \quad (227)$$

Equating $\mu_1(\psi)$ found from equations (225) and (227) at $\psi = \psi_1$ introduced in equation (226) leads to

$$g_1 = -\frac{g_0}{(1+\alpha)} \frac{(3+2\alpha)}{\left[4\delta (1 - \cos 2\delta)^{(1+\alpha)} + 2(3+2\alpha) \Phi_1(\psi_1) \right]} \quad (228)$$

Equations (225) and (227) determine the shape of the velocity discontinuity surface shown in Figure 35.

The amount of velocity jump across the velocity discontinuity surface is determined by $[u]_1 = |\mathbf{u}_r - \mathbf{u}_{pl}|$ where the velocity vectors should be taken at the surface. Substituting equations (215), (218) and (219) into this equation and using equation (221) give

$$[u]_1 = U \sqrt{(1 - \sin 2\psi)^2 + (\alpha\nu + k_0 + k_1 \cos 2\psi)^2 + [(1 + \alpha)\mu_1(\psi) - g_0 - g_1 \cos 2\psi]^2} \quad (229)$$

The infinitesimal area element of the velocity discontinuity surface can be found in the form

$$ds = dx \sqrt{(dy)^2 + (dz)^2} = H^2 d\nu \sqrt{(d\mu)^2 + (d\eta)^2} = H^2 d\nu \sqrt{(d\mu)^2 + 4 \cos^2 2\psi (d\psi)^2} \quad (230)$$

Substituting the derivative $d\mu/d\psi$ from equation (220) into equation (230) leads to

$$ds = \frac{2H^2\nu \cos 2\psi}{(1 - \sin 2\psi)} \sqrt{(1 - \sin 2\psi)^2 + [(1 + \alpha)\mu_1(\psi) - g_0 - g_1 \cos 2\psi]^2} d\psi d\nu \quad (231)$$

The energy dissipation rate at the velocity discontinuity surface $\mu = \mu_1(\psi)$ is finally given by

$$\dot{E}_1 = \frac{2\sigma_0 U H^2}{\sqrt{3}} \int_0^{\pi/4} \int_{\nu_1(\psi)}^{B/H} \frac{\cos 2\psi}{(1 - \sin 2\psi)} \times \sqrt{(1 - \sin 2\psi)^2 + [(1 + \alpha)\mu_1(\psi) - g_0 - g_1 \cos 2\psi]^2} \frac{[u]_1}{U} d\nu d\psi \quad (232)$$

The function $\nu_1(\psi)$ will be determined later.

A similar analysis can be carried out for the velocity discontinuity surface corresponding to the unit normal vector $\mathbf{m}$. As a result, the equation for this surface is

$$\nu_1(\psi) = -\frac{k_0}{\alpha} - \frac{4k_1}{(1 - 2\alpha)} \left(\psi - \frac{\pi}{4} \right) \quad (233)$$

in the interval $\pi/4 \geq \psi \geq \psi_1$ and

$$\nu_1(\psi) = -\frac{k_0}{\alpha} + (1 + \sin 2\psi)^\alpha \left[\frac{k_0}{\alpha} - 2k_1 \Phi_2(\psi) \right], \Phi_2(\psi) = \int_0^\psi \frac{\cos^2 \gamma}{(1 - \sin 2\gamma)^{(1+\alpha)}} d\gamma \quad (234)$$

in the interval $\psi_1 \geq \psi \geq 0$. Constant k_1 should be excluded and expressed as

$$k_1 = \frac{k_0(1-2\alpha)(1-\cos 2\delta)^\alpha}{\alpha \left[4\delta + 2(1-2\alpha)(1-\cos 2\delta)^\alpha \Phi_2(\psi_1) \right]} \quad (235)$$

Thus, the function $v_1(\psi)$ involved in equation (232) is determined by equations (233) and (234). In particular, $v_1(\pi/4) = v_0 = -k_0/\alpha$. The energy dissipation rate at the velocity discontinuity surface $v = v_1(\psi)$ is given by

$$\dot{E}_2 = \frac{2\sigma_0 U H^2}{\sqrt{3}} \int_0^{\pi/4} \int_{\mu_1(\psi)}^{W/H} \frac{\cos 2\psi}{(1-\sin 2\psi)} \times \frac{1}{\sqrt{(1-\sin 2\psi)^2 + [\alpha v_1(\psi) + k_0 + k_1 \cos 2\psi]^2}} \frac{[u]_2}{U} d\mu d\psi \quad (236)$$

where

$$[u]_2 = U \sqrt{(1-\sin 2\psi)^2 + [\alpha v_1(\psi) + k_0 + k_1 \cos 2\psi]^2 + [(1+\alpha)\mu - g_0 - g_1 \cos 2\psi]^2} \quad (237)$$

The geometry of the rigid and plastic zones at $\eta = 1$ is illustrated in Figure 36.

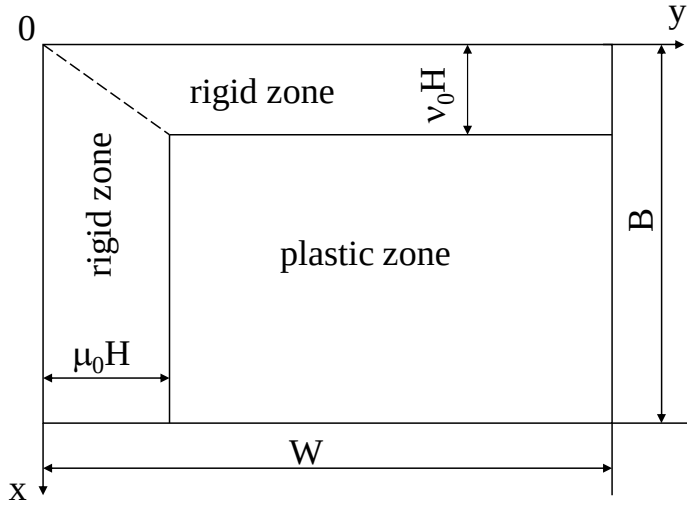


Figure 36. Configuration of plastic and rigid zones at $z = H$

Another velocity discontinuity surface appears at $\eta = 1$ (or $\psi = \pi/4$) in the region $\mu_1(\pi/4) \leq \mu \leq W/H$ and $v_1(\pi/4) \leq v \leq B/H$ or, with the use of equations (225) and (233), $g_0/(1+\alpha) \leq \mu \leq W/H$ and $-k_0/\alpha \leq v \leq B/H$. The amount of velocity jump

across this velocity discontinuity surface is equal to $[u]_3 = \sqrt{u_x^2 + u_y^2}$ where u_x and u_y should be taken at $\eta = 1$. Then, it follows from equation (215) that

$$[u]_3 = U \sqrt{(k_0 + \alpha v)^2 + [g_0 - (1 + \alpha)\mu]^2} \quad (238)$$

The infinitesimal area element is just $ds = H^2 d\mu dv$. Therefore, with the use of equation (238), the energy dissipation rate at this velocity discontinuity surface is given by

$$\dot{E}_3 = \frac{\sigma_0 U H^2}{\sqrt{3}} \int_{-k_0/\alpha}^{B/H} \int_{g_0/(1+\alpha)}^{W/H} \sqrt{(k_0 + \alpha v)^2 + [g_0 - (1 + \alpha)\mu]^2} d\mu dv \quad (239)$$

Using the velocity field (215) and equation (221) the equivalent strain rate and the infinitesimal volume element are determined by

$$\begin{aligned} \zeta_{eq} &= \frac{U}{\sqrt{3}H} \sqrt{4(1 + \alpha + \alpha^2) + (k_1^2 + g_1^2) \tan^2 2\psi}, \\ dV &= H^3 dv d\mu d\eta = 2H^3 \cos 2\psi dv d\mu d\psi \end{aligned}$$

Therefore, the energy dissipation rate in the plastic zone is

$$\dot{E}_{pl} = \frac{2U\sigma_0 H^2}{\sqrt{3}} \int_0^{\pi/4} \int_{\mu_1(\psi)}^{\pi/4} \int_{v_1(\psi)}^{B/H} \sqrt{4(1 + \alpha + \alpha^2) + (k_1^2 + g_1^2) \tan^2 2\psi} \cos 2\psi dv d\mu d\psi$$

Integration with respect to v and μ can be carried out with no difficulty to give

$$\dot{E}_{pl} = \frac{2U\sigma_0 H^2}{\sqrt{3}} \int_0^{\pi/4} \left(\frac{B}{H} - v_1(\psi) \right) \left(\frac{W}{H} - \mu_1(\psi) \right) \times \sqrt{4(1 + \alpha + \alpha^2) + (k_1^2 + g_1^2) \tan^2 2\psi} \cos 2\psi d\psi \quad (240)$$

In the case under consideration,

$$\sigma_0 \iiint_V \zeta_{eq} dV = \dot{E}_{pl}, \quad \frac{\sigma_0}{\sqrt{3}} \iint_{S_d} [u_\tau] dS = \dot{E}_1 + \dot{E}_2 + \dot{E}_3$$

Therefore, equation (9) becomes

$$F_{u1}^{(0)} U = 4(\dot{E}_1 + \dot{E}_2 + \dot{E}_3 + \dot{E}_{pl}) \quad (241)$$

where $F_{u1}^{(0)}$ is the upper bound on the load applied for the specimen with no crack and based on the velocity field (215). It is convenient to rewrite equation (241) in the following dimensionless form

$$f_{u1}^{(0)} = \frac{F_{u1}^{(0)}}{4\sigma_0 WB} = \frac{E_1 + E_2 + E_3 + E_4}{U\sigma_0 WB} \quad (242)$$

The right hand side of this equation can be calculated by means of equations (225), (227), (229), (232), (233), (234), (236), (237), (239), and (240), and, with the use of equations (228) and (235), can be represented as a function of three parameters, α , k_0 and g_0 . The right hand side of equation (242) should be minimized with respect to these three parameters to find the best upper bound based on the kinematically admissible velocity field chosen.

The right hand side of equation (242) should be minimized numerically. The minimization has been performed in the domain $2 \leq W/H \leq 10$ and $2 \leq B/H \leq 10$. It has been assumed that $\delta = 0.001$ in equation (226). Note that W and B are involved in equation (242) in a symmetric manner. Therefore, $f_{u1}^{(0)}$ is an even function of $W - B$. Using this property the numerical solution can be fitted to a polynomial as

$$\begin{aligned} f_{u1}^{(0)} = \Omega_1 = & 1.122 - 0.0264 \frac{(W-B)^2}{H^2} + 0.1035 \frac{(W+B)}{H} + \\ & + 0.00095 \frac{(W-B)^2(W+B)}{H^3} + 0.00025 \frac{(W+B)^2}{H^2} \end{aligned} \quad (243)$$

For the specimen with a sufficiently small ratio W/H , which is equivalent to a sufficiently large crack for the specimen with the crack, another solution used in many previous studies can be proposed by assuming the kinematically admissible velocity field consisting of two rigid blocks (in the domain $y \geq 0$ and $z \geq 0$) separated by a velocity discontinuity surface (straight line in yz planes), similar to that used in Section 3.1.5. After some simple algebra (see Alexandrov and Kocak, 2008), the upper bound on the limit load based on this special velocity field can be found as

$$f_{u2}^{(0)} = \Omega_2 = \frac{F_{u2}^{(0)}}{4\sigma_0 WB} = \begin{cases} \frac{2}{\sqrt{3}}, & \text{if } \beta \geq \frac{\pi}{4} \\ \frac{2}{\sqrt{3} \sin 2\beta}, & \text{if } \beta < \frac{\pi}{4} \end{cases} \quad (244)$$

where $\beta = \arctan(H/W)$. Since equations (16) at $a = 0$ and (244) are based on kinematically admissible velocity fields applicable for three dimensional deformation, it

follows from the upper bound theorem that the upper bound limit load for the specimen with no crack is

$$f_u^{(0)} = \Omega = \min \{ \Omega_1, \Omega_2, \Omega_3 \} \quad (245)$$

where Ω_3 is f_u from (16) at $a = 0$.

In order to find the limit load for the specimen with a crack, it is just necessary to combine the solution found and (12).

The variation of Ω found from (245) with w_1 at different values of B/H is shown in Figure 37. In the case of specimens with no crack, $w = w_1$ and $\Omega(w) = \Omega(w_1) = f_u^{(0)}(w)$, giving an upper bound of the dimensionless force. It can be seen from Figure 37 that the function $f_u^{(0)}(w)$ can attain a maximum at some value of w . However, $F_u^{(0)}(w)$ is a monotonic function of w . The single curve (including its extension shown by a dashed line) independent of B/H corresponds to the plane-strain solution whereas five different curves corresponding to five different values of B/H demonstrate an effect of three dimensional deformation. Note that the actual effect of three dimensional deformation is even larger than that shown in Figure 37. For, the solution (16) is an approximation of a numerical solution satisfying all field equations and boundary conditions, whereas equation (243) is based on the minimization of a function of three variables according to the upper bound theorem. The exact three dimensional solution would result in a smaller limit load and would therefore shift the curves corresponding to three dimensional deformation (Figure 37) down.

6. CONCLUSION

The present chapter concerns with a number of upper bound limit load solutions for highly undermatched welded joints. Special attention is devoted to non-standard approaches to constructing kinematically admissible velocity fields. Most of solutions are given in an analytic form. Therefore, the solutions can be directly used in engineering applications such as flaw assessment procedures.

The singular behavior of the actual velocity field given by equation (10) is of special importance for highly undermatched welded joints because the interface between the base and weld materials is usually a velocity discontinuity surface, and the equivalent strain rate follows the rule (10) in the vicinity of such surfaces.

In some cases, the limit load for a given structure can be found with the use of the limit load solution for a simpler structure and an additional term (or multiplier) which can be easily found. This approach is also applicable for overmatched structures.

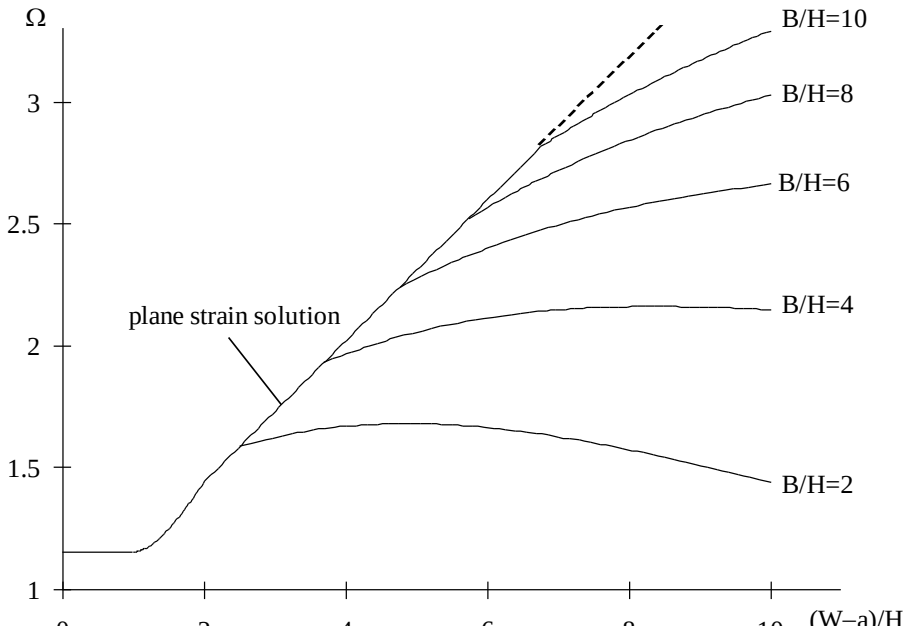


Figure 37. Variation of Ω with geometric parameters of the specimen

ACKNOWLEDGMENT

The research described has been supported by grant RFBR-08-01-00700.

REFERENCES

- Aleksandrov, SE; Konchakova, NA. *J Machinery Manufacture Reliability*, 2007, 36, 50-56.
 Alexandrov, S. *J Mater Proc Technol*, 2000, 105, 278-283.
 Alexandrov, S. *J Appl Mech Techn Phys.*, 2008, 49, 340-345.
 Alexandrov, S. *Mater Sci Forum.*, 2010, 638-642, 3821-3826.
 Alexandrov, S; Chung, KH; Chung, K. *Fat Fract Engng Mater Struct*, 2007, 30, 333-341.
 Alexandrov, S; Chicanova, N; Kocak, M. *Engng Fract. Mech*, 1999^a, 64, 383-399.
 Alexandrov, SE; Goldstein, RV; Tchikanova, NN. *Fat Fract Engng Mater Struct*, 1999^b, 22, 775-780.
 Alexandrov, SE; Goldstein, RV. *Fat Fract Engng Mater Struct*, 1999, 22, 975-979.
 Alexandrov, S; Goldstein, R. *Mech Solids*, 2005, 40, 36-41.
 Alexandrov, S; Gracio, J. *Fat Fract Engng Mater Struct*, 2003, 26, 399-403.
 Alexandrov, S; Kocak, M. *Fat Fract Engng Mater Struct*, 2007, 30, 351-355.
 Alexandrov, S; Kocak, M. *Proc IMechE Part C J Mech Engng Sci.*, 2008, 222, 107-115.
 Alexandrov, S; Kontchakova, N. *Mater Sci Engng.*, 2004, A387-389, 395-398.
 Alexandrov, S; Kontchakova, N. *Engng Fract Mech.*, 2005, 72, 151-157.
 Alexandrov, S; Richmond, O. *Int J Non-Linear Mech*, 2001, 36, 1-11.

- Alexandrov, S; Tzou, GY. *Key Engng Mater*, 2007, 345-346, 425-428.
- Alexandrov, S; Tzou, GY; Hsia, SY. *Engng Fract Mech.*, 2008, 75, 3131-3140.
- Avitzur, B. *Metal Forming: the Application of Limit Analysis*, Dekker: New York, NY, 1980.
- Bramley, AN. *J Mater Process Technol.*, 2001, 116, 62-66.
- Capsoni, A; Corradi, L; Vena, P. *Int J Solids Struct*, 2001^a, 38, 3945-3963.
- Capsoni, A; Corradi, L; Vena, P. *Int J Plast*, 2001^b, 17, 1531-1549.
- Drucker, DC; Prager, W; Greenberg, HJ. *Quart Appl Math*, 1952, 9, 381-389.
- Hao, S; Cornec, A; Schwalbe, KH. *Int J Solids Struct*, 1997, 34, 297-326.
- Hill, R. *The Mathematical Theory of Plasticity*, Clarendon Press: Oxford, 1950.
- Hill, R. *J Mech Phys Solids*, 1956, 5, 66-74.
- Joch, J; Ainsworth, RA; Hyde, TH. *Fat Fract Engng Mater Struct*, 1993, 16, 1061-1079.
- Kachanov, LM. *Foundations of Plasticity Theory*, GITTL: Moscow, 1956 [in Russian].
- Kim, YJ; Schwalbe, KH. *Engng Fract Mech*, 2001^a, 68, 163-182.
- Kim, YJ; Schwalbe, KH. *Engng Fract Mech*, 2001^b, 68, 183-199.
- Kim, YJ; Schwalbe, KH. *Engng Fract Mech*, 2001^c, 68, 1137-1151.
- Kotousov, A; Jaffar, MFM. *Engng Failure Anal*, 2006, 13, 1065-1075.
- Miller, AG. *Int J Press Ves Pip*, 1988, 32, 197-327.
- Tzou, GY; Alexandrov, S. *J Mater Process Technol.*, 2006, 177, 159-162.
- Zerbst, U; Ainsworth, RA; Schwalbe, KH. *Int J Press Ves Pip.*, 2000, 77, 855-867.

Chapter 6

FRACTURE AND FATIGUE ASSESSMENT OF WELDED STRUCTURES

S. Cicero* and F. Gutiérrez-Solana

University of Cantabria, Materials Science and Engineering Department,
Santander, Cantabria, Spain

ABSTRACT

The presence of damage in engineering structures and components may have different origins and mechanisms, basically depending on the type of component, loading and environmental conditions and material performance. Four major modes or processes have generally been identified as the most frequent causes of failure in engineering structures and components: fracture, fatigue, creep and corrosion (including environmental assisted cracking), together with the interactions between all of these. As a consequence, different Fitness-for-Service (FFS) methodologies have been developed with the aim of covering the mentioned failure modes, giving rise to a whole engineering discipline known as structural integrity.

At the same time, welds can be considered as singular structural details, as they may have, among others features, noticeably different mechanical properties from the base material (both tensile properties and toughness), geometrical singularities causing stress concentrations, and residual stresses with specific profiles depending on the type of weld and welding process. Traditional approaches to the assessment of welds have consisted in making successive conservative assumptions that lead to over-conservative results. This has led to the development, from a more precise knowledge of weld behavior and performance, of specific Fitness-for-Service (FFS) assessment procedures for welds which offer great improvements with respect to traditional approaches and lead to more accurate (and still safe) results or predictions.

The main aim of this chapter is to present these advanced Fitness-for-Service (FFS) tools for the assessment of welds and welded structures in relation to two of the above-mentioned main failure modes: fracture and fatigue.

* Corresponding author: Email: ciceros@unican.es

1. INTRODUCTION

The presence of damage in engineering structures and components may have different origins and mechanisms, basically depending on the type of component, loading and environmental conditions and material performance. Four major modes or processes have generally been identified as the most frequent causes of failure in engineering structures and components (together with the interactions between all of these):

- Fracture: the failure occurs when the applied driving force acting to extend a crack (the crack driving force) exceeds the material's ability to resist the extension of that crack. This material property is called the material's fracture toughness or fracture resistance [1]. The final fracture of structural components is associated with the presence of macro or microstructural defects that affect the stress state due to the loading conditions.
- Fatigue: type of failure that involves initiation and propagation of cracks in components subjected to cyclic loading that, in general, do not exceed the yield stress of the material. In case there is a pre-existing flaw, it basically consists in crack growth in the presence of cyclic stresses; if there is no pre-existing flaw, fatigue involves a crack initiation process plus the crack growth.
- Creep: components and structures that operate at high temperatures (relative to the melting point of the material) may fail through slow, stable extension of a macroscopic crack.
- Corrosion (including environmental assisted cracking): due to electrochemical processes causing the degradation of the material, metal loss, appearance of defects (e.g., pits) and/or flaw propagation.

As a consequence, different Fitness-for-Service (FFS) methodologies have been developed with the aim of covering the mentioned failure modes, giving rise to a whole engineering discipline known as structural integrity. These methodologies are generally implemented in well known FFS/structural integrity procedures. Some examples are BS7910 [2], SINTAP [3] R5 [4], R6 [5] or API579/ASME FFS [6]. Most of them are focused on one specific failure mode (e.g., R5 analyses creep processes) and/or one industrial sector (e.g., the original field of API579 is the petrol sector, although it can be used in other situations). In order to provide a wider scope of analysis, and as part of the V EU Framework Program, the European Fitness-for-Service Network [7] devised the FITNET FFS Procedure [1], a document which defines a structural integrity assessment procedure for analysis against the four above mentioned main failure modes: fracture-plastic collapse, fatigue, creep and corrosion.

At the same time, welds can be considered as singular structural details, as they may have, among others features, noticeably different mechanical properties from the base material (both tensile properties and toughness), geometrical singularities causing stress concentrations, and residual stresses with specific profiles depending on the type of weld and welding process. Traditional approaches to the assessment of welds have consisted in making successive conservative assumptions that lead to over-conservative results (e.g., use of minimum values of yield strength – base material vs. weld - in the joint, or assuming a

uniform tensile residual stress field equal in magnitude to the maximum yield stress of the base material or weld material). This has led to the development, from a more precise knowledge of weld behavior and performance, of specific Fitness-for-Service (FFS) assessment procedures for welds which offer great improvements with respect to traditional approaches and lead to more accurate (and still safe) results or predictions. Some of the most significant advances have been the development of mismatch analysis procedures (both overmatch and undermatch), the definition of adjusted residual stress profiles for the main types of welds, the definition of stress concentrations in case of weld misalignment and the consideration of peak stresses (instead of nominal stresses) in the fatigue analysis of welds. All these issues have been considered in FITNET FFS, and the review provided in this chapter is mostly based on the contents of this procedure.

Hence, the main aim of this chapter is to present advanced Fitness-for-Service (FFS) tools for fracture and fatigue assessment of welds and welded structures, following the guidelines provided by FITNET FFS Procedure, as one of the most updated Fitness-for-Service assessment procedures. For further knowledge on creep and corrosion, as well as their analysis in welded structures along the same lines as the contents of this chapter, the reader is referred to specialized bibliography and assessment procedures, including the FITNET FFS itself.

2. FRACTURE ASSESSMENT OF WELDED STRUCTURES

2.1. Brief Overview of Ordinary Fracture Assessments (Welded and Non-Welded)

The fracture analysis of the component containing a crack or crack-like flaw is expected to be controlled by the following three parameters:

- (a) The fracture resistance of the material
- (b) The component and crack geometry
- (c) The applied stresses including secondary stresses such as residual stresses.

Usually two of these parameters are known and, therefore, the third can be determined by using the relationships of fracture mechanics [1].

There are two main approaches for determining the integrity of cracked structures and components: the first uses the concept of a Failure Assessment Diagram (FAD) [8,9]; the second a diagram which uses a crack driving force (CDF) curve [8,9]. Both approaches are based on the same scientific principles and give identical results when the input data are treated identically, so they are totally equivalent approaches [10].

The basis of both approaches is that failure is avoided so long as the structure is not loaded beyond its maximum load bearing capacity defined using both fracture mechanics criteria and plastic limit analysis. The fracture mechanics analysis involves comparison of the crack tip driving force with the material's fracture toughness or fracture resistance. The crack tip loading must, in most cases, be evaluated using elastic-plastic concepts and is dependent on the geometry (of both the structure and the crack), the material's tensile properties and the loading. In the FAD approach, both the comparison of the crack tip driving force with the

material's fracture toughness and that of the applied load with the plastic load limit are performed at the same time. In the CDF approach the crack driving force is plotted and compared directly with the material's fracture toughness. Separate analysis is carried out for the plastic limit analysis [1].

The FAD is a plot of the failure envelope of the cracked structure, defined in terms of two parameters, K_r and L_r . The former is defined as the ratio of the applied linear elastic stress intensity factor, K_I , to the material's fracture toughness, K_{mat} ; the latter is defined as the ratio of the total applied load (F) giving rise to the primary stresses, to the plastic limit load (F_e) of the flawed structure [1]. Solutions of K_I and F_e are available in the literature for a wide range of geometries, also depending on the stress distribution in the structural section being analyzed.

The failure envelope is called the Failure Assessment Line (FAL), which is basically dependent on the material's tensile properties, through the equation:

$$K_r = f(L_r) \quad (1)$$

It incorporates a cut-off at $L_r = L_r^{max}$, which defines the plastic collapse limit of the structure. $f(L_r)$, which is actually a plasticity correction function, is provided by assessment procedures and presents different expressions depending on the data available regarding the stress-strain curve of the material.

The component being assessed is represented in the FAD through the co-ordinates (K_r, L_r), calculated under the loading conditions applicable (given by the loads, crack size, material properties), which are then compared with the Failure Assessment Line. Figure 1a [1] shows an example for a structure analyzed using the fracture initiation levels of analysis, and Figure 1b [1] provides an example for a structure that may fail by ductile tearing. Assessment points lying on or within the area defined by the FAL and the coordinate axes indicate that the structure is acceptable against this limiting condition. A point which lies outside this envelope indicates that the structure as assessed has failed to meet this limiting condition. Margins and factors can be determined by comparing the assessed condition with the limiting condition.

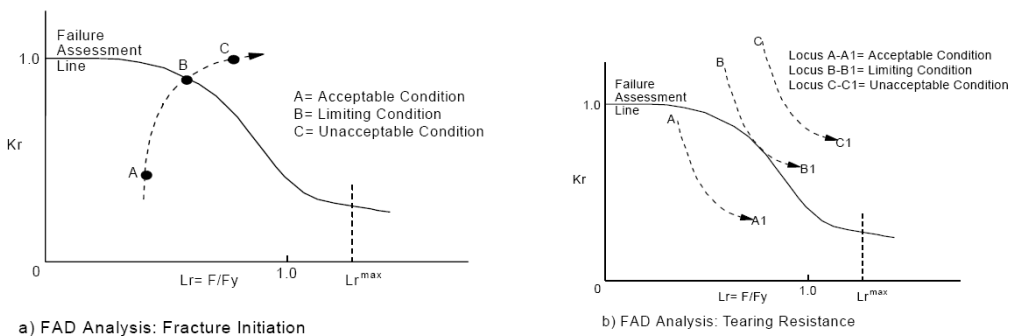
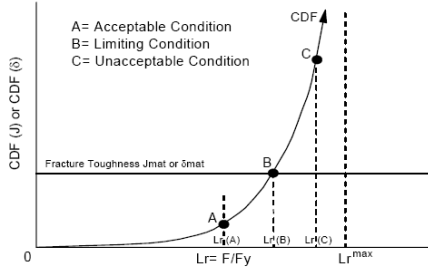
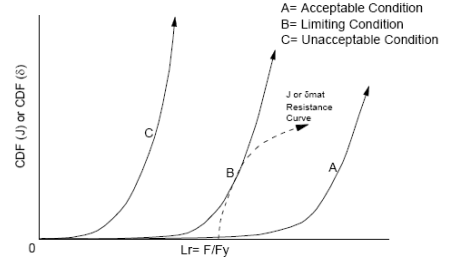


Figure 1. FAD analysis for fracture initiation and ductile tearing (taken from [1])



a) CDF Analysis: Fracture Initiation



b) CDF Analysis: Tearing Resistance

Figure 2. CDF analysis for fracture initiation and ductile tearing (taken from [1])

The CDF approach requires the calculation of the crack driving force on the cracked structure as a function of L_r . The crack driving force may be calculated in units of J , equation (2), or in units of crack opening displacement, equation (3). Both are derived from the same basic parameters used in the FAD approach, the linear elastic stress intensity factor, K_r and L_r . In their simplest forms J is given by [1]:

$$J = J_e [f(L_r)]^{-2} \quad (2)$$

where $J_e = K_I^2/E'$ and

$$\delta = \delta_e [f(L_r)]^{-2} \quad (3)$$

$$\delta_e = \frac{K^2}{E'R_e} \quad (4)$$

R_e is the material's yield or proof strength and E' is Young's modulus, E for plane stress, and $E/(1-\nu^2)$ for plane strain, ν being the Poisson's ratio.

To use the CDF approach, for the basic option of analysis (initiation), the CDF is plotted as a function of L_r to values of $L_r \leq L_r^{\max}$, and a horizontal line is drawn at the value of CDF equivalent to the material's fracture toughness. The point where this line intersects the CDF curve defines the limiting condition. A vertical line is then drawn at the L_r value given by the loading condition being assessed. The point where this line intersects the CDF curve defines the assessed condition for comparison with the limiting condition. Figure 2a gives an example of such a plot.

To use the CDF approach (in terms of J -integral or Crack Tip Opening Displacement, δ) for the higher option of analysis required for ductile tearing, it is necessary to plot a CDF curve as a function of crack size at the load to be assessed. The material's resistance curve is then plotted, as a function of crack size originating from the crack size being assessed. The limiting condition is defined when these two curves meet at one point only (if the resistance curve is extensive enough, this will be at a tangent). Figure 2b gives an example of this type

of plot. As for the FAD approach, margins and factors can be assessed, by comparing the assessed condition with the limiting condition [1].

The choice of approach (FAD vs. CDF) is left to the user, and there is no technical advantage in using one approach over the other. For this reason (and for simplicity), this chapter is based on the CDF approach (more specifically, using δ), but a similar reasoning could be developed following the FAD approach (and totally analogous reasoning for the CDF-J integral approach).

2.2. Mismatch Analysis

2.2.1. An introduction to mismatch

In weldments where the difference in yield or proof strength between weld and parent material is smaller than 10%, the homogeneous (ordinary) assessment procedure explained above can be used for both undermatching (yield stress of weld lower than yield stress of parent material, as is common in Al-alloy welds) and overmatching (yield stress of weld higher than yield stress of parent material, as in most steel and Ti-alloy welds). In these cases, the lower of the base or weld metal tensile properties should be used. For higher degrees of mismatch, a specific mismatch analysis should be used (i.e, Option 2 in FITNET FFS Fracture Module, based on mismatch procedures provided in [3,11]), given that the predictions for undermatching cases may be unsafe if base metal properties are used, while the predictions for overmatching cases would yield over-conservative predictions (but the analysis will be safe). In both cases, actually, the joint behaves as a heterogeneous bi-metallic joint, in which the plastic zone develops as shown in Figure 3. It can be seen that in case of overmatching there is remote plasticity at the base material (this protects the crack against fracture), while in case of undermatching the plastic zone is confined within the weld zone. In energy terms, overmatched welds allow extra plastic energy to be developed in the joint (increasing the load bearing capacity), whereas the undermatched welds limit the amount of plastic energy developed (and also the load bearing capacity). Therefore, it is essential to provide additional shielding mechanisms for such flaws to promote damage tolerant behavior. Development of efficient joint design and “local engineering” methods (e.g. strengthening of the weld area) are required to overcome the loss of the load carrying capacity of undermatched welds in almost all geometries.

2.2.2. Assessment of mismatched structures

Fitnet Ffs provides some guidance on whether the application of the mismatch option is likely to be useful. The following points may be noted:

- The maximum benefit arises in collapse dominated cases and is at most equal to the ratio of the flow strength of the highest strength material in the vicinity of the crack to that of the weakest constituent
- There is little benefit for values of $L_r < 0.8$
- There is little benefit for cracks in undermatched welds under plane stress conditions

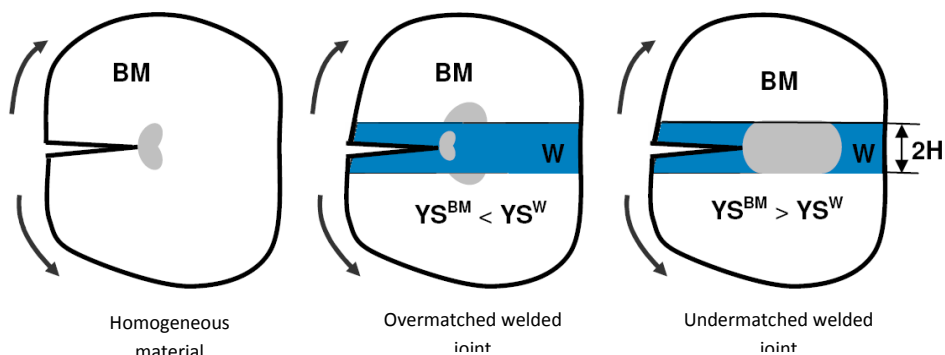


Figure 3. Development of plastic zone in cracked welded joints. BM: Base Material; W: Weld; YS: Yield Stress

This requires knowledge of the yield or proof strengths and tensile strengths of both the base and weld metals, and also an estimate of the mismatch yield limit load. It is, however, possible to use the procedures for homogeneous materials even when mismatch is greater than 10%; and provided that the lower of the yield or proof stress of the parent material or weld metal is used, the analysis will be conservative [1].

Three combinations of stress strain behavior are possible:

- Both base and weld metal exhibit continuous yielding behavior.
- Both base and weld metal exhibit a lower yield plateau.
- One of the materials exhibits a lower yield plateau and the other has a continuous stress strain curve.

The mismatch analysis is performed using FADs and CDFs derived using values of L_r and $f(L_r)$ for an equivalent material with tensile properties derived under the mismatch conditions. In general, for all combinations of yield behavior, this requires the calculation of the following parameters:

- The mismatch ratio, $M = R_e^W/R_e^B$ ($M < 1$ for undermatching; $M > 1$ for overmatching)
- The mismatch limit load, F_e^M , following FITNET FFS nomenclature
- The value for $L_r^{\max}$ under the mismatch conditions
- The value for the lower bound strain hardening exponent N of an equivalent material

All of these are defined in FITNET FFS, as explained below. Advice on calculating the mismatch limit load is given in mismatch assessment procedures [1,3,11], and these also contain solutions for some typical geometries. Here, it is important to note that the mismatch limit load depends not only upon the mismatch ratio but also on the location of the flaw within the weldment (e.g., centre line of the weld material or interface between weld and base material).

Finally, it should be mentioned that mismatch effects can also be considered implicitly by defining the $f(L_r)$ function (equations (1) to (3)) through the full stress strain curves of both

the base material and the weld material. This methodology is gathered in FITNET FFS [1] as Option 3 in the Fracture Module.

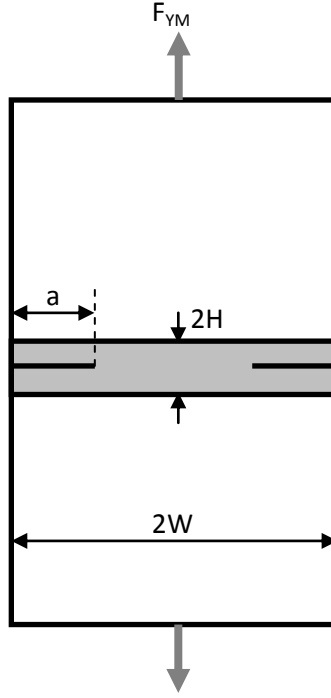


Figure 4. Definition of geometrical parameters for Double Edge Cracked (DEC) panels

As an example, the complete formulation for the simple case of a Double Edge Cracked (DEC) undermatched panel in tension, with a total width W , thickness B and the crack length a is presented (see Figure 4). The height of the central region is normalized by:

$$\psi = \frac{W - a}{H} \quad (5)$$

The limit load for the panel made wholly of base material and for plane stress conditions is given by [1,11]:

$$F_e^B = \beta \cdot 2 \cdot R_e^B \cdot B \cdot (W - a) ; \quad \beta = \begin{cases} 1 + 0.54 \left(\frac{a}{W} \right) & \text{for } 0 \leq \frac{a}{W} \leq 0.286 \\ \frac{2}{\sqrt{3}} & \text{for } 0.286 \leq \frac{a}{W} \leq 1 \end{cases} \quad (6)$$

Then, the mismatch corrected yield load solution, F_e^M , is:

$$F_e^M = M \cdot F_e^B \quad \text{for all } \psi \quad (7)$$

For plane strain conditions, the yield load is given by [1,11]:

$$F_e^B = \beta \cdot \frac{4}{\sqrt{3}} \cdot R_e^B \cdot B \cdot (W - a) ; \quad \beta = \begin{cases} 1 + \ln\left(\frac{2W - a}{2(W - a)}\right) & \text{for } 0 \leq \frac{a}{W} \leq 0.884 \\ 1 + \frac{\pi}{2} & \text{for } 0.884 \leq \frac{a}{W} \leq 1 \end{cases} \quad (8)$$

Then, assuming that yielding occurs within the weld material, the mismatch corrected yield load solution, F_e^M , is given by:

$$F_e^M = \begin{cases} M \cdot F_e^B & \text{for } 0 \leq \psi \leq 0.5 \\ \min\{F_e^{M(1)}, F_e^{M(2)}\} & \text{for } 0.5 \leq \psi \end{cases} \quad (9)$$

$$F_e^{M(1)} = \left[1 - (1 - M) \frac{0.5}{\psi} \right] \cdot F_e^B \quad (10)$$

$$F_e^{M(2)} = \begin{cases} M \left[\beta + A(\psi - 0.5) + B(\psi - 0.5)^2 \right] F_e^B / \beta & \text{for } 0.5 \leq \psi \leq \psi_0 \\ M(0.25\psi + 2.2172) F_e^B / \beta & \text{for } \psi \geq \psi_0 \end{cases} \quad (11)$$

$$A = \begin{cases} 0.25 - \frac{\beta - 2.3422}{\psi_0 - 0.5} & \text{for } 0 < \frac{a}{W} < 0.35 \\ 0.25 - \frac{2(\beta - 2.3422)}{\psi_0 - 0.5} & \text{for } 0.35 < \frac{a}{W} \end{cases} \quad (12)$$

$$B = \begin{cases} 0 & \text{for } 0 < \frac{a}{W} < 0.35 \\ \frac{\beta - 2.3422}{(\psi_0 - 0.5)^2} & \text{for } 0.35 < \frac{a}{W} \end{cases} \quad (13)$$

$$\psi_0 = 16.3 - 35.2 \left(\frac{a}{W} \right) + 19.9 \left(\frac{a}{W} \right)^2 \quad (14)$$

Analogous formulae are provided in [1,3,11] for a number of components, types of cracks and crack positions.

In any case, once the mismatch yield or limit load is defined, the analysis following the CDF route using δ would continue with the calculation of δ :

$$\delta = \delta_e [f(L_r)]^{-2} \quad (15)$$

with the elastic part of CTOD, δ_e :

$$\delta_e = \frac{K^2}{mR_e^W E} \quad (16)$$

K denotes the elastic stress intensity factor, the parameter m ($m=1$ for plane stress and $m=2$ for plane strain, as defined in [1,3,11]) is considered a constraint parameter, E' is E for plane stress and $E/(1-\nu^2)$ for plane strain, and

$$L_r = \frac{F}{F_e^M} \quad (17)$$

is the ratio of externally applied load, F , and the mismatch yield load, F_e^M . The plasticity correction function, $f(L_r)$, is subdivided into different options within the different procedures and is dependent on the extent of the material data input and on the case analyzed (homogeneous or heterogeneous with strength mismatch). For a strength mismatched configuration (and following FITNET FFS Fracture Module, Option 2), the plasticity correction function, $f(L_r)$, is defined as:

$$f(L_r) = \left[1 + \frac{1}{2} L_r \right]^{-1/2} \times [0.3 + 0.7 \exp(-\mu_M L_r^6)] \quad \text{for} \quad 0 \leq L_r \leq 1 \quad (18)$$

$$f(L_r) = f(L_r = 1) \times L_r^{(N_M - 1)/2N_M} \quad \text{for} \quad 1 < L_r < L_r^{\max} \quad (19)$$

where,

$$\mu_M = \frac{M - 1}{(F_e^M / F_e^B - 1) / \mu_W + (M - F_e^M / F_e^B) / \mu_B} < 0.6 \quad \text{else} \quad \mu_M = 0.6 \quad (20)$$

$$\mu_B = 0.001 \frac{E}{R_e^B} < 0.6 \quad \text{else} \quad \mu_B = 0.6 \quad (21)$$

$$\mu_W = 0.001 \frac{E}{R_e^W} < 0.6 \quad \text{else} \quad \mu_W = 0.6 \quad (22)$$

$$L_r^{\max} = \frac{1}{2} \left(1 + \frac{0.3}{0.3 - N_M} \right) \quad (23)$$

Strain hardening exponents for mismatch, N_M , base, N_B , and weld materials, N_W , are defined as follows [3,12,13]:

$$N_M = \frac{M - 1}{(F_e^M / F_e^B - 1) / N_W + (M - F_e^M / F_e^B) / N_B} \quad (24)$$

$$N_B = 0.3 \left(1 - \frac{R_e^B}{R_m^B} \right) \quad (25)$$

$$N_W = 0.3 \left(1 - \frac{R_e^W}{R_m^W} \right) \quad (26)$$

R_m denotes the ultimate tensile strengths of base (superscript B) and weld (superscript W) materials.

Summing up, once the mismatch effect has been considered through the previous parameters, the analysis (CDF route) has the following steps (as in homogeneous materials):

- (a) Calculate δ_e as a function of the applied loads on the structure at the initial flaw size of interest, a_0 , where δ_e has been defined above (equation (16))
- (b) Plot the $CDF(\delta)$ using the appropriate expression for $f(L_r)$ (equation (18))
- (c) Calculate L_r for the loading on the structure at the flaw size of interest and draw a vertical line at this value to intersect the $CDF(\delta)$ curve at $\delta = \delta_{str}(a_0)$
- (d) Repeat the above steps a), b) and c) for a series of different flaw sizes above and below the initial flaw size of interest, a_0 , to give a range of values of δ_{str} as a function of flaw size
- (e) On the axes of δ versus flaw size, a , plot the $CDF(\delta)$ as a function of flaw size where the $CDF(\delta)$ is given by the values $\delta = \delta_{str}(a)$ obtained from steps c) and d) above. Terminate this curve at any point where $L_r = L_r^{\max}$
- (f) Plot $\delta_{mat}(a)$ on this diagram, originating from a_0 , the initial flaw size of interest. This material parameter must be obtained for the same base material-weld-crack configuration (e.g., the analysis of a crack in the centre line of the weld material in a given component would require fracture toughness tests with the crack performed in the centre line of the specimen weld)

Then, if the $CDF(\delta)$ intersects the $\delta_{mat}(a)$ curve, the analysis has shown that the structure is acceptable in terms of the limiting conditions imposed. If this curve only touches the $\delta_{mat}(a)$ curve, or lies wholly above it, the analysis has shown that the structure is unacceptable in terms of these limiting conditions (Figure 2).

A number of applications of this kind of assessments can be found in the literature (e.g., [14-17]), for both undermatched and overmatched situations.

2.3. Consideration of Residual Stresses

2.3.1. An introduction to residual stresses

Residual stresses are those that remain in the structure or component after their original cause has been removed. They are a consequence of interactions between time, temperature, deformation, and microstructure [18]. They occur for several reasons (e.g., thermal gradients), including welding processes. In fact, welding is one of the most significant causes of residual stresses and usually produces great tensile stresses whose maximum value is, in many cases, quite close to the yield stress of the materials being joined. Such tensile stresses are balanced by lower compressive residual stresses elsewhere in the component. In other words, residual stresses are self-equilibrating (net force and bending moment are zero).

Tensile residual stresses may reduce the performance of structures and components. They may increase the rate of damage by fatigue, creep or environmental degradation, and may reduce the load bearing capacity by contributing to failure by brittle fracture. On the other hand, compressive residual stresses are generally beneficial (although they may decrease the buckling load).

2.3.2. Assessment of welds containing residual stresses

When performing structural integrity analyses, it is necessary to define or make assumptions about the stresses in the component being analyzed. This includes normal operational stresses, transient stresses (associated with start-up and shut-down or system upsets), the existence of multiaxial stress states and, of particular importance here, residual stresses at welds (or on cold-worked surfaces) [1]. In any case, the loads or resulting stresses must be separated into primary and secondary: the former arise from loads which contribute to plastic collapse while the latter arise from loads which do not contribute to plastic collapse, since they are caused by strain/displacement limited phenomena. Such a categorization is a matter of some judgment but, in general, primary stresses are produced by applied external loads (e.g., pressure, deadweight, etc), whereas secondary stresses are produced by questions such as thermal gradients and welding processes. However, it should be noted that there are situations where residual (and thermal) stresses can act as primary ones [1], and their consideration as secondary stresses would lead to an underestimation of the stresses causing plastic collapse.

In those cases where the FFS assessment is performed through the CDF (δ) approach, the residual stresses are taken into account in the definition of δ_e . Thus, equation (4) is substituted by:

$$\delta_e = \frac{(K_I^P(a) + K_I^S(a))^2}{E'R_e} \quad (27)$$

where $K_I^P(a)$ is the linear elastic stress intensity factor calculated for all primary stresses and $K_I^S(a)$ is the linear elastic stress intensity factor calculated for all secondary stresses. Assessment procedures (e.g., [1-6]) provide K_I solutions for many typical geometries and, once the residual stresses are known, the definition of $K_I^S(a)$ is totally analogous to the

definition of $K_I^p(a)$. Equation (27) is then introduced in the corresponding expression for δ (substituting equation (3)):

$$\delta = \delta_e [f(L_r) - \rho]^{-2} \quad (28)$$

The parameter ρ takes account of the plasticity corrections required to cover interactions between primary and secondary stresses and depends not only on flaw size but also on the magnitude of the primary stresses (i.e., on L_r) [1]. Procedures such as FITNET FFS [1], SINTAP [3] and R6 [5] provide relatively simple methods for calculating ρ .

In case the residual stresses are of a primary nature, they also affect the plastic collapse analysis, increasing the primary load, F , which is compared to the yield load, F_e , in the definition of L_r (which is also increased). This occurs when residual stresses are long-range residual stress, which are those exhibiting significant elastic follow-up. Under such loading, both the ligament net stress (i.e. reference stress) and the stress intensity factor increase with increasing crack length. Long-range residual stresses usually develop from global or imposed boundary restraint effects, which commonly arise during the fabrication of complex multi-component structures [1].

As shown in Section 2.1, the fracture-plastic collapse assessment of a structure or component can be performed following FAD or CDF approaches, which are compatible equivalent methodologies. Therefore, analogous procedures and formulations to that shown in equations (27) and (28) would be used when performing the assessment of structural components with residual stresses following FAD or CDF (J) approaches.

2.3.3. The magnitude of residual stresses

Once the procedure that includes the residual stresses in the assessment is known, it is necessary to define such residual stresses. The definition of their magnitude to be included in the assessment is a difficult matter and depends, among others, on material, weld design and procedures, structural geometry and (if any) post weld heat treatment (PWHT).

FITNET FFS [1] (Annex C) presents a compendium of recommended residual stress profiles for a range of different configurations of as-welded structural weldments and is principally based on the Section II.7 of the R6 [5] as well as BS7910 [2] and SINTAP [3], although FITNET has provided an update of a number of residual stress profiles, in particular those concerning laser beam and friction stir welded joints. FITNET FFS distinguishes between three types of through-wall residual stress profile, leading to Levels 1–3 of analysis:

- Level 1 profiles readily enable an initial conservative assessment of a defect to be made by assuming a uniform, tensile residual stress field equal in magnitude to the maximum yield stress of the plate or weld material [1]
- Level 2 profiles provide a more detailed but conservative through-wall characterization [1]
- Level 3 profiles represent a more realistic estimate of the specific weld through-wall residual stress distribution based on experimental measurements combined with detailed analysis [1]

A majority of the residual stress profiles recommended in FITNET FFS are essentially upper bounds to available measured and predicted residual stress data. It should be noted that although Level 2 and Level 3 through-wall profiles do not represent realistic self-balancing stress distributions, they do provide a starting point for the quantification of residual stresses that is less conservative than a Level 1 assumption, in almost all cases.

In general, the residual stress field in a welded structure can be characterized by components of stress in the weld longitudinal and transverse directions, $\sigma_{yy}(x,z)$ and $\sigma_{xx}(x,z)$ respectively, and the spatial variation of these components in the transverse (x) and through-thickness (z) directions. The component of residual stress in the through-thickness direction $\sigma_{zz}(x,z)$ is generally small and frequently assumed to be negligible. However, where the $\sigma_{xx}(x,z)$ stress is a concern, or where spatial variations of stress in the longitudinal direction (y) are important, Section II.7.5 in R6 [5] must be consulted. It should be noted that the terms “transverse” and “longitudinal” refer to the welding direction and not the component geometry (i.e., in a pipe circumferential butt weld, the longitudinal and transverse directions coincide with the hoop and axial pipe directions, respectively) [1].

The starting point in the definition of the residual stresses acting on the welded structure is to characterize the residual stress profile at room temperature, either in the as-welded state, or after PWHT. Once the room temperature residual stress distribution has been defined, the effect of mechanical stress relief, assessment temperature or/and historical operation at high temperatures should be considered. An outline of the process is provided here, basically as it is gathered in FITNET FFS [1]:

- (a) *As-welded distribution*: Following FITNET FFS Procedure, three approaches, denoted as Level 1, 2, or 3, for determining the magnitude and spatial distribution of as-welded residual stress are available.

Simple estimates (Level 1) of residual stress magnitude enable an initial conservative assessment of a defect to be made without having to characterize the through-wall distribution. For a weld that has not been stress-relieved, the assumption is that both the longitudinal and transverse components of residual stress are tensile and uniformly distributed in both the through-thickness and transverse directions, with a magnitude equal to the material yield strength at room temperature. In general, Level 1 estimates of residual stress are expected to be conservative for fracture assessments. If adequate safety margins are not achieved using these estimates, then the more detailed characterization approach (Level 2) is recommended.

Level 2 is based on published compendia of conservative residual stress profiles, $\sigma_{yy}(x,z)$ and $\sigma_{xx}(x,z)$ for a range of as-welded structures. The residual stress profiles are given as transverse stresses, σ_R^T (stresses normal to the weld run) and longitudinal stresses, σ_R^L (stresses parallel to the weld run), providing the variation of stresses with through wall distance and normal distance from the weld centre-line. Stresses acting on the through thickness direction are assumed to be negligible [1].

Two approaches for defining Level 2 residual stress profiles are provided in FITNET FFS, depending on the available information about welding conditions:

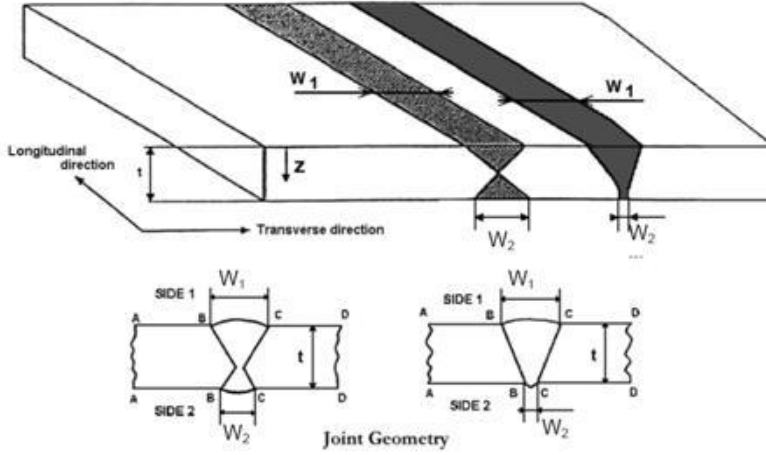
- If the welding conditions are known or can be estimated, then residual stress profiles (e.g., Figure 5) may be used in association with the size parameters of the plastic zone (r_0 , y_0 , as defined in FITNET FFS Annex C).

- If the welding conditions are unknown, then polynomial functions provided by the procedure should be used. Equation (29) provides an example for longitudinal through-thickness residual stresses (σ_R^L) in plate butt and pipe seam welds performed in austenitic steels:

$$\frac{\sigma_R^L}{\sigma_y^W} \left(\frac{z}{t} \right) = 0.95 + 1.505 \left(\frac{z}{t} \right) - 8.287 \left(\frac{z}{t} \right)^2 + 10.57 \left(\frac{z}{t} \right)^3 - 4.08 \left(\frac{z}{t} \right)^4 \quad (29)$$

- t being the thickness, σ_y^W the yield stress of weld material and z as defined in Figure 5.

Material: Ferritic and Austenitic steels and Aluminium (only (a))
Geometry



(a) Longitudinal Surface Residual Stress, σ_R^L

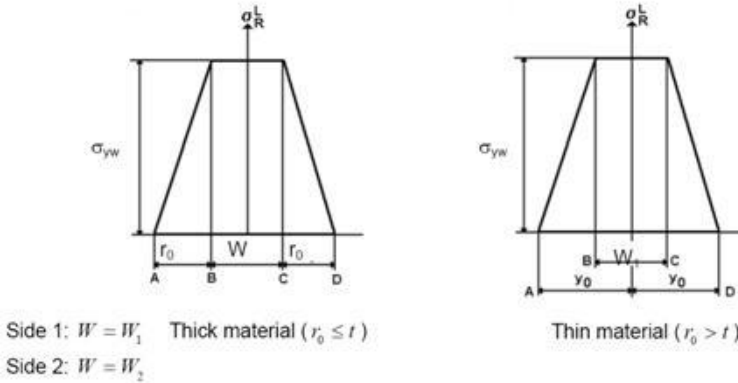


Figure 5. FITNET FFS Level 2 residual stress profile for plate butt welds [1] (r_0 being the radius of yield zone, depending on material properties and welding procedure)

If welds have been repaired, a bounding residual stress profile associated with the repair geometry must be defined. Repairs have the greatest influence on the transverse component of residual stress [1]. Transverse stresses are increased in magnitude, have a more uniform through-wall distribution that can penetrate beyond the repair depth, and have a long transverse range of influence. Simple guidance on defining a bounding stress field is provided in FITNET FFS. A more detailed review covering the effects of section thickness, and the length and depth of the repair is given in [19], and further insight into the effects of repair weld length can be found in [20].

If adequate margins are still not achieved, FITNET FFS (Annex C) provides guidance on how the magnitude and spatial distribution of residual stress can be determined through a combination of analysis and experimental measurements (Level 3). This Level 3 characterization approach is expected to lead to less conservative results but is more complex, more time consuming, and requires detailed information about weld construction, although some validated Level 3 profiles are given in the procedure. FITNET FFS outlines the methods in order of increasing complexity. It is only necessary to proceed to a later step if the earlier, simpler methods do not lead to adequate margins of safety in the assessment.

- (b) *Effect of PWHT*: Welded structures are often post-weld heat treated to improve the metallurgical properties of the weld region and to reduce residual stress. The magnitude and distribution of the residual stresses after PWHT will depend on the initial residual stress state in the body, the weld geometry, the creep behavior of the weld and parent materials, and the nature of the PWHT. Three approaches for characterizing the residual stress field are provided in FITNET FFS. Simple estimates of residual stress magnitude after PWHT enable an assessment of a defect to be made without having to characterize the spatial distribution. A second approach provides guidance on analytical methods for estimating the relaxation in as-welded residual stress. The third approach requires the application of detailed finite element analysis in conjunction with Level 3 as-welded residual stress profiles [1].

The mechanism of stress relief may cause creep damage, cause prior crack tip plasticity in the case of pre-existing defects, or adversely affect the microstructure. For all these cases, the influence of the heat treatment on fracture toughness and crack growth mechanisms must be accounted for in the assessment [1].

- (c) *Effect of mechanical treatments*: Mechanical treatments are often applied to engineering components to improve structural performance (e.g., proof-test). This effect arises from a positive change to the internal residual stress field. However, the effect of mechanical stress relief on fracture depends on whether the structure is cracked or uncracked prior to treatment.

For uncracked structures, the redistribution of residual stress following a proof test depends on weld geometry, the parent and weld material's behavior, the initial stress in the body and the nature of the proof test loading [1]. A simple expression is provided in British Standard BS 7910 [2] for estimating a reduced magnitude of Level 1 residual stress after proof stress loading. The formula is based on idealized uniaxial behavior and also on factors allowing for local weld geometry and work

hardening, but has limited validation and should be used with caution. In cases of uncertainty, the as-welded residual stress profiled should be assumed. For more accurate characterization of the relaxed stress field, detailed Level 3 methods should be applied.

In cracked structures, applications of a prior overload to a cracked structure can enhance the fracture toughness at lower operating temperatures owing to warm pre-stressing effects, and provide assurance of integrity during subsequent operation at lower load. British Standard BS7910 provides an alternative formula for estimating residual stress relaxation effects arising from proof test loading. At present, however, the application of this formula for quantifying the benefits of prior overload is not recommended [1].

- (d) *Effect of assessment temperature:* A uniform increase in the temperature of a welded component usually reduces as-welded residual stress at higher temperatures. This is caused by two mechanisms: first, the elastic modulus falls with rising temperature giving a proportional decrease in elastic stress for the same elastic strain; secondly, a fall in material yield strength with increasing temperature can lead to conversion of elastic strain into plastic strain [1]. The benefit of these temperature effects can be included in the assessment. Thus, Level 1, 2, and 3 estimates of residual stress should be based on the room temperature yield stress multiplied by the ratio of elastic modulus at the assessment temperature to that at room temperature. In case the yield stress at the assessment temperature is lower than the magnitude of the stress factored for elastic modulus, then either the Level 1 stress estimate, the Level 2 peak tensile stress, or the Level 3 peak tensile and peak compressive stress may be reduced to this value [1].
- (e) *Effect of historical operation at high temperatures:* For uncracked structures, it is possible to argue that part of the residual stresses need not be considered in the assessment, providing they have been relieved by historical operation at elevated temperatures. Thus, either the Level 1 stress estimate, the Level 2 peak tensile stress, or the Level 3 peak tensile stress and peak compressive stresses can be reduced to a yield stress value (factored for elastic modulus) that is less than the assessment temperature yield stress, as explained in [1]. In addition, if the component has operated at temperatures within the creep range for the material, residual stresses will further relax due to the accumulation of creep strain, and methods are provided in FITNET FFS to quantify the corresponding stress reduction. However, it is conservative to neglect any relaxation of residual stress due to creep in service [1]. For cracked structures, the relaxed residual stress profile associated with the uncracked structure may be used in the integrity assessment, providing the adverse effects of any creep damage are accounted for in the fracture toughness values and the crack growth laws used. Alternatively, the uncracked relaxed residual stress profile at a conservative estimate of the time when the crack first appears may be used. Volumes 4/5 and 7 of R5 [4] provide methods by which the time-scale for stress relaxation in a cracked structure can be calculated.

The whole procedure proposed by FITNET FFS is illustrated schematically in the flow chart in Figure 6 [1].

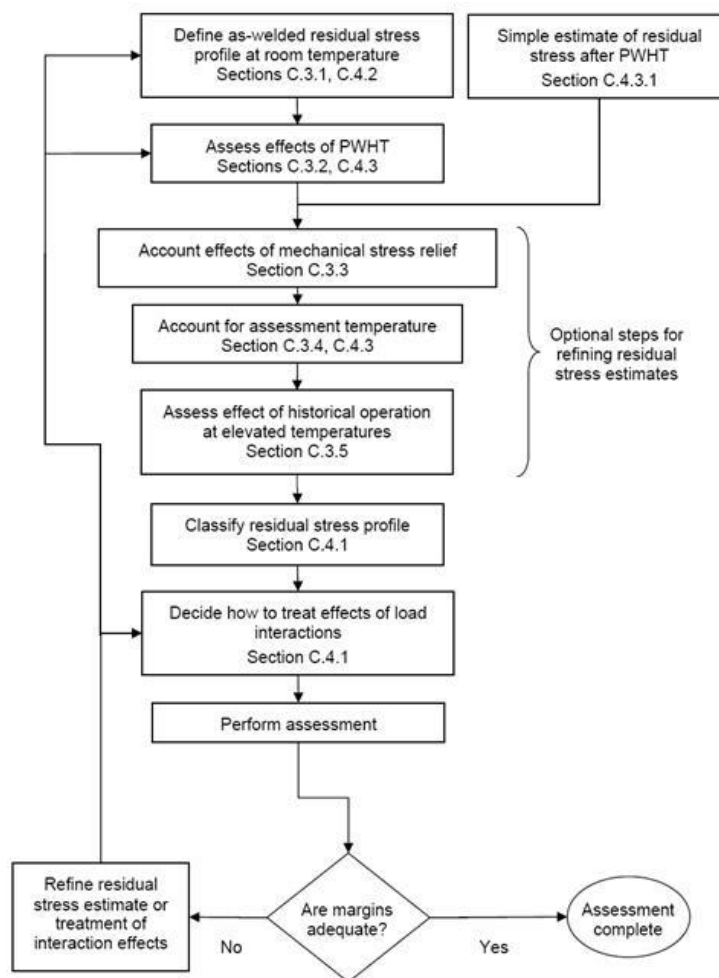


Figure 6. FITNET FFS flow chart for treatment of residual stress (Annex C in [1])

2.4. Consideration of Weld Misalignment

The last issue here analyzed concerning the fracture assessment of welded structures or components is weld misalignment, which is produced when the centerlines of the pieces being joined do not coincide. This causes stress concentrations that should be considered on structural integrity assessments. The presence of misalignment, axial (eccentricity) or angular (Figure 7), or both, at a welded joint can cause an increase (or decrease) in stress at the joint when it is loaded, due to the introduction of local bending stresses [21-23] that usually do not make great contributions to static overload failure (provided the material is ductile). However, they do increase the risk of brittle failure. Thus, there are authors (e.g., [8]) who suggest that misalignment stresses should be treated in the same way as residual stresses (i.e., secondary

stresses) when performing FAD or CDF structural integrity assessments, affecting K_r but not L_r . However, as mentioned above, there are circumstances where residual stresses can act as primary stresses (thus, affecting L_r), as misalignment stresses also do. FITNET FFS proposes, as a general conservative assumption, that misalignment stresses affect both the stress intensity factors (thus, K_r) and the reference stresses/yield loads (and consequently L_r).

In those situations where more than one type of misalignment exists (e.g., both axial and angular), the total induced bending stress is the sum of the bending stresses due to each type. Both tensile (positive) and compressive (negative) stresses will arise as a result of misalignment, depending on the surface or through-thickness position being considered, and special caution should be taken with the relevant sign when calculating the net effect of combined misalignments and when calculating the total stress due to applied and induced stresses [1].

Moreover, misalignment stresses depend not only on their type and extent, but also on factors that influence the ability of the welded joint to rotate under the induced bending moment [1] (e.g., loading and boundary conditions, section shape and the presence of other members, providing local stiffening). The quantification of their corresponding effects requires special analysis (e.g. finite element stress analysis). FITNET FFS postulates that unless it can be demonstrated that restraint on the joint reduces the influence of misalignment, the induced bending stress should be calculated assuming no restraint.

Finally, FITNET FFS provides formulae (e.g., those shown in Figure 8) for calculating the bending stress, σ_s , as a function of the applied membrane stress, P_m , for a number of cases of misalignment, based on the solutions provided in [21-23]. For joints that experience combined membrane and bending stresses, the formulae are used in conjunction with the membrane stress component only.

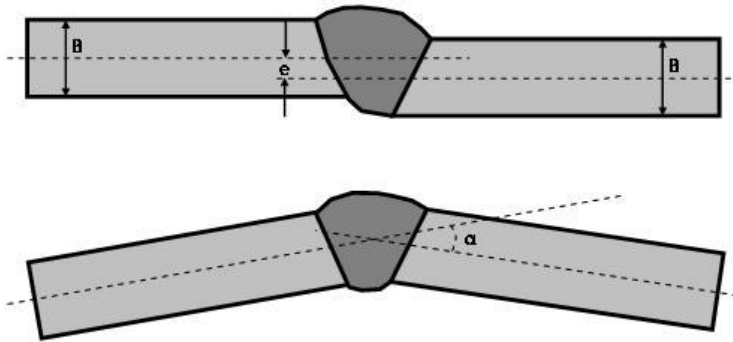


Figure 7. Examples of weld misalignment: axial (top) and angular (bottom)

	$\frac{\sigma_s}{P_m} = \frac{6e}{B_1} \left(\frac{B_1^n}{B_1^n + B_2^n} \right)$	Relates to remotely loaded, unrestrained joints Use $n = 1.5$ supported by tests

Figure 8. Formulae for calculating the bending stress due to axial misalignment between flat plates of different thicknesses [1]

3. FATIGUE ASSESSMENT OF WELDED STRUCTURES

3.1. Brief Overview of Ordinary Fatigue Assessments (Welded and Non-Welded)

The contents gathered in Section 2 refer to structures or components subjected to static or monotonic loading. However, the presence of cyclic stresses may cause initiation (in case there is no pre-existing flaw) and subcritical propagation of cracks that could eventually reach their critical size causing the structural failure. This process is known as fatigue and occurs at stress values well below the material's ultimate tensile stress, and often below the yield stress limit of the material [24].

Summing up, two fatigue analysis approaches are usually distinguished, depending on the existence or not of a crack in the component being analyzed:

- (a) *Fatigue of uncracked components*: there are no pre-existing cracks and the fatigue process leading to fracture is controlled by the (crack) initiation stage. The goal of the fatigue analysis is to determine the accumulation of fatigue damage at a critical location and the basic approach is to determine the fluctuating stress range at the location in question and to relate this to appropriate fatigue life curves. At the same time, depending on the applied stress level, two situations may be distinguished:

High Cycle Fatigue: corresponding to those situations where fatigue stresses are below the material yield stress. This usually leads to more than 10000 cycles to fracture.

The fatigue life curves used in the analysis of this phenomenon are known as S-N curves (as those shown in Figure 9), which provide the number of cycles to failure (N) as a function of the applied stress amplitude ($\Delta\sigma$).

Low Cycle Fatigue: stresses over the material yield stress, usually leading to less than 10000 cycles to fracture. Here, the appropriate fatigue life curves represent the number of cycles to failure as a function of the strain range ($\Delta\epsilon$).

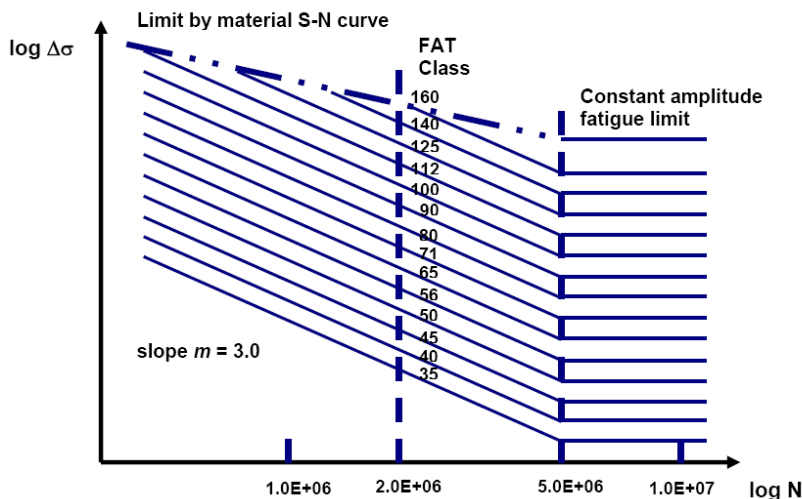


Figure 9. Fatigue resistance S-N curves for $m=3.00$, normal stress (steel) [1]

- (b) *Fatigue of cracked components*: there are pre-existing cracks and the final fracture depends on the crack propagation process. In such cases, the goal of the fatigue analysis is to determine the fatigue life of the component, which is obtained through the Paris law or similar expressions.

The reader is submitted to specific fatigue bibliography (e.g., [25-27]) for further knowledge on this phenomenon and the theoretical background sustaining the different approaches and tools (Paris law, Miner's rule, Coffin-Manson's law, load histogram definition, etc) used for its analysis.

Regardless of the specific fatigue analysis situation (pre-existing flaw or not, high cycle vs. low cycle, etc), the assessment of welded structures and components present specific questions that need to be addressed. Basically, their treatment is quite similar to that provided for non-welded structures, but presenting specific curves or factors attending to their singularities. As was done for the fracture analysis, the following sections dealing with the fatigue analysis of welded structures are based on the treatment given by FITNET FFS Procedure [1] to this phenomenon. The overall scheme of FITNET FFS fatigue assessment procedure is shown in Figure 10. It can be observed that FITNET FFS distinguishes five different routes:

- (a) *Route 1* - Fatigue damage assessment using nominal stresses
- (b) *Route 2* - Fatigue damage assessment using either structural hot spot stress or notch stress
- (c) *Route 3* - Fatigue damage assessment using a local stress-strain approach
- (d) *Route 4* - Fatigue crack propagation
- (e) *Route 5* - Non-planar flaw assessment

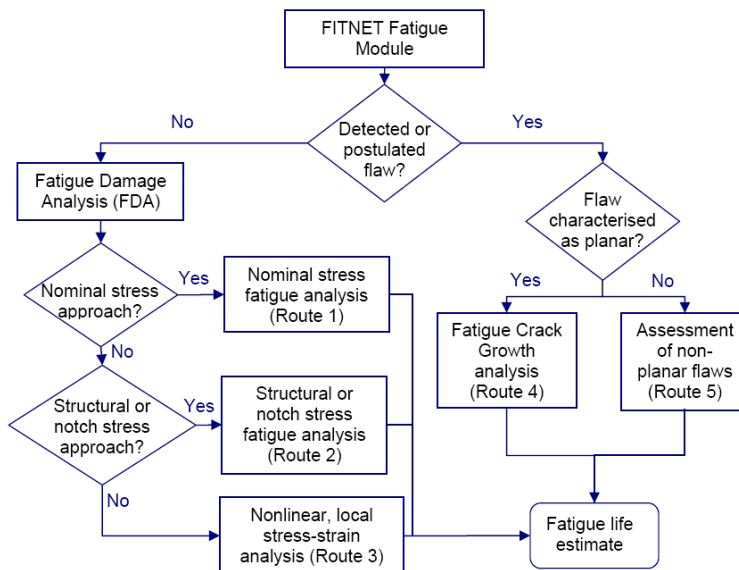


Figure 10. Selection of a fatigue assessment route when using FITNET FFS [1]

The first selection criterion is whether the component is to be analyzed in the presence of an established crack (detected or postulated). If negative, Routes 1 to 3 are followed (Fatigue Damage Assessment, FDA) and if positive, Routes 4 (Fatigue Crack Growth Assessment) and 5 (Non-Planar Flaw Assessment), depending on whether the defect is plane or not [1].

Consequently, Routes 1 and 2 correspond to the above-mentioned high cycle fatigue analysis of uncracked components, Route 3 corresponds to low cycle fatigue analysis of uncracked components and Route 4 refers to the fatigue analysis of cracked components. In case the component presents non-crack-like initial defects, FITNET FFS provides an additional assessment route (Route 5).

Figure 11 shows the basic steps used in applying the five assessment routes, while the scope and background of them are briefly described below [1].

3.2. Particularities in the Fatigue Assessment of Welded Structures

As mentioned above, the fatigue assessment of welded components is analogous to that in non-welded components, using the same tools (e.g., S-N curves, crack propagation laws, etc) which are adapted to address weld specific features such as residual stresses, local geometries, microstructure, etc.

In the following, the specific treatment given by FITNET FFS to welded structures will be presented. The main novelty of the Fatigue Module (Chapter 7, Volume I) of the FITNET FFS Procedure is that it provides clear updated guidelines for carrying out the various types of existing fatigue analyses according to the varying knowledge of the state of the defects.

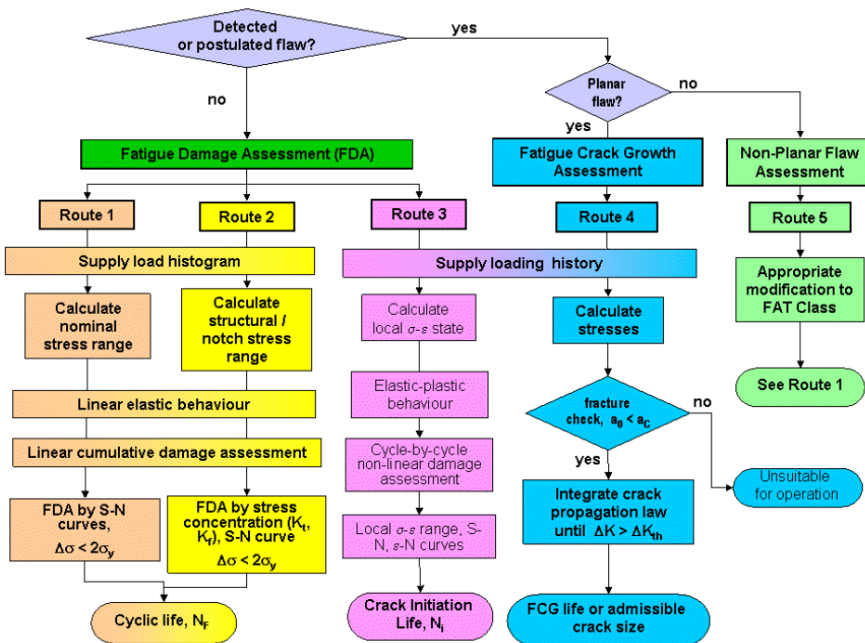


Figure 11. Basic steps in the FITNET fatigue assessment routes [1]

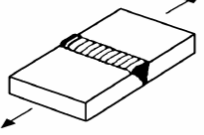
No.	Structural Detail	Description (St.=Steel; Al.=Aluminium)	FAT St.	FAT Al.	Requirements and Remarks
216		Transverse butt welds welded from one side without backing bar, full penetration root controlled by NDE no NDE	 71 36	 28 12	Misalignment <10%

Figure 12. Fatigue resistance values for structural details in steel and aluminum assessed on the basis of nominal stresses [1]

3.2.1. Fatigue assessment of welded structures following FITNET FFS Route 1 (FDA using nominal stresses)

This Route considers the nominal elastic stress values in the location of interest.

In welded components, the fatigue life is determined using a set of S-N curves (Figure 9) classified according to different levels of fatigue resistance for $2 \cdot 10^6$ cycles or FAT Classes (depending on the geometry and the material) provided in Annex G of the Procedure [1], as shown in Figure 12.

These FAT solutions have been taken from [28]. The S-N curve of the component is a straight line which passes through the point corresponding to the FAT value and to $2 \cdot 10^6$ cycles with a slope of 3 (5 for tangential stresses) and becomes constant, with an endurance value (stress variation below which fatigue life is considered to be infinite), when this straight line reaches $5 \cdot 10^6$ cycles (10^8 in the case of tangential stresses). The fatigue curves for welds are based on representative experimental investigations and thus include the effects of [1]:

- Structural hot spot stress
- Concentrations due to the detail shown
- Local stress concentrations due to the weld geometry
- Weld imperfections consistent with normal fabrication standards
- Stress direction and welding residual stresses
- Metallurgical conditions and welding process (fusion welding, unless otherwise stated)
- Inspection procedure (NDE), if specified
- Post weld treatment, if specified

The FAT of the component must also be corrected according to the relation between the minimum and maximum load (R) and the component's thickness. FITNET FFS provides appropriate formulae for these modifications.

In the case of variable load amplitudes, Palmgren-Miner is applied. Figure 13 shows the corresponding working scheme.

It can be observed (Step 4) that FITNET FFS proposes that fatigue assessment is not required when the stress range does not exceed a certain threshold (e.g., in steel components, this occurs when the highest nominal design stress range is lower than $36/\gamma_M$ MPa, γ_M being a partial safety factor taken from an applicable design code).

From this analysis, a nominal stress permissible for the component's life is derived, which is compared with the stress applied to it.

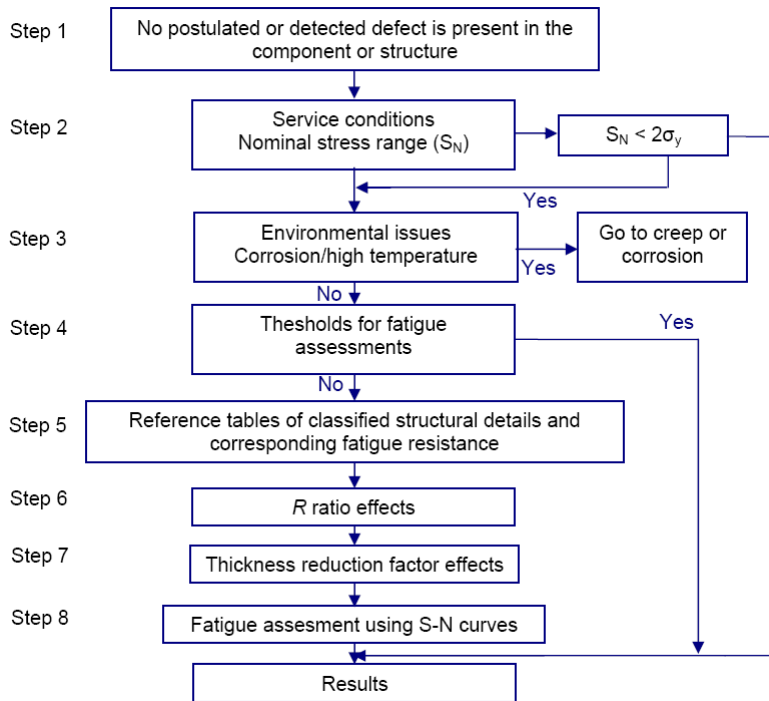


Figure 13. FITNET FFS Route 1 of fatigue analysis in welded components [1]

3.2.2. Fatigue assessment of welded structures following FITNET FFS Route 2 (FDA using either structural hot spot stress or notch stress)

This Route, appropriate for components with stress concentrators, analyses fatigue using two different approaches:

- Calculation of the hot spot stress [29] and application of specific S-N curves (included in the Procedure for a good number of cases).
- Calculation of the notch stress using stress concentration factors such as K_t or K_f [30] and use of specific S-N curves.

In the case of variable load amplitudes, Palmgren-Miner is applied.

Figure 14 shows the definition of the stresses used in this assessment route and Figure 15 shows the assessment scheme for the case of welded components.

The hot spot stress can be obtained analytically from the stresses obtained using finite element techniques at certain reference points (located at a certain distance from the stress concentration which is a function of the thickness). Following [29], the hot spot stress is obtained by multiplying by one stress concentration factor (SCF_{HS}) the nominal stress value (Figure 14). FITNET provides SCF_{HS} expressions for different stress gradient situations and FAT solutions for a number of common cases (Figure 16).

At the same time, the notch stress can be calculated directly by finite elements using linear elastic theory (direct approach) or analytically by multiplying the SCF_{HS} by a new stress concentration factor (SCF_{notch}) which is a function of the weld geometry and cofiguration and post weld treatments (if any) and can easily be obtained from the tabulated

values listed in the procedure. Finally, the corresponding fatigue curves when using notch stresses are also defined in FITNET FFS. As an example, the FAT class for any kind of welded joint when using the direct approach is 225 for steels and 75 for aluminums.

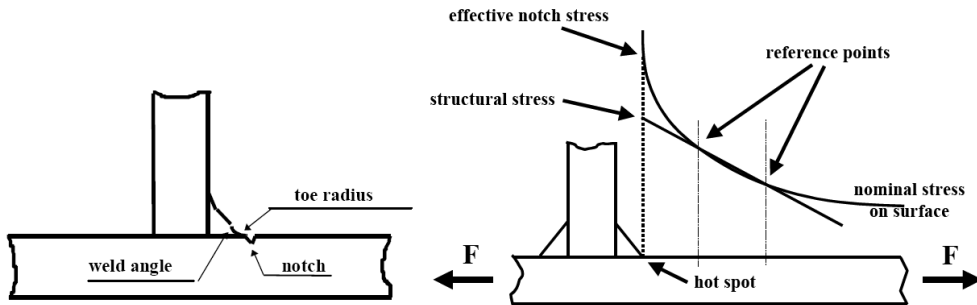


Figure 14. Hot Spot stress (or Structural Stress) and Notch Stress in a welded joint

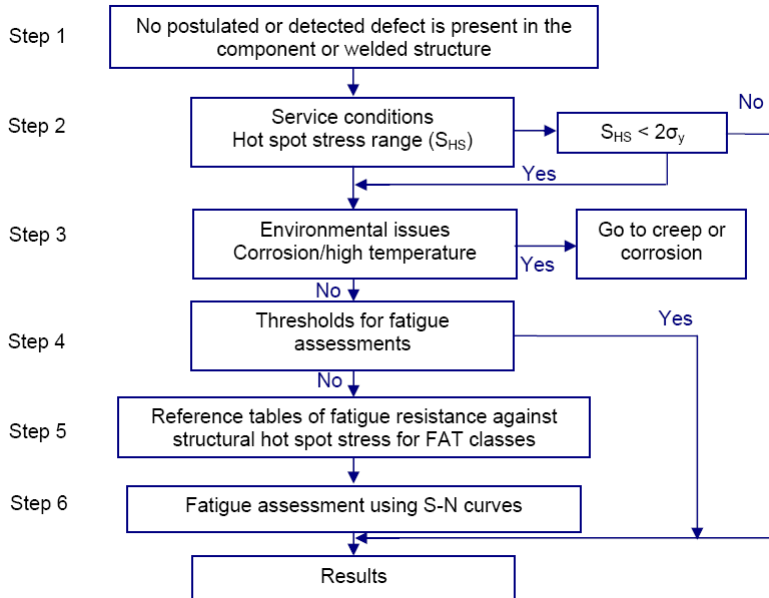


Figure 15. FITNET FFS Route 2 of fatigue analysis in welded components [1]

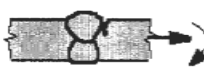
No	Structural detail	Description	Requirements	FAT Steel	FAT Alu.
1		Butt joint	As welded, NDT	100	40

Figure 16. Fatigue resistance values for structural details in steel and aluminum assessed on the basis of hot spot stresses [1]

3.2.3. Fatigue assessment of welded structures following FITNET FFS Route 3 (FDA using local stress-strain approach)

This route is mainly directed at non-welded components and uses a direct calculation of strains at a critical point, making use of the elastoplastic behavior of the material. As there is no specific application to welded components, the reader is referred to the procedure for further information on this assessment Route.

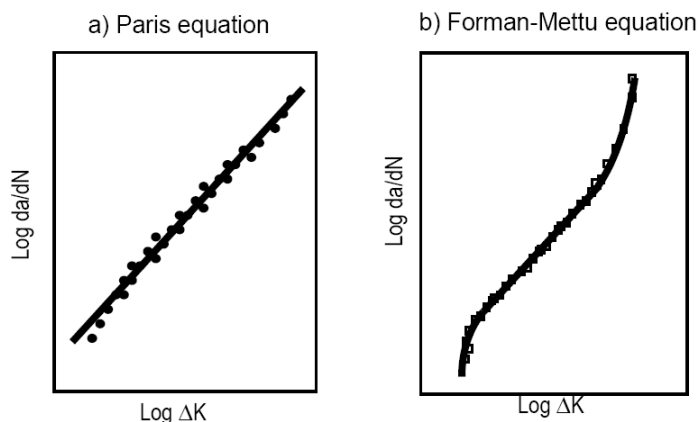


Figure 17. Schematic showing how the fatigue crack growth rate is represented by the Paris or the Forman-Mettu equations [1]

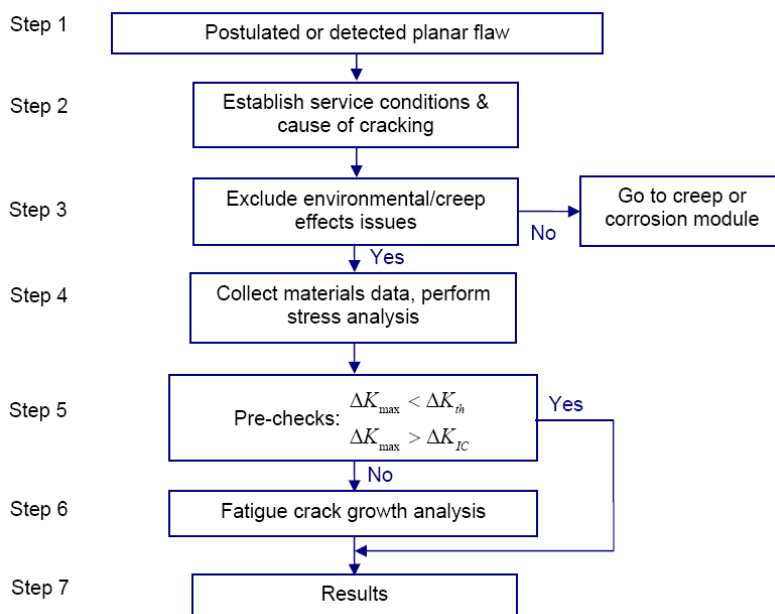


Figure 18. FITNET FFS Route 4 of fatigue analysis in welded components [1]

3.2.4. Fatigue assessment of welded structures following FITNET FFS Route 4 (Fatigue crack growth assessment)

This Route allows a detected or postulated plane flaw to be analyzed. The basic methodology proposed for propagation analysis is the Paris Law but a more sophisticated approach is proposed, based on the Forman-Mettu equation [32], which predicts the fatigue behavior of the material from stress variations typical of the propagation threshold up to those close to fracture (see Figure 17).

Figure 18 shows the corresponding flowchart. The presence of welds is mainly considered in Step 4, on which the materials relevant to the feature to be assessed should be defined, including, in the case of weldments, the weld metal and heat affected zone (HAZ) structures. This means that in case the crack in the component being analyzed is located in the weld material (or in the HAZ), the corresponding crack propagation laws, the fatigue threshold and the fracture toughness should be obtained from standard fatigue specimens with the crack located on the weld material (or in the HAZ).

3.2.5. Fatigue assessment of welded structures following FITNET FFS Route 5 (Non-planar flaw assessment)

Non-plane defects can be assessed as plane flaws following Route 4, obtaining conservative results given that they are not crack-like. However, there are cases in which they can be assessed following Routes 1 and 2 using the S-N curves for welded joints provided that the size of the defects is not greater than certain limits specified in the Procedure. Thus, basically, if the imperfections are not greater than the limits specified by the procedure, Routes 1 and 2 may be applied. If they are, Route 4 should be followed (treatment as crack-like defects).

The overall flowchart for Route 5 is shown in Figure 19.

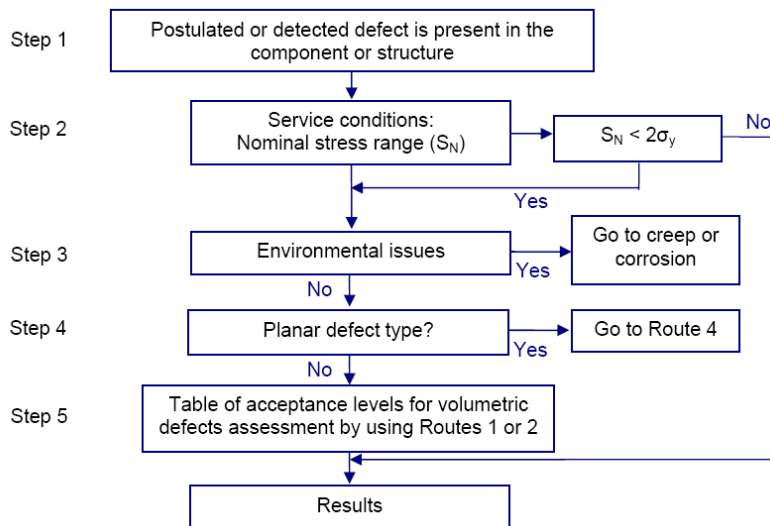


Figure 19. FITNET FFS Route 5 of fatigue analysis in welded components [1]

Material/Weld	Fatigue class	Allowable undercut u/t	
		butt welds	fillet welds
Welds in steel	100	0.025	not applicable
	90	0.05	not applicable
	80	0.075	0.05
	71	0.10	0.075
	63	0.10	0.10
	56 and lower	0.10	0.10

Figure 20. FITNET FFS Route 5 acceptance levels for undercuts [1]. Notes: undercut deeper than 1 mm should be assessed as crack-like imperfection (Route 4); values given are applicable only to plate thicknesses from 10 to 20 mm

At present, this approach is available only for assessing a limited amount of defect types in steel or aluminum alloy butt and fillet welds.

The types of imperfections covered in FITNET FFS Route 5 are the following:

- (a) Imperfect shape: Undercuts (groove melted into the base metal adjacent to the weld toe/root and left unfilled by weld metal). The basis for the assessment of undercut is the ratio u/t (ratio of depth of undercut to plate thickness, as indicated in Figure 20).
- (b) Volumetric discontinuities (gas pores and cavities of any shape; solid inclusions such as isolated slag, slag lines, flux, oxides and metallic inclusions). Acceptance levels for various FAT classes are gathered in FITNET FFS analogously to those shown in Figure 20.

3.2.6. FITNET FFS advices for fatigue life improvement and special options for fatigue analysis

Finally, regarding the fatigue assessment of components, the FITNET FFS covers aspects such as the description of methodologies that improve fatigue life (Burr Grinding, Hammer Peening ...), as well as special analysis options gathering advanced methodologies for fatigue assessments. The former consists in different post weld improvement techniques that may increase the fatigue strength of welded joints that are likely to fail from cracking from the weld toe. Such techniques rely on two main principles:

- (a) Reduction of the severity of the weld toe stress concentration. The primary objective is to remove or reduce the size of the weld toe flaws. A secondary objective is to reduce the stress concentration effect of the weld profile. A variety of techniques belong to this group as shown in Figure 21.
- (b) Introduction of beneficial compressive residual stress, keeping the weld toe in a state of compression with the result that an applied tensile stress must first overcome the residual stress before it becomes damaging. An overview of techniques in the residual stress group is shown in Figure 22.

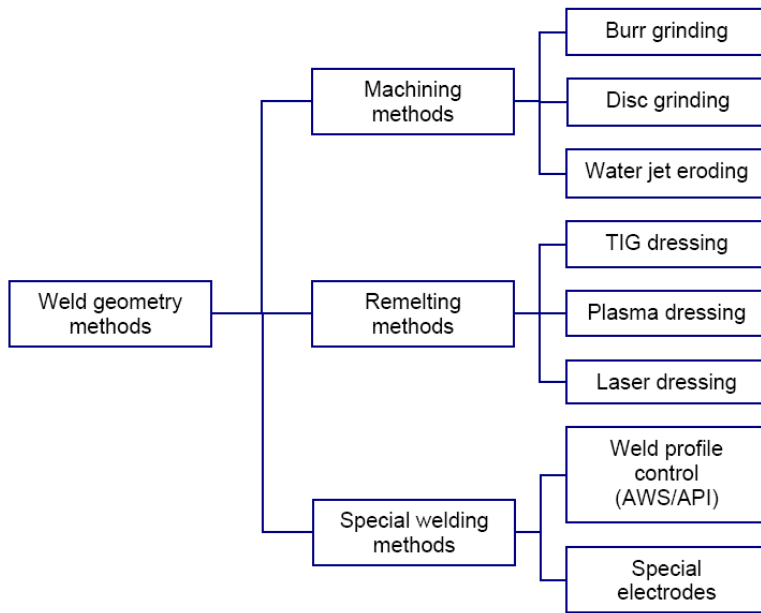


Figure 21. Techniques for reduction of stress concentration factors [1]

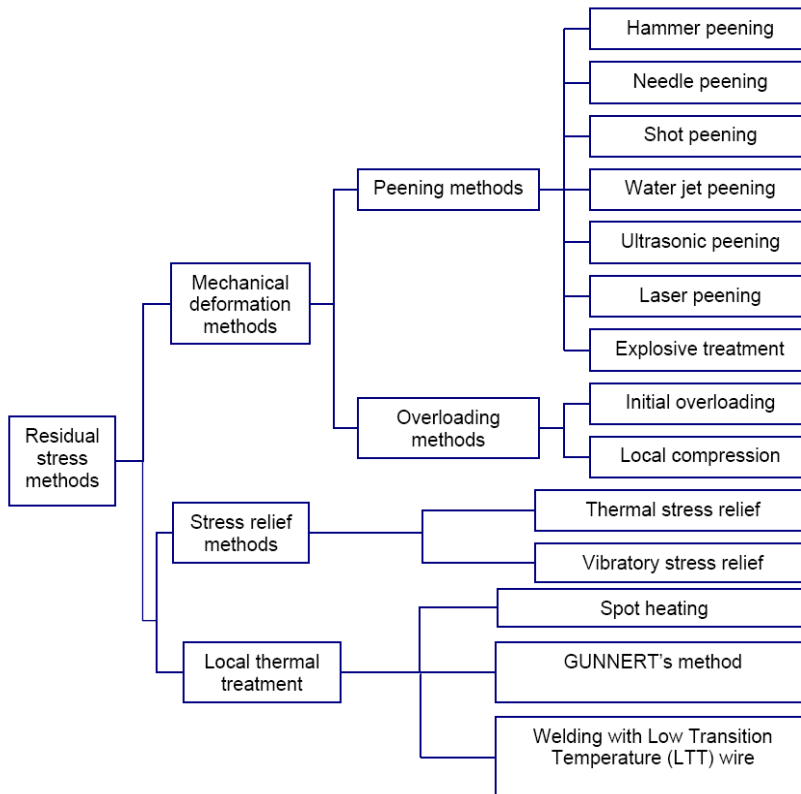


Figure 22. Techniques for modification of residual stress [1]

Annex L in FINTET FFS includes further contents dealing with some of these techniques: burr grinding, TIG dressing, hammer peening and needle peening. Moreover, the proper Procedure proposes (Chapter 7) specific S-N curves for joints that have been improved by any of the four above-mentioned techniques.

Concerning the fatigue special analysis options, FITNET FFS has specific sections dealing with the Dang Van criterion [32,33], multiaxial analysis, rolling contact fatigue, fatigue-creep and fatigue-corrosion interactions and the growth of short cracks.

4. CONCLUSION

The chapter has provided an in-depth insight into the singularities arising when performing fracture and fatigue assessments of welded structures. The methodologies presented here suggest that significant improvements can be obtained when using specific assessment procedures addressing the particular nature of weldments, rather than using traditional overconservative assumptions.

Also, it has been shown how FITNET FFS procedure deals with the fracture and fatigue assessment of welded structures, covering fundamental questions such as mismatching, residual stresses and weld misalignment, in case of fracture assessments, as well as the definition of specific S-N curves and/or crack propagation laws when performing any of the different possible fatigue analysis approaches. Finally, FITNET FFS procedure has shown its ability to deal with the Fitness-for-Service analysis of welded structures, constituting a truly valuable updated engineering tool.

REFERENCES

- [1] FITNET, *Fitness-for-Service (FFS) Procedure - Volume I*. M; Kocak, S; Webster, JJ; Janosch, RA; Ainsworth, R; Koers, Ed; Geesthacht, Germany, 2008.
- [2] British Standard BS 7910: *Guide on Methods for Assessing the Acceptability of Flaws in Metallic Structures*, BSi, London, UK, 2000.
- [3] SINTAP, *Structural Integrity Assessment Procedure for European Industry*, SINTAP BRITE-EURAM Project BRPR-CT95-0024, 1999.
- [4] R5, *Assessment Procedure for the High Temperature Response of Structures*, British Energy Generation, Issue 3, 2003.
- [5] R6, *Assessment of the Integrity of Structures Containing Defects*, British Energy Generation, Report R/H/R6, Revision 4, 2001.
- [6] API 579-1/ASME FFS-1 2007 *Fitness-For-Service*, American Petroleum Institute, 2001.
- [7] FITNET, *European Fitness-for-Service Network*, EU's Framework 5, Proposal No. GTC1-2001-43049, Contract No. G1RT-CT-2001-05071.
- [8] Anderson, TL. *Fracture Mechanics: Fundamentals and Applications*, 3rd edition; CRC Press: Boca Raton, FL, 2005.
- [9] Broek, D. *Elementary Engineering Fracture Mechanics*; 3rd Edition; Martinus Nijhoff: The Hague, The Netherlands, 1982.

- [10]Ruiz Ocejo, J; Gutiérrez-Solana, F; González-Posada, MA; Gorrochategui, I. *Failure Assessment Diagram-Crack Driving Force Diagram COMPATIBILITY*, SINTAP Task 5, Report SINTAP/UC/05, University of Cantabria, 1997.
- [11]Schwalbe, KH; Kim, YJ; Hao, S; Cornec, A; Koçak, M. EFAM ETM-MM 96: The ETM Method for Assessing the Significance of Crack-Like Defects in Joints with Mechanical Heterogeneity (Strength Mismatch), GKSS Report 97/E/9, *Geesthacht*, Germany, 1997.
- [12]Ruiz Ocejo, J; Gutiérrez-Solana, F. *On the Strain Hardening Exponent Definition and its Influence within SINTAP*, Report SINTAP/UC/07, University of Cantabria, 1998.
- [13]Ruiz Ocejo, J; Gutiérrez-Solana, F. *Validation of Different Estimations of N*, Report SINTAP/UC/09, University of Cantabria, 1998.
- [14]Seib, E; Kocak, M. *Fracture Analysis of Strength Undermatched Welds of Thin-Walled Aluminium Structures Using FITNET Procedure*, IIW Doc. X-1577-2005, 2005.
- [15]Kim, YJ; Koçak, M; Ainsworth, RA; Zerbst, U. *Engineering Fracture Mechanics*, 2000, vol. 67, 529-546.
- [16]Cicero, S; Yeni, Ç; Koçak, M. *Fatigue and Fracture of Engineering Materials and Structures*, 2008, vol. 31, 738-753.
- [17]Dzioba, I; Neimitz, A. *International Journal of Pressure Vessels and Piping*, 2007, vol. 84, 475-486.
- [18]Bhadeshia, HKDH. In: ASM International, *Handbook of Residual Stress and Deformation of Steel*; ASM International: Materials Park, OH, 2001, 3-10.
- [19]Bouchard, PJ. *A Review of Residual Stresses at Repair Welds*, Nuclear Electric Report EPD/DNB/REP/0054/96, 1996.
- [20]Dong, P; Zhang, J; Bouchard, PJ. *Journal of Pressure Vessels Technology*, 2002, vol. 124, 74-80.
- [21]Maddox, SJ. *Fitness for Purpose Assessment of Misalignment in Transverse Butt Welds Subject to Fatigue Loading*, London: International Institute of Welding. IIW document XIII-1180-85, 1985 (Unpublished)
- [22]Andrews, RM. *Fatigue and Fracture of Engineering Materials and Structures*. 1996, vol. 19, 775-768.
- [23]Berg, S; Myhre, H. *Norwegian Maritime Research*, 1977, vol. 5, 29-39.
- [24]Ashby, MF; Jones, DRH. *Engineering Materials 1: An Introduction to Properties, Applications and Design*, 3rd edition; Elsevier: Boston, MA, 2005.
- [25]Suresh, S. *Fatigue of Materials*, 2nd edition; Cambridge University Press: Cambridge, UK, 1995.
- [26]Bannantine, JA. *Fundamentals of Metal Fatigue Analysis*, Prentice Hall, 1989.
- [27]Stephens, RI; Fatemi, A; Stephens, RR; Fuchs, HO. *Metal Fatigue in Engineering*, 2nd edition; Wiley Interscience, 2000.
- [28]Hobbacher, A. *Recommendations for fatigue design of welded joints and components*, IIW document XIII-1965-03/XV-1127-03, 2004.
- [29]Niemi, E; Fricke, W; Maddox, SJ; *Structural hot spot stress approach to fatigue analysis of welded components-Designers Guide*, IIW doc. XIII-1819-00/XV-1090-01, 2000.
- [30]Bureau Veritas rules for steel ships classification – *Fatigue check of structural details* – Part B, Chapter 7, Section 4, 2003.
- [31]Forman, RG; Mettu, SR. In: *Fracture Mechanics 22th Symposium 1 American Society for Testing and Materials ASTP STP 1131*, HA; Ernst, A; Saxena, DL; McDowell, ASTM: Philadelphia, PA, 1992, 519-646.

- [32]Dang Van, K. Introduction to Fatigue Analysis in Mechanical Design by the Multiscale Approach, CISM Courses and Lectures, *Springer Verlag Wien*, New York, NY, 1999, Vol. 392.
- [33]Dang Van, K. Criterion for High Cycle Fatigue Failure under Multiaxial Loading. In: *Proceedings of International Confefence on Multiaxial Fatigue*, Sheffield, 1986.

Chapter 7

LASER TRANSMISSION WELDING: A NOVEL TECHNIQUE IN PLASTIC JOINING

***Bappa Acherjee^{1,*}, Arunanshu S. Kuar¹,
Souren Mitra¹ and Dipten Misra²***

¹Department of Production Engineering, Jadavpur University, Kolkata, India

²School of Laser Science & Engineering, Jadavpur University, Kolkata, India

ABSTRACT

Plastics are found in a wide variety of products from the very simple to the extremely complex, from domestic products to food and medical product packages, electrical devices, electronics and automobiles because of their good strength to weight ratio, ease of fabrication of complex shapes, low cost and ease of recycling. Laser transmission welding is a novel method of joining a variety of thermoplastics. It offers specific process advantages over conventional plastic welding techniques, such as short welding cycle times while providing optically and qualitatively high-grade joints. Laser transmission welding of plastic is also advantageous in that it is non-contact, non-contaminating, precise, and flexible process, and it is easy to control and automate.

This chapter discusses all major scientific and technological aspects concerning laser transmission welding of thermoplastics that highlights the process fundamentals and how processing affects the performance of the welded thermoplastic components. With the frame of this discussion the different strategies of laser transmission welding of plastic parts are also addressed. Finally, applications of laser transmission welding are presented, which demonstrates the industrial implementation potential of this novel plastic welding technology.

* Corresponding author: a.bappa@yahoo.com, bappa.rana@gmail.com.

1. INTRODUCTION

Laser welding was first demonstrated on thermoplastics in the 1970's [1] but it found a place in industrial-scale situations only in the last decade. In 1987, Nakamata [2] patented the laser transmission welding technique, as a process in which, the laser beam penetrates the upper transparent plastic part and is converted into heat by the absorbing lower plastic part. The melt is created only where it is needed, in the joining area of the both partners, to form the weld.

Laser transmission welding technique often provides solutions where conventional plastic joining techniques have failed or required to be improved upon. Laser's versatility permitted the replacement of plastic welding techniques based on ultrasonic energy, friction, vibration, electric resistance and heated tool. The gradual replacement of conventional tools by laser in welding in the plastic industries can be justified by the reproducibility of the process, simplicity of processing, decrease of rejection rate and increase of productivity [3]. Laser welding of plastics is suitable for diverse areas of applications.

This chapter presents an overview of the process of laser transmission welding of plastics. The objective is to provide a deeper insight into the laser transmission welding process fundamentals and strategies. The main focus is set on the material properties and process parameters that govern the welding process and the principal phenomena that affect the quality of the joint. In addition to that, a number of applications of laser transmission welding process, which have already been transferred into industrial production, are also reported.

2. LASER TRANSMISSION WELDING PROCESS

Laser beam can be used to weld plastics in two general ways: either by irradiating the surface of a laser-absorbing plastic and welding by fusion or by transmitting a laser beam through a laser-transparent material and welding at the interface with the laser-absorbing material. The former technique is known as direct laser welding and the latter is described as the laser transmission welding process. Laser sources of 2.0 - 10.6 μm wavelength are generally used for direct laser welding process. In laser transmission welding, a laser beam is aimed at two overlapping thermoplastic parts of different optical properties. The first part is designed to be transparent to the radiation at the laser wavelength and the second part is to be absorbent of that radiation. Depending on the thickness and absorption coefficient of the absorbing part, the transmitted energy is absorbed over a certain depth of that material and converted to heat. The heat generated in this way is transported to the transparent part; consequently, both the parts are melted at the joining interface and results in a firm joint as the weld seam. Laser sources of 0.8-1.1 μm wavelength are used for the laser transmission welding process, as plastics have a high transmittance at this wavelength range.

Figure 1 illustrates the working principle of the laser transmission welding process in lap joint geometry. The top part of the plastic is transparent to the infrared laser. The bottom part is either transparent or opaque to the infrared laser. For the case of transparent bottom part, a layer of infrared absorbing dye coating is used as laser absorbing medium. Laser transmission welding can be used for thin as well as thick plastic materials.

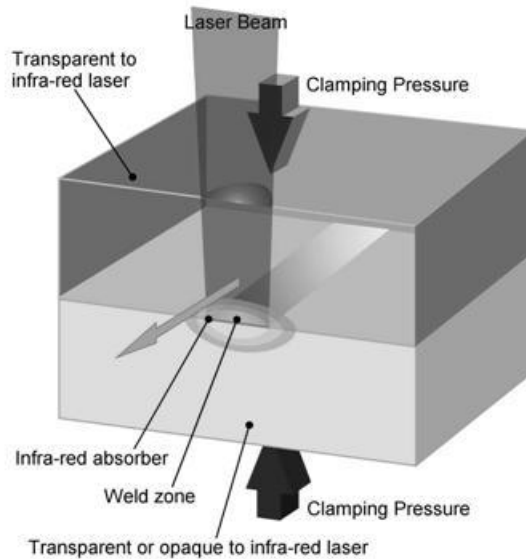


Figure 1. Principle of laser transmission welding process (Reproduced by permission TWI Ltd)

2.1. Laser – Material Interaction

Laser is a concentrated beam of coherent monochromatic radiation. Ordinary light consists of several colors and waves. Therefore, it is not possible to collimate ordinary light without losing its intensity. However, using a monochromatic light source as laser that provides all the waves in single phase, it is possible to concentrate the laser beam using an optical lens to a spot of any desired size without appreciably losing any of its intensity. Thus, laser has become an appropriate radiant energy source to heat and melt the joint for welding of materials.

When the radiant energy strikes a material surface, part of the radiation is reflected, part is absorbed, and part is transmitted.

$$\rho + \alpha + \tau = 1 \quad (1)$$

Where, reflectivity, ρ is the fraction of the radiant energy reflected, absorptivity, α is the fraction absorbed and transmissivity, τ is the fraction of the transmitted radiant energy.

The application of laser beam in welding depends on the thermo-optic interaction between the beam and the work material. So, it is obvious that the work surface should not reflect back too much of the incident laser beam energy. Reflectivity of metals is pretty high, sometimes about 90% for high quality polished surfaces at the operating wavelength of the CO₂ laser. Metals have a relative low reflectance at the wavelengths of Nd:YAG and diode lasers, which makes these lasers more efficient related to the process. The part of light, which is not reflected, enters to the material. The absorbed light propagates into the medium and its energy is gradually transferred to the lattice atom in the form of heat.

$$E_{\text{photon}} = h\gamma_{\text{photon}} = E_1 - E_2 \quad (2)$$

E_{photon} is the energy of each absorbed photon of the laser beam and, E_1 and E_2 , are two energy states of absorbing material. The non-reflected laser beam is absorbed in the metal surface within a thickness of less than a micron and converted into heat. The heat generated at the substrate may finally leads to heating, melting, vaporization or even ionization of the material, which is required for heat treatment, welding or cutting of the metal with the application of laser [4].

Plastics in their natural state are transparent to the laser radiation at the wavelength of Nd:YAG or diode lasers. Plastic parts are rendered laser absorbing by compounding it with appropriate additives. If a small quantity of absorbing additive is used in plastic, then the radiation energy will be absorbed over a broad layer of that material. This phenomenon is termed as volumetric absorption. In this case, absorbed light energy, converted to heat, is considered to be equivalent to total internal heat generation in the plastic. The ability of absorption of radiant energy in absorbing plastic is determined by the Beer-Lambert's law, which states that the intensity of a beam of monochromatic radiation in an absorbing medium decreases exponentially with penetration distance.

$$I(z) = I_{(z=0)} e^{-Kz} \quad (3)$$

Where, I is radiation intensity (W/m^2), z is distance within the material and K is the total extinction coefficient (m^{-1}) caused by the laser beam absorption and scatter. For the case of amorphous polymers, the extinction is determined by the absorption only. The absorption coefficient depends on the quality and the color of the plastic material. It is defined as the reciprocal value of the optical penetration depth dp [5].

$$K_{\text{absorption}} = dp^{-1} \quad (4)$$

Thus, the optical penetration depth has great influence on the laser transmission welding process.

Plastics containing sufficient amount of laser absorbing additives, absorb the radiation energy in a very thin layer of that material. This phenomenon is known as surface absorption. In this case, the absorbed light energy is converted into heat at the surface itself, similar to metal, and the laser beam may be considered to be equivalent to a surface heat flux.

The volumetric heat generated or the surface heat flux deposited in this way is transported by thermal conduction into the deeper layers of absorbing part and also into the part that is transparent to the laser beam. Some part of that heat is also transferred to the surroundings through convection and radiation. When heated to a temperature above the melting point, or melting range, it causes melting of a thin layer of plastic in both parts. Clamping pressure ensures the contact between the parts to be joined and also an increase in molten metal flow at the weld zone. Molecular diffusion occurs and a solid joint is formed as the melt layer solidifies.

2.2. Basic Requirements

The two plastic parts to be welded together must have different optical properties, that is, one of the plastic parts should be transparent to the wavelength of the laser beam and other one absorbent of the laser beam. In some cases, where both the plastic parts are transparent, a laser absorbing third material is to be placed at the area of heat generation between the joining surfaces. Sufficient contact between the mating parts is needed to allow for the heat to be conducted from the absorbing material to the transparent material. Both the materials must have chemical compatibility and the difference between the melting temperatures of those materials should not be too high [6].

2.3. Laser Used

Three types of lasers generally are used for laser transmission welding of plastics: Nd:YAG, diode and fiber lasers, operating in the wavelength range of $0.8\text{ }\mu\text{m}$ - $1.1\text{ }\mu\text{m}$, where the plastics have minimal intrinsic absorption, permitting successful laser transmission welding of parts having millimeter thickness. In early 90's, the Nd:YAG lasers ($1.064\text{ }\mu\text{m}$ wavelength) were mostly used as the laser source for welding of plastics. These machines took up a large amount of space and required a great deal of maintenance. The application of laser for welding of plastics remained limited partly due to the high investment cost and low efficiency of these laser systems. The replacement of Nd:YAG laser by the modern diode laser has appreciably increased the interest for applying laser in welding of plastics. These lasers are compact, reliable, comparatively inexpensive and flexible – emitting radiation in the range between $0.8\text{ }\mu\text{m}$ and $1.0\text{ }\mu\text{m}$. The modern diode lasers have air cooling system, which replaces the complex and expensive water cooling systems and reduces energy consumptions. Another advantage to note is that the electrical-to-optical efficiency of diode laser at 37-50% is much higher than that for CO_2 lasers, about 10% and very much higher than that for Nd:YAG lasers, 3-5%. However, diode lasers have relatively low beam quality than Nd:YAG lasers. Fiber lasers (of $1.1\text{ }\mu\text{m}$ wavelength) have emerged as the direct alternative of Nd:YAG lasers in the field of plastic welding as they are operated at very close to Nd:YAG lasers' wavelength, with equivalent beam quality but greater efficiency [4, 7].

Most of the plastics absorb CO_2 laser beam ($10.6\text{ }\mu\text{m}$ wavelength) within a very short depth of the material. Because of that the CO_2 lasers are not suitable for laser transmission welding process. CO_2 lasers are mainly used for welding of thin plastic films in packaging industries.

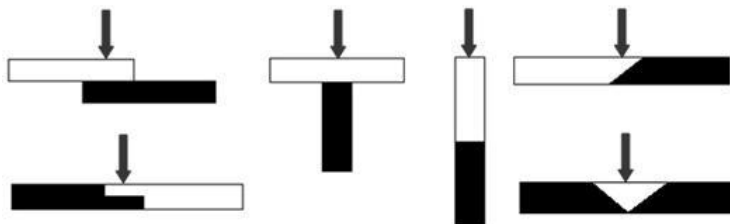


Figure 2. Various possible joint configurations for laser transmission welding process

Infrared lamps are also used as the non laser source for some plastic welding applications such as butt welding of plastic pipes. These systems have disadvantages of longer cycle time, less energy efficiency, less control over energy input to the weld and limited lamp life.

2.4. Joint Design

The joints must be designed in such a way that the sufficient laser energy reach to the joint interface and there be an appropriate area for pressure to be applied to the joint. The materials should be of high finish to reduce any possible air gap between the mating parts to ensure contact conduction. Figure 2 shows different possible joint configurations for laser transmission welding process. The most preferred joint configuration for laser transmission welding process is lap joint, where a transmitting polymer is placed on top of an absorbing polymer. T-joint is also a common configuration used in laser transmission welding. However, butt joint is rather difficult to achieve, as it requires high optical penetration depth in the transparent medium. Applying weld pressure is also difficult for butt welding. Meltdown may occur, specially, in butt and T-joint welding, due to squeezing out of molten material under clamping pressure when the joint interface softens or melts.

2.5. Process Variants

Contour welding- In this laser welding process, either a focused laser beam is moved over the workpiece surface along a predefined path or the laser source is kept fixed and the workpiece moves to make a continuous weld following the weld seam geometry as shown in Figure 3. For a moving laser source system, the optical radiation is delivered to the workpiece via an optical fiber cable mounted on a gantry or a robotic arm system. The workpiece is moved with an X-Y table or by a robotic system, when the laser source is fixed at a position. The contour laser welding process is simple and flexible. Moreover, this process is easy to control and cost effective.

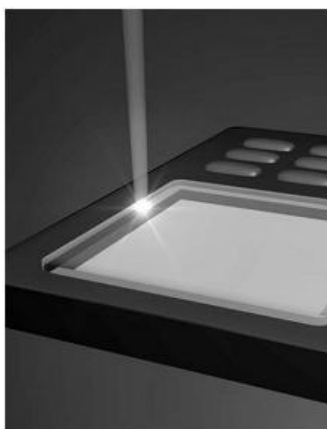


Figure 3. Working principle of contour welding (Reproduced by permission Leister Technologies)

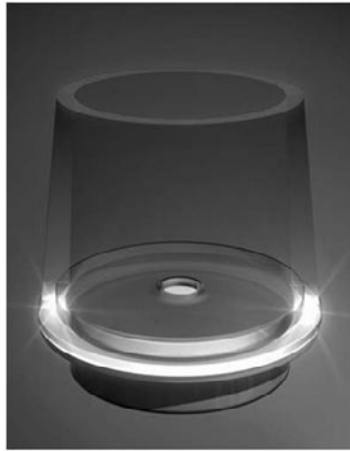


Figure 4. Working principle of simultaneous welding process (Reproduced by permission Leister Technologies)

Simultaneous welding- In this process an array of diode laser modules with homogeneous intensity distribution are arranged in such a way that they irradiate the entire weld line simultaneously in a single exposure of requisite cycle time. The main concern is to arrange the diode laser module in a way to avoid overlapping of beam spots and the non-irradiated areas in the entire weld seam contour. The laser intensity must be homogeneous over the complete weld zone to ensure uniform weld and to avoid material decomposition or lack of fusion. This process is fast but complex, expensive and less flexible compared to contour laser welding process. Figure 4 illustrates the working principles of simultaneous laser welding process.

Quasi-simultaneous welding- In this welding process, neither the laser head nor the workpiece moves along the desired weld contour. The laser beam scans the workpiece several times, along the weld lines, by a galvo mirror system, at a very high speed. Because of the low thermal conductivity of the polymers the entire weld seam heat up gradually and about equally, such that the material along the weld seam melts quasi-simultaneously. The high welding speed minimizes the heat loss to the surroundings, which prevents the molten material from re-solidification during the process. This process is particularly suitable for the two-dimensional welding contours and has found applications in manufacturing of automotive sensors and electronic housings. Another limitation of this process is the maximum working area that the scanning device can cover. Figure 5 demonstrates the working principles of simultaneous laser welding process.

Mask welding- In this laser welding process a mask is used to ensure that the laser beam reaches only to the exposed surface. The mask is con-formal with the desired weld seam and placed between the laser and the workpiece. The open areas of the mask are then laser scanned simultaneously or by a line scan, as shown in figure 6. This process offers an alternative solution for the simultaneous welding process regarding the limitation of the seam geometry. The efficiency of the process is reduced because some part of laser beam is blocked by the mask and could not be used for the process [4].

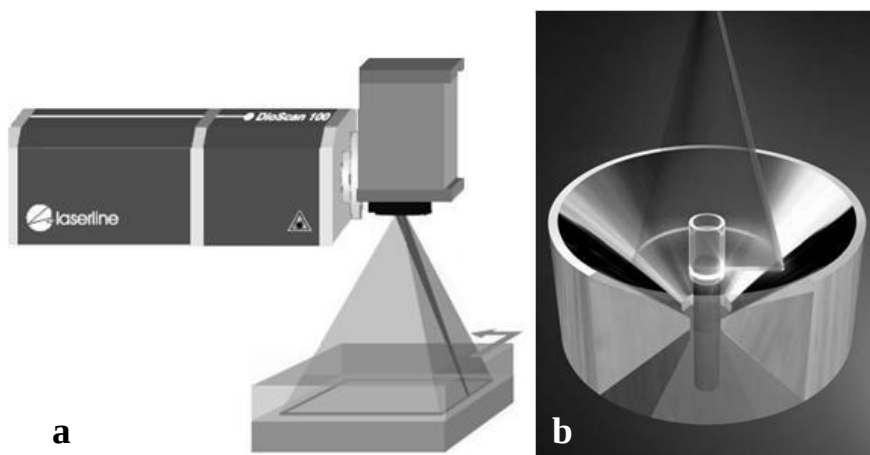


Figure 5. Working principle of quasi-simultaneous welding process (a) orthogonal weld contour (Reproduced by permission Laserline GmbH) and (b) circular weld contour (Reproduced by permission Leister Technologies)

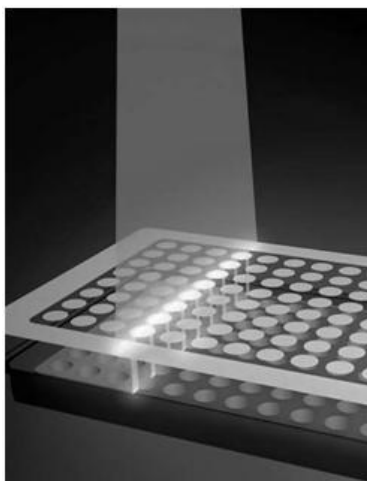


Figure 6. Working principle of mask welding process (Reproduced by permission Leister Technologies)

2.6. Relevant Material Properties

Proper assessment of the following materials and optical properties of the plastics are very important for the functionality of laser transmission welding process, and also for the design and manufacturing of plastic parts by laser transmission welding process [8, 9]:

1. Polymer composition- fiber-glass reinforcement, mineral fillers, impact-modifiers, heat stabilizers, and other additives content by % wt. in polymer matrix
2. Colorants- type of colorants, and content, by % wt.
3. Thermo-physical properties- density, specific heat and thermal conductivity of the plastic

4. Plastic condition before welding- dry as molded, moisture content, by % wt.
5. Constituent (polymer and additives) properties- polymer crystallinity, polymer melting point and additive particle size
6. Optical properties- laser energy transmission and absorption, polymer's and additive's refractive indices

3. EFFECTS OF PLASTIC COMPOSITIONS

The efficiency of laser transmission welding strongly depends on the optical properties of the plastic parts to be joined. The basic composition of polymer matrix, colorants and additives affect laser energy absorption, reflection and transmission and finally to the mechanical performance of the weld.

Most of the polymers in their natural state are transparent to the infrared wavelength. When the laser beam strikes the transparent plastic part, a fraction of the incident light is reflected back from the top surface of the part and the remaining light energy transmitted through the material. A portion of the incoming radiation may be absorbed in the bulk of the transparent material due to the possible scattering.

Presence of reinforcements, mineral fillers, impact modifiers and some heat stabilizers in polymer matrix lower the transmissivity of polymer due to increased scattering effect [6]. The laser transmission decreases with increase of the fiber-glass content in the specimen due to the increased light scattering [9, 10]. The addition of mineral filler is more detrimental to the laser transmission than that of fiber-glass reinforcement, because the filler has a great number of scattering centers for the same weight of reinforcement content [29]. Increasing the fiber-glass content in the laser transmitting polymer, increases the tensile strength of the part but reduces at the weld. The weld width increases with fiber-glass content, due to the increased scattering at the transparent part. This results in increase of laser spot diameter at weld interface and decrease of energy density [11]. Minimum power requirement for welding is proportional to the fiber-glass content in transparent polymer. It is studied that the increase of fiber-glass content from 6 - 45 % wt. in 3.2 mm thick nylon 6 plastic parts, increases the minimum power requirement from 12 - 44 W/cm to make a weld in 2 seconds [12].

The impact modifier can reduce the light transmission significantly, even more than the fiber-glass reinforcement of same levels. Laser transmission is reduced by about 50% in natural 3.2 mm thick nylon part, due to scattering by small inhomogeneities introduced by the modifier, depending on the type used.

The use of the flame retardant also has substantial effect on the laser transmission. The addition of flame retardant in polymer matrix diminishes transmission by 60 - 70 % relative to the natural nylon 6.

Colorants are used in plastics to introduce color either for decoration or for some functional needs. Pigments and dyes are different types of colorants. Pigments do not dissolve but dyes dissolve into the polymeric application medium. Pigments are generally classified as organic or inorganic. Organic pigments generally show better transmittance than that of inorganic pigments (such as carbon black and titanium oxide) because organic pigments have smaller particle size with low refractive index than inorganic. Use of colorants in polymer influence the optical properties not only in the visible region but also the near infrared, where the emission of diode and Nd:YAG lasers take place [4, 13]. Kagan et al. [9] investigated the

effects of colorants, namely: green, yellow, red, white and black pigments on 3.2 mm thick nylon 6 plastic. They observed that the transmission of red color specimen is similar to natural color while white, yellow and green colors reduce transmission by 75 - 85%. They believed that the red is most likely an organic, while others are likely to be the inorganic pigments. The plastics, which are rendered black by using carbon black pigments, show very low transmission while non-carbon black plastics have relatively greater transmission. Titanium oxide, the most important white pigment, provides high degree of opacity because of maximum light scattering but with minimum absorption, which needs more laser energy for welding [13]. They also studied the effects of colorants on mechanical performance of the weld for PA6 specimens. Red colored PA 6 shows maximum tensile strength at weld followed by blue, black (non-carbon black), grey and natural PA.

The bottom polymer part of the assembly for laser transmission welding process is rendered laser absorbing through compounding it with colorant such as carbon black, in appropriate proportions. The absorption coefficient of the absorbing polymer increases with the carbon black content in the polymer matrix. The laser penetration depth decreases with the increase of carbon black content in polymer matrix [14]. When a low pigmented (% wt.) absorbing polymer is used, only moderate temperature rise is obtained at the interface and most of the energy deposited inside the absorbing material causes only a limited heat transfer to the transparent part that induces the asymmetric shape of the weld seam into two materials. A high content of pigment favors absorption at polymer interface, and therefore thermal diffusion occurs equally in both polymers, which results in symmetric weld in both the parts [15, 16]. Increase in the carbon black content in laser absorbing polymer part causes decrease in the thickness of heat affected zone and increase in the melt temperature and weld strength [17]. Jansson et al. [18] observed that increasing the carbon black content in absorbing polymer from 0.5 - 1.5 % wt. causes a slight increase in the minimum weld strength. But further increase of carbon black content does not have an effect on maximum weld strength.

4. EFFECTS OF PART THICKNESS

Thickness of plastic part also has influence on the optical properties, especially for semi-crystalline materials. Kagan et al. [9] observed that the degree of laser transmission is a function of plastic part thickness for nylon 6, a semi-crystalline material. For natural and red color nylon 6, laser energy decreases monotonically from 85 - 42% with an increase in the thickness of plastic part from 0.8 - 6.25 mm. While for yellow, green and white plastics a reduction of 60 - 3% is observed with the increase in the thickness over the same range for same input laser power.

5. EFFECTS OF WAVELENGTH

The influence of wide range of infrared wavelengths (from 0.83 - 1.064 μm) on the optical properties of thermoplastics is evaluated by Kagan et al. [10] for unfilled, filled and reinforced polyamide 6, 66 and amorphous grades. At near infrared spectral wavelength, natural (uncolored) plastics absorb a very small portion of the laser energy. Natural polyamide absorbs up to maximum 20% of the laser energy of diode and Nd:YAG laser,

working at near infrared wavelengths from 0.8 - 1.064 μm . Adding organic green colorant to polyamide increase its absorption to 60 - 90% depending on wavelength in the range of study. Except for the green specimen, decreasing the wavelength from 1.064 - 0.83 μm slightly decreases the transmittance of yellow, white and natural state unfilled PA based plastics. Highly transmissible optical polymers, such as acrylic, polycarbonate, methylmethacrylate styrene and polystyrene in their natural state are non-sensitive to wavelength change in the range from 0.4 - 1.08 μm [10].

6. WELDING PARAMETERS AND THEIR EFFECTS

The laser power density and the laser interaction time are the most important parameters for any laser material processing applications. The most important independent process parameters for the contour welding are laser power, welding speed, size of the laser beam spot on the work-piece and clamping pressure [19]. In quasi-simultaneous welding the principal process parameters are laser power, scanning speed and the number of scans [18]. The temperature field inside the weld during welding can be controlled with these process parameters [16].

The energy density used during welding combines the process parameters of laser power density and the laser interaction time. It is determined by the laser power, size of the laser beam spot on the work-piece and the laser irradiation time or welding speed.

$$\text{Power density} = \frac{\text{Power}}{\text{Spot size}} \quad (5)$$

$$\text{Energy density} = \frac{\text{Power} \times \text{Time}}{\text{Spot size}} \quad (6)$$

$$\text{Energy density} = \frac{\text{Power}}{\text{Spot size} \times \text{Speed}} \quad (7)$$

The weld strength is restricted by very high energy density, which causes overheating and partial decomposition of the material, and a very low energy density results in lack of fusion [20]. The optimum weld strength can be achieved at a favorable value of energy density with an appropriate combination of laser power and welding speed. The same energy density can be achieved by combining either low power with low welding speed or high power with high welding speed. In the case of low welding speed required by the relatively low laser power, the heat transfer comes into play here more as the heat conduction losses have greater impact at these slow speeds. At high speeds, using higher laser power for the same energy density, heat loss is minimized due to the less time available for heat dissipation, and maximum of the input energy is deposited at the weld zone [21].

Prabhakaran et al. [22] studied the effect of contour laser welding parameters on meltdown and weld strength for T-joint welded 30% glass reinforced Nylon 6. It is observed that the melt down is a direct function of laser input energy, defined as the ratio of the laser

power to welding speed. It is also found that optimum weld strength can be achieved by an appropriate combination of laser power and welding speed values. For the range of weld parameter studied, meltdown increases but weld strength decreases with increase of the weld pressure. Baylis et al. [6] investigated the effect of laser welding parameters on the laser transmission weld quality, defined by weld width and strength for lap welded thermoplastic elastomers to polypropylene. They observed that the track width (the heat affected zone plus weld width) and weld strength increase with line energy i.e., the laser input energy per unit length (J/mm). Douglass and Wu [23] considered laser power, welding speed and clamping pressure as input parameters and determined their effect on the lap shear strength of lap welded soft and hard polyolefin elastomer (POE) to thermoplastic polyolefin (TPO). The regression analysis resulting equations for lap shear strength of soft and hard POEs to TPO welded specimens confirm that the power and speed have the most significant effects on the welding. By increasing the power and decreasing the travel speed the joint strength can be increased. Next most dominant is that of the combination of power and speed i.e., line energy, which tends to be positive with respect to strength. For both the materials, pressure has the little positive effect on the strength.

Acherjee et al. [21] presented a detailed study on the effects of laser transmission contour welding parameters on the weld quality of acrylics. It is observed that, both, weld strength and weld width, increase with laser power. Increasing the laser power increases the heat input to the weld zone, thus, more base material is melted, resulting higher weld strength and wider seam width. However, weld strength increases until the critical temperature of decomposition is reached. It is also found that welding speed has a negative effect on weld strength as well as weld width. This is so, because the energy deposition and heat diffusion into the material in laser transmission welding depends on the laser power density and the irradiation time. Higher the speed, lower is the irradiation time, causing low heat input to the weld zone, resulting narrow and weak weld. Clamp pressure showed a little positive effect on the both, weld strength and weld width. Clamp pressure ensures good contact between the parts to be welded. This enhances the conduction of heat from the absorptive material to the transparent part and also promotes the molten fluid flow, required for intermixing and cross linking of the polymer chains to combine towards weld formation. Laser beam spot diameter also showed a very significant effect on the weld quality as it controls the power density over the irradiation zone.

Coelho et al. [3] observed that the weld quality does not depend only on the energy delivered to the sample, but also on the spot shape of the laser beam. This is due to the fact that the amount of molten material contributed to the weld seam increases with the size of laser affected zone and the irradiation time of that zone. A spherical lens and a cylindrical lens as alternate laser beam focusing system is used for producing circular and elliptical focal spot, respectively, to study the effect of laser spot shape on welding result. They observed that no welding can be achieved at the required speed with the sample placed at the beam focal spot. With a defocused beam, the width of the interaction area increases, which increases seam width, thickness of molten layer and improve weld tensile strength. Defocusing also leads to a more uniform energy distribution by decreasing the energy gradient inside the spot. For an elliptical beam spot, less power is necessary to achieve critical specific energy for good weld which lead to higher welding speed for same maximum power than that for circular spot. Using an elliptical beam spot with its larger dimension directed along the movement direction, welding of transparent high density polyethylene sample moving at 7

m/s is observed. Whereas, a circular laser beam of approximately the same spot area is capable of producing the weld at a maximum speed of 5 m/s for the same laser power.

7. EFFECTS OF MOISTURE CONTENT

Moisture absorption by plastic can lead to a change in some of its mechanical and physical properties, which may also effect the performance of welds. The amount of moisture pick up depends on the type of plastics, as well as the environmental conditions. Very few research works have been oriented towards the study of the influence of accumulated moisture on optical and mechanical properties of laser-welded plastics.

Kocheny et al. [24] investigated the effects of moisture content on the efficiency of laser transmission welding process and compared the weld strength of laser welded specimens to those welded by vibration, hot plate and ultrasonic welding technology at different environmental conditions. They used laser transmissible and laser absorbing grades of unfilled and 33 wt.% fiber-glass reinforced nylon 6. The samples used for the experiments that were sealed before welding, were kept into an environmental chamber where the relative humidity were maintained at 62%, were submerged into a tank of water to results in samples with 100% relative humidity. It is found that absorption of moisture in plastic have not any significant effect on the mechanical performance of the laser transmission welded parts. Similar trends are observed for the effect of moisture on optical and mechanical performance of laser welded polyamide, studied by the same research group [25]. They mentioned that the moisture is not a barrier for the laser transmission welding applications and it does not have any detrimental effect on the mechanical performance of laser welded components. They reported that, laser transmission welding technology is more efficient in the welding of wet nylon and polyamide than ultrasonic welding and gives a similar mechanical performance to linear vibration welded material. In both the studies no evidence is found relating the significance of moisture to laser energy transmission in polyamide and nylon 6. It is also observed that the samples, which are welded in the dry-as-molded condition exhibit brittle fracture, either in the weld or in the base materials. Whereas, the samples that contained increased amount of moisture exhibit ductile fracture within the weld region because moisture in thermoplastics serves as a plasticizer that reduces the material strength and increases its ductility.

8. BRIDGING THE AIR GAP

In laser transmission welding, the heat generated in the laser absorbing polymer is transported to the transparent polymer by thermal conduction. Therefore, the presence of air gap at the joining interface is a major concern for the process. If the air gap between the mating parts is excessively high, no heat transfer will take place. The presence of air gap is always disadvantageous for the functionality of the laser transmission welding process.

The gap between the mating parts occurs as a result of the poor dimensional accuracy of the parts and also due to poor clamping and joint design. For bridging the air gap, the absorptive material must be heated slowly to allow more heat to be conducted into the bulk of the material. This results in large volume expansion and the two parts to be fused. In this way,

a certain range of air gap can be bridged using the thermal expansion of the material during laser transmission welding. The acceptable air gap is highly dependent of the thermo-physical, mechanical and optical properties of the polymer and of the weld geometry. Therefore, the gap bridging capability can be optimized, by controlling the melt volume by seam width and absorption length by increasing the penetration depth, through changing the absorption coefficient of the absorbing material [4, 26].

Jansson et al. [19] reported that the weld strength decreases with increase in air gap between the parts but the maximum weld strength for different air gaps are achieved with approximately the same line energy as in experiments without an air gap. In the air gap experiments, higher weld strength is achieved with a lower welding speed. Higher irradiation time contributed to a wider and deeper weld, i.e., a larger welding volume which leads to larger gap bridging capability.

The simultaneous and quasi-simultaneous welding techniques exhibit better potential in bridging air gaps than contour welding technique as the above two techniques creates higher increase in volume of the melt than that of the latter laser transmission welding variant [27]. Jansson et al. [18] observed that quasi-simultaneous welding technique can bridge upto 0.3 mm air gap without any critical decrease in the maximum tensile strength per length of the weld. It is also noticed that the higher volume increase of polypropylene (PP) favors the gap bridging capability of the polypropylene compared to polycarbonate (PC) welded to acrylonitrile butadiene styrene/polycarbonate (ABS+PC) alloy. A relatively high laser power also creates more molten material and thereby, a wider and deeper weld. This leads to a better gap bridging capability.

9. CLEARWELD[®] - PLASTIC WELDING TECHNOLOGY

The Clearweld[®] process is invented, and has been patented by TWI [28]. It is being commercialized by Gentex Corporation and became commercially available in 2002. This technology is used to join colored and uncolored, but optically transparent thermoplastics. It can produce high quality weld without the use of opaque materials or the addition of unwanted colors. This process produces joints almost invisible to the human eye. The Clearweld[®] process uses an almost colorless dye made up of near infrared absorbing materials dissolved in a variety of solvents that are used to transport the absorber to the joint interface. These dyes absorb the laser light, and through an exothermic reaction, convert the energy to heat, which melts the joining interface to make the weld. The infrared absorbing medium is either printed or painted onto one surface of the joint, encompassed into the bulk plastic, or produce in the form of a film that can be inserted into the joint. These dyes have slight green tint before welding for locating the weld zone but after laser welding with the optimized processing condition it becomes colorless, similar to the sample presented in Figure 7. The Clearweld[®] dye materials have a maximum absorption range between 0.94 to 1.064 μm . Both diode and Nd:YAG lasers can be used for this process. Clearweld[®] process depends upon accurate and repeatable application of the near infrared absorbing layer at the localized joint interface, compatibility of the absorbing material with substrate material, process parameters and joint design. This process is especially suitable where the appearance of product is important. Applications of Clearweld[®] process can be found through the plastic

welding industry including medical devices, packaging, automotive components, consumer products, textiles and electronics [29-32].

The Clearweld® process has gained a great interest among the researchers in studying various aspects of this novel plastic joining technique. Jones et al. [30] successfully welded two clear sheets of acrylic (polymethylmethacrylate) of 3 mm thickness using Clearweld® technique with a Nd:YAG laser. A 12 µm methylmethacrylate film containing approximately 0.02% infrared absorbing dye is placed at the interface. Both pieces are clamped together and welded with an applied power of 100 W at a welding speed of 8 mm/s. The laser beam used is of 6 mm diameter, larger than the film strip width of 5 mm. The maximum failure force achieved is 50 N per mm of the weld. The failures are occurred at the parent material near the weld, and implies stronger weld. The appearance of the weld is found as clear as the parent material and has a very little effect of residual color. Hoult and Burrell [33] studied the effects of diode laser wavelength on the Clearweld® process. Clear acrylic samples are welded to each other using a range of different infrared absorbing dye concentrations. It is found that combination of diode laser and infrared absorbing inks can produced satisfactory full strength joints over a wide range of laser parameters. Higher dye concentration absorbs more energy and produce stronger joints in shorter times when all other variables remain constant. For the particular type of inks used in this study, the longer wavelength 0.977 µm absorbed laser energy most efficiently under this relevant laser irradiation condition. Hertly et al. [34] studied the Clearweld® technique with polycarbonate, polyamide and polystyrene samples. They found that the Clearweld® technique is capable of producing not only an aesthetically but also mechanically sound weld. Woosman and Burrell [31] studied the effects of welding parameters on strength of the Clearweld®ed thermoplastic parts using a methoxy-propanol based ink for welding polypropylene and an ethanol based ink for welding acrylic in a butt joint configuration. They reported that the strongest welds are achieved with highest powers (250-300 W) and clamp pressure (4.5 MPa), used for the study. They concluded that the users can choose to work with a mid-range power because the weld strengths are found less sensitive to the variation of welding speed. Kagan and Woosman [35] studied the efficiency of Clearweld® technology for various non-reinforced and short-fiber reinforced nylons. The Clearweld® process is performed using a diode laser of input power 150 watt and wavelength 0.94 µm with a rectangular beam. The beam size is varied between 2.5 and 4.5 mm to produce the optimum energy density based on each material while keeping laser power and speed as constant. They used an optimized clamping pressure of the range 1.0-1.2 MPa. The efficiency is determined by the ratio of tensile strength of T-joint Clearweld®ed plastic materials to the tensile strength of parent plastic materials. The analysis shows that the Clearweld® technology is highly efficient for use with various transparent nylon grades. The tensile strength of the T-type butt joints is found similar to the result achieved for nylon with other advanced plastic joining methods such as linear vibration, orbital vibration, hot plate and regular laser transmission technologies.



Figure 7. Welding of two transparent plastics using Clearweld[®] technology (Reproduced by permission TWI Ltd)

Woosman et al. [36] studied the Clearweld[®] process for different polymers using the Clearweld[®] resin additives. The laser absorbing additives are compounded with the polymer to render the transparent plastic laser absorbing. The compatibility of the additive to the specific plastic is an important issue for this type of application. Burrell et al. [37] used Clearweld[®] technique to weld polycarbonate parts with Clearweld[®] resin additives. They conducted a set of experiments to optimize welding parameters based on additive concentrations. The additive concentration and laser power intensity have shown the most influence on the weld strength. Haberstroh and Hoffman [38] used two different types of commercially available resin additives (Clearweld[®] and Lumogen[®]) in welding of transparent micro plastic parts for application in micro-technology. Polycarbonate samples containing additive concentration of 0.01 wt.% are used for this study. A higher concentration of additives is not applied, as it obstructs visible transparency of the transparent polycarbonate parts. They observed that that Lumogen[®] leads to a rather high absorbance of more than 90%, while Clearweld[®] additives results in a lower absorbance of about 50%. The maximum absorption for Lumogen[®] is found in a wavelength range of 0.78 μm to 0.82 μm , whereas Clearweld[®] is more effective at 0.94 μm wavelength radiation. They concluded that these additives are not suitable for the application in micro-parts due to the pronounced volume absorption caused by these additives, and suggested that an appropriate laser absorbing thin intermediate layer with high absorbance can be used to avoid such problems.

10. ADVANTAGES AND LIMITATION

Laser transmission welding has several advantages over other conventional plastic welding processes, as follows:

1. Non-contact, non-contaminant, flexible joining process,
2. Produces optically and qualitatively high-grade joints,
3. Low thermal and mechanical stress,
4. Localized heat affected zone,
5. Absence of vibration of the parts,

6. No particulate development,
7. No flash or marks on outer surface of the material,
8. Minimal limitations of part geometry and the size to be joined,
9. Amorphous, crystalline and thermoplastic elastomers are weldable,
10. Ability to weld dissimilar plastics,
11. Capable of gas-tight, hermetic sealing,
12. Equipped to weld 3D joint lines,
13. High processing rates,
14. Quick changeover,
15. High process repeatability,
16. No tool wear,
17. Low tooling costs, and
18. High integration capabilities and potential of automation.

Laser transmission welding process has some process limitations as well:

1. It depends too much on materials' optical properties. The part at the top must be laser transparent and the bottom one should be laser absorbent,
2. When welding two transparent materials, an IR absorbing intermediate layer is required to be placed at the weld interface. This increases cost per unit,
3. High equipment cost,
4. Intimate contact required between mating parts, and
5. Part thickness limitation for crystalline materials.

11. APPLICATIONS

Laser transmission welding of polymer is at the evolving stage for wide industrial applications. However, several applications have already been adapted into industrial production. At present, many industries are investigating this process to replace conventional plastic joining processes. Laser transmission welding is now used in a wide range of application areas, including medical devices, automotive components, electrical and electronic devices, packaging, light and displays, house hold goods, and textiles industries.

A number of applications are there in automotive industries for welding automotive parts such as connectors, front and rear lights assemblies, bumpers, pump and turbine housings, liquid containers, dash board components, remote door keys, flood lights, automotive intake manifolds, etc. Laser welding technology is successfully applied for contour welding of mobile phone cover and cosmetic packages. The use of laser transmission welding continues to expand to other applications such as sensors and switches in the electronic industries; biomedical sensors, dialysis components and medical packaging fabrication in the medical industries; plastic window, doors, dowels in the building trade; plastic dishes and shavers in the house hold good industries; and product packaging and air tight sealing in the packaging industries [39].

Applications of laser welding of polymers that have been advertised by Laserline GmbH include automatic gear-shift sensor, gear-shift console, pneumatic pump module, filter housing, air flow sensor, car key, electronic housings, automatic gear box, mats from plastic

materials, liquid pouring device, weld of foam on plastics, welding of hydraulic tanks etc. Some of these are presented in Figure 8.

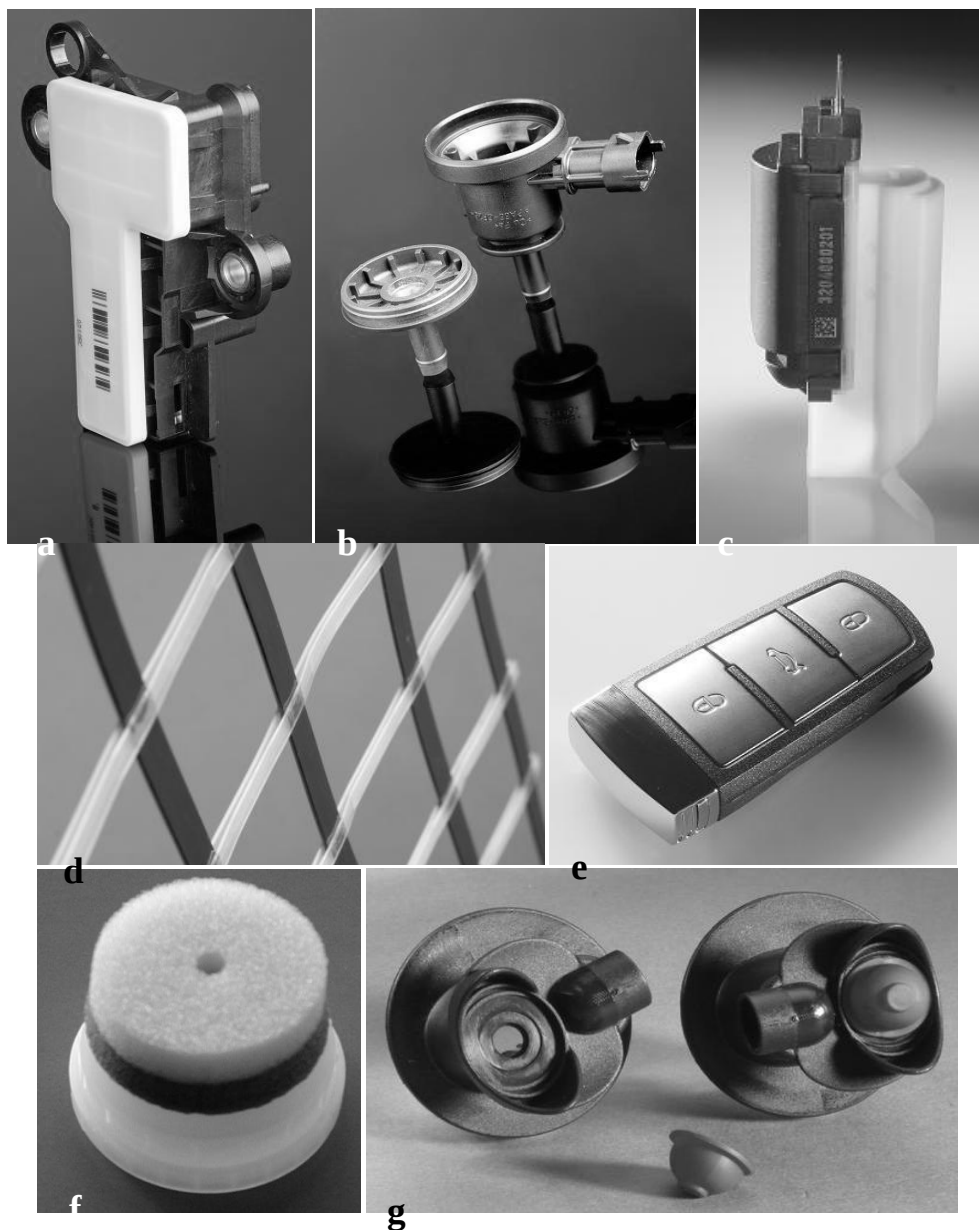


Figure 8. Application of laser transmission welding in (a) automatic gear-shift sensor, (b) filter housing, (c) air flow sensor, (d) car key, (e) mats from plastic materials, (f) liquid pouring device and (g) weld of foam on plastics (reproduced by permission Laserline GmbH)



Figure 9. Laser transmission welded micro-fluidic device (reproduced by permission Leister Technologies)

There is also continuous growth in the use of laser transmission welding technique in the manufacture of micro parts such as joining of micro-fluidic devices (Figure 9). The electronics and medical devices industry require of micro joining of dissimilar materials for the majority of their applications. In joining biomedical products, the joining process should not make use of any third material, which is not biocompatible. The laser transmission welding process meets this condition. Being a non-contact process, the laser transmission welding does not lead to contamination at the functional areas of the bio-medical products. Laser transmission welding process is now used to join biomedical implants and for encapsulation of biomedical devices due to its high precision and biocompatibility property. Laser welding of metal to plastic, ceramics to plastic and glass to plastic are also successfully demonstrated [40-42].

12. CONCLUSION

Laser transmission welding is a novel and promising technology for many industries, those involved the joining of plastics. Laser sources of 0.8-1.1 μm wavelength are generally used for the laser transmission welding process, as the plastics have a high transmittance at this wavelength range. To date, the three main types of industrial lasers namely, Nd:YAG, diode and fiber lasers have been used for laser transmission welding of plastics. Advantages of diode laser have contributed to a cost effective welding alternative to traditional plastic welding techniques, which, significantly increased the interest for applying laser in welding of plastics.

The efficiency of laser transmission welding process strongly depends on the optical properties of the plastic parts to be joined and the types of laser used. The basic composition of polymer matrix, colorants and additives affect laser energy absorption, reflection and transmission and finally to the mechanical performance of the weld. Thickness of plastic part has also influence in optical properties, especially for semi-crystalline materials. The most important process parameters for laser transmission welding process are laser power density, irradiation time and clamping pressure. The temperature field inside the weld during welding

can be controlled with these process parameters. Moisture content in plastic part does not create any difficulties for the laser transmission welding applications and neither it has not any detrimental effect on the mechanical performance of laser welded components. An allowable air gap between the parts is very much dependent of the thermo-physical, mechanical and optical properties of the polymer and of the weld geometry. A certain range of air gap can be bridged using the thermal expansion of the material during laser transmission welding, and by selecting the suitable process variant and parameters. The Clearweld® technology is the latest addition in the field. This innovative technology is capable of joining colored and uncolored, but optically transparent thermoplastics without using of opaque materials or the addition of any unwanted colors.

The laser transmission welding process offers several process advantages over the other conventional plastic joining technologies. The application of laser transmission welding is expanding rapidly. A number of applications have already been shaped into industrial production. The process is now successfully applied for welding of plastics to metal, ceramics and glasses. However, extensive research work is necessary to explore various aspects of this relatively newer joining process for plastics. Future research works may be directed towards the development of newer process friendly materials and pigments, newer application strategies and optimization of the process. This will lead to more effective utilization of the process yielding better weld quality.

REFERENCES

- [1] Silvus, H. J. Jr. & Wachtell, S. (1970). Perforating, welding, and cutting plastic films with a continuous CO₂ laser. Pennsylvania State University, *Engineering Proceedings*, 88-97.
- [2] Nakamata, H. (1987). Process for joining different kinds of synthetic resins. *US Patent*, 4636609.
- [3] Coelho, J. M. P., Abreu, M. A. & Pires, M. C. (2000). High-speed laser welding of plastic films. *Optics and Lasers in Engineering*, vol. 34, 385-395.
- [4] Bachmann, F. G. & Russek, U. A. (2002). Laser welding of polymers using high power diode lasers. *Proceedings of SPIE*, vol. 4637, 505-518.
- [5] Bonten, C. & Tüchert, C. (2002). Welding of plastics-Introduction into heating by radiation. *Journal of Reinforced Plastics and Composites*, vol.21(8), 699-710.
- [6] Baylis, B. (2002). Welding thermoplastic elastomers to polypropylene with a diode laser. *Proceedings of the 21st International Congress on Applications of Lasers & Electro-Optics*, Scottsdale, Arizona, USA.
- [7] Bryden, B. (2000). High power diode laser transmission welding of plastics. *Assembly Automation*, vol. 20(2), 136-139.
- [8] Kagan, V. A. & Bray, R. G. (2001). Advantages and limitations of laser welding technology for semi-crystalline reinforced plastic. *Proceedings of the 20th International Congress on Applications of Lasers & Electro-Optics*, Jacksonville, Florida, USA.
- [9] Kagan, V. A., Bray, R. G. & Kuhn, W. P. (2002). Laser transmission welding of semi-crystalline thermoplastics: part I: Optical characterization of nylon based plastics. *Journal of Reinforced Plastics and Composites*, vol. 21(12), 1101-1122.

-
- [10] Kagan, V. A., Bray, R. & Chambers, A. (2003). Forward to better understanding of optical characterization and development of colored polyamides for the infra-red/laser welding: part I - Efficiency of polyamides for infra-red welding. *Journal of Reinforced Plastics and Composites*, vol. 22(6), 533-547.
- [11] Kagan, V. A. & Pinho, G. P. (2004). Laser transmission welding of semicrystalline thermoplastics – part II: Analysis of mechanical performance of welded nylon. *Journal of Reinforced Plastics and Composites*, vol. 23(1), 95-107.
- [12] Grewell, D., Rooney, P. & Kagan, V. A. (2004). Relationship between optical properties and optimized processing parameters for through-transmission laser welding of thermoplastics. *Journal of Reinforced Plastics and Composites*, vol. 23(3), 239-247.
- [13] Kagan, V. A., Chambers, A. & Bray, R. (2003). Forward to better understanding of optical characterization and development of colored polyamides for the infra-red/laser welding, part II – Family of colored polyamides. *Journal of Reinforced Plastics and Composites*, vol. 22(7), 593-603.
- [14] Haberstroh, E., Hoffmann, W. M., Poprawe, R. & Sari, F. (2006). 3 laser transmission joining in microtechnology. *Microsystems Technology*, vol. 12, 632-639.
- [15] Potente, H., Korte, J. & Becker, F. (1999). Laser transmission welding of thermoplastics: analysis of heating phase. *Journal of Reinforced Plastics and Composites*, vol. 18(10), 914-920.
- [16] Abed, S., Laurens, P., Carrétéro, C., Deschamps, J. R. & Duval, C. (2001). Diode laser welding of polymers: microstructures of the welded zones for polypropylene. *Proceedings of the 20th International Congress on Applications of Lasers & Electro-Optics*, Jacksonville, Florida, USA.
- [17] Haberstroh, E. & Luetzeler, R. (2001). Influence of carbon black pigmentation on the laser beam welding of plastics micro parts. *Journal of Polymer Engineering*, vol. 21(2-3), 119-129.
- [18] Jansson, A., Kouvo, S. & Kujanpää, V. (2004). Preliminary investigations of laser welding of plastics in massproduction. *Proceedings of the 23rd International Congress on Applications of Lasers and Electro-Optics*, San Francisco, California, USA.
- [19] Jansson, A., Kouvo, S., Salminen, A. & Kujanpää, V. (2003). The effect of parameters on laser transmission welding of polymers. *Proceedings of the 22nd International Congress on Applications of Lasers & Electro-Optics*, Jacksonville, Florida, USA.
- [20] Acherjee, B., Kuar, A.S., Mitra, S. and Misra, D. (2010) Selection of process parameters for optimizing the weld strength in laser transmission welding of acrylics, *Proc. IMechE Part B: Journal of Engineering Manufacture*, vol. 224, in press, doi: 10.1243/09544054JEM1
- [21] Acherjee, B., Misra, D., Bose, D. & Venkadeshwaran, K. (2009). Prediction of weld strength and seam width for laser transmission welding of thermoplastic using response surface methodology. *Optics & Laser Technology*, vol. 41(8), 956-967.
- [22] Prabhakaran, R., Kontopoulou, M., Zak, G., Bates, P. J. & Baylis, B. K. (2006). Contour laser – Laser-transmission welding of glass reinforced nylon 6. *Journal of Thermoplastic Composite Materials*, vol.19, 427-439.
- [23] Douglass, D. M. & Wu, C. Y. (2003). Laser welding of polyolefin elastomers to thermoplastic polyolefin. *Proceedings of the 22nd International Congress on Applications of Lasers & Electro-Optics*, Jacksonville, Florida, USA.

- [24] Kocheny, S. A., Kagan, V. A. & Macur, J. (2004). Through-transmission laser welding of nylon – Breaking the moisture barrier. *ANTEC 2004 Conference proceedings*, Chicago, IL, USA.
- [25] Kagan, V. A., Kocheny, S. A. & Macur, J. E. (2005). Moisture effects on mechanical performance of laser-welded polyamide. *Journal of Reinforced Plastics and Composites*, vol. 24(11), 1213-1224.
- [26] Van de Ven, J. D. & Erdman, A. G. (2007). Bridging gaps in laser transmission welding of thermoplastics. *Journal of Manufacturing Science and Engineering*, vol. 129, 1011-1018.
- [27] Jansson, A., Kouvo, S. & Kujanpää, V. (2005). Quasi-simultaneous laser welding of polymers - the process and applications for mass-production. *Proceedings of the 24th International Congress on Applications of Lasers & Electro-Optics*, Miami, Florida, USA.
- [28] Jones, I. A. & Wise, R. J. (2003). Welding method. *European patent*, 1117502.
- [29] Jones, I. A., Taylor, N. S., Sallavanti, R. & Griffiths, J. (2000). Use of infrared dyes for transmission laser welding of plastics. *ANTEC 2000 Conference proceedings*, Orlando, USA.
- [30] Jones, I. A., Hilton, P. A., Sallavanti, R. & Griffiths, J. (1999). Use of infrared dyes for transmission laser welding of plastics. *Proceedings of the 18th International Congress on Applications of Lasers & Electro-Optics*, San Diego, CA, USA.
- [31] Woosman, N. M. & Burrell, M. M. (2003). A study of the effect of weld parameters on strengths of ClearweldTM thermoplastics. *Proceedings of the 22nd International Congress on Applications of Lasers & Electro-Optics*, Jacksonville, Florida, USA.
- [32] *Clearweld plastics*. <http://clearweld.com>, accessed on April 20, 2010.
- [33] Hoult, A. P. & Burrell, M. (2002). The effect of diode laser wavelength on the clearweldTM welding process. *Proceedings of the 21st International Congress on Applications of Lasers & Electro-Optics*, Scottsdale, Arizona, USA.
- [34] Hartley, S. & Sallavanti, R. A. (2003). ClearweldTM laser transmission welding of thermoplastic polymers: light transmission and color considerations. *Proceedings of SPIE*, vol. 4830, 63-68.
- [35] Kagan, V. A. & Woosman, N. M. (2004). Efficiency of clearwelding technology for polyamides. *Journal of Reinforced Plastics and Composites*, vol. 23(4), 351-359.
- [36] Woosman, N., Curtis, M., Cawley, W. & Verespy, J. (2005). ClearweldTM resins: alternative options for TTIR clearwelds. *ANTEC 2005 Conference proceedings*, Boston, MA, USA.
- [37] Burrell, M. M., Cawley, W. H. & Verespy, J. P. (2007). Design of experiment to optimize absorber in resin welding parameters. *ANTEC 2007 Conference proceedings*, Cincinnati, OH, USA.
- [38] Haberstroh, E. & Hoffmann, W. M. (2007). Laser transmission welding of transparent plastics parts in micro technology. *3rd International Conference on Multi-Material Micro Manufacture (4M 2007)*, Borovets, Bulgaria.
- [39] Russek, U. A., Poggel, M., Otto, G. & Koeppe, A. (2003). Advances in laser beam welding of polymers and automotive prospects. *Proceedings of the 9th International Conference: TPOs in Automotive*, Maastricht, The Netherlands.
- [40] Katayama, S. & Kawahito, Y. (2008). Laser direct joining of metal and plastic. *Scripta Materialia*, vol. 59, 1247-1250.

-
- [41] Kawahito, Y., Niwa, Y. & Katayama, S. (2009). Laser Direct Joining of Ceramic and Engineering Plastic. *Proceedings of the 28th International Congress on Applications of Lasers & Electro-Optics*, Orlando, FL., USA.
 - [42] Sultana, T., Georgiev, G. L., Baird, R. J., Auner, G. W., Newaz, G., Patwa, R. & Herfurth, H. J. (2009). Study of two different thin film coating methods in transmission laser micro-joining of thin Ti-film coated glass and polyimide for biomedical applications. *Journal of Mechanical Behavior of Biomedical materials*, vol. 2, 237-242.

Chapter 8

EFFECT OF IN SITU REACTION ON THE PROPERTY OF PULSED Nd:YAG LASER WELDING SiCp/A356

Kelvii Wei Guo* and Hon Yuen Tam

Department of Manufacturing Engineering and Engineering Management,
City University of Hong Kong, Kowloon Tong, Kowloon, Hong Kong

ABSTRACT

The effect of in situ reaction on the properties of pulsed Nd:YAG laser welded joints of particle reinforcement aluminum matrix composite SiCp/A356 with Ti filler was studied, and its corresponding temperature field was simulated. Results shows that in situ reaction during the laser welding restrains the pernicious Al_4C_3 forming in the welded joints effectively. At the same time, the in situ formed TiC phase distributes uniformly in the weld, and the tensile strength of welded joints is improved distinctly. Furthermore simulation results illustrate that in addition to the lower heat-input into the substrate because of Ti melting, in situ reaction as an endothermic reaction decreases the heat-input further, and its temperature field distributes more smoothly with in situ reaction than that of laser welding directly. Also, the succedent fatigue test shows the antifatigue property of welded joints with in situ reaction is superior to that of traditional laser welding. It demonstrates that particle reinforcement aluminum matrix composite SiCp/A356 was successfully welded by pulsed Nd:YAG laser with in situ reaction.

Keywords: In situ reaction; Nd:YAG laser; SiCp/A356; Ti; Simulation; Fatigue.

1. INTRODUCTION

The high specific strength, good wear resistance and corrosion resistance of aluminum matrix composites (AMCs) have led to a number of industrial applications [1–5]. For

* E-mail address: guoweichinese@yahoo.com; metamhy@cityu.edu.hk

example, AMCs are widely used in automobile, aerospace industries, structural components, and heat and wear resistant parts such as automotive brake discs. Owing to the typical characteristics of production methods, the distribution of the reinforcement in stir-cast AMCs is generally inhomogeneous [1, 5]. Furthermore, the ceramic reinforcement may be in the form of particles, short fibers, or whiskers [5, 6]. The discontinuous nature of the reinforcement creates several problems in imparting strength and quality to weld joints. Although there are several welding techniques currently available for joining AMCs [7–15], there still exist quality problems due to the factors such as (i) reinforcement distribution in the weld [16–18]; (ii) Interface between particle reinforcements and aluminum matrix [19–20].

This work studies the technique of welding the stir-cast aluminum matrix composite $\text{SiC}_p/\text{A356}$ by Nd:YAG laser with pure titanium as filler. The effect of in situ reaction on the properties of welded joints has been investigated using Scanning Electron Microscope (SEM+EDX), Transmission Electron Microscope (TEM) and X-ray diffraction (XRD) and simulated by the finite element method (FEM).

2. EXPERIMENTAL MATERIAL AND PROCESS

2.1. Experimental Material

Stir-cast $\text{SiC}_p/\text{A356}$ aluminum matrix composite (AMC), reinforced with 20 % volume fraction SiC particle of 12 μm mean size, was used as the welding specimens. The tensile strength of the specimen was 240 MPa. Figure 1 shows the microstructure of the sample and Table 1 lists the chemical composition of the matrix alloy. Pure titanium was used as the filler metal.

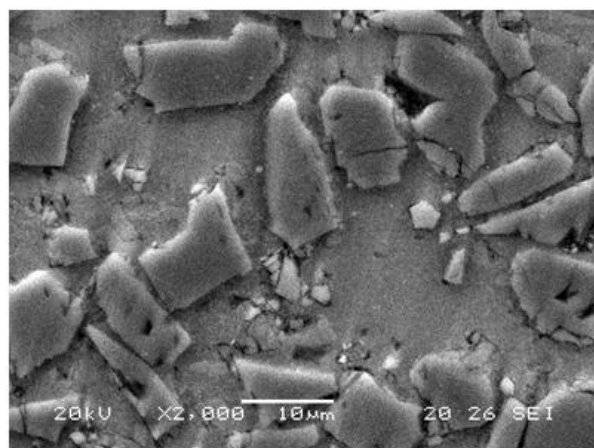


Figure 1 Microstructure of $\text{SiC}_p/\text{A356}$ aluminum matrix composite

Table 1. Composition of A356

Composition (wt %)			
Si	Mg	Ti	Al
6.5~7.5	0.3~0.5	0.08~0.2	Bal.

2.2. Experimental Process

The stir-cast AMC specimens were individually wire-cut to $3 \text{ mm} \times 10 \text{ mm} \times 35 \text{ mm}$ size. The quench-hardened layer induced by wire-cut and the oxide on the surfaces of specimens were removed by polishing on 400 # ($35 \text{ }\mu\text{m}$ in average) emery cloth. The pure titanium filler was then machined to $3 \text{ mm} \times 10 \text{ mm}$ size with thicknesses of 0.15, 0.3, 0.45, 0.5, 0.6 and 0.75 mm, respectively. The specimens were ultrasonically cleaned in acetone at 28-34 Hz frequency for 5 minutes, then carefully pure ethyl alcohol rinsed and blow dried before welding. Finally, the specimens were mounted into a clamping device on the platform of a GSI Lumonics Model JK702H Nd:YAG TEM₀₀ mode laser system.

A repeated cleaning process was used for machined titanium, and the titanium filler was carefully sandwiched between the two composite specimens in the clamp. Thereafter, specimens were welded immediately by the Nd:YAG laser with wavelength of $1.06 \text{ }\mu\text{m}$, defocused distance of 10 mm so as to give a focus spot diameter of approximately 1.26 mm on the samples.

Tensile strength of the joint was measured on a MTS Alliance RT/30 electron-mechanical material testing machine with a straining velocity of 0.5 mm/min. The cross-section of welded joints was wire-cut for Optical Microscopy (OM), Scanning Electron Microscopy (SEM) and Transmission Electron Microscopy (TEM). SEM was used to analyze the microstructure at the weld joints and the fractured tensile test-pieces of the joints. Optical microscope was used for observing the structure of a large area. TEM and Energy Dispersive X-ray analysis (EDX) were used to analyze the interface between the newly-formed phases and the matrix, the distribution of chemical elements and spectra at the joints. Moreover, the Nd:YAG laser with similar setting conditions and processing parameters was also used to weld the AMC specimens without filler.

3. RESULTS AND DISCUSSION

3.1. Microstructures and Properties of Welded Joints

The microstructure (Figure 2) of the traditional Nd:YAG laser weld without filler shows that acicular Al_4C_3 with various sizes is formed in the weld, which led to a lower joint tensile strength (Figure 3) of 91 MPa (about 37.9 % parent AMC). The corresponding fracture surface is shown in Figure 4. It shows in addition to some bare reinforcement particles (SiC) scattering on the fracture surface, a lot of Al_4C_3 is also distributed on the fracture surface. It illustrates that the reinforcement particles have not been perfectly wet. At the same time, the reinforcement particles lose its advantage effect instead of being as newly-formed harmful phase Al_4C_3 , resulted in decreasing the tensile strength of welded joints.

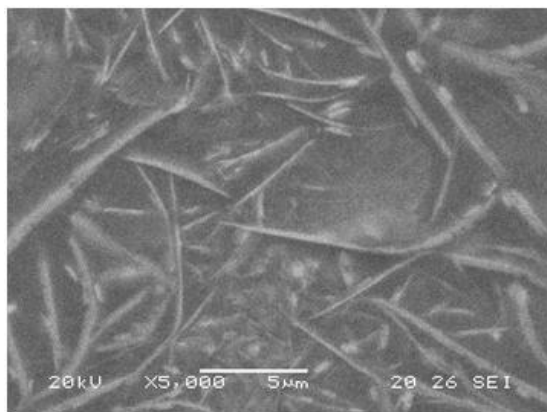


Figure 2. Microstructure of the weld without Ti filler

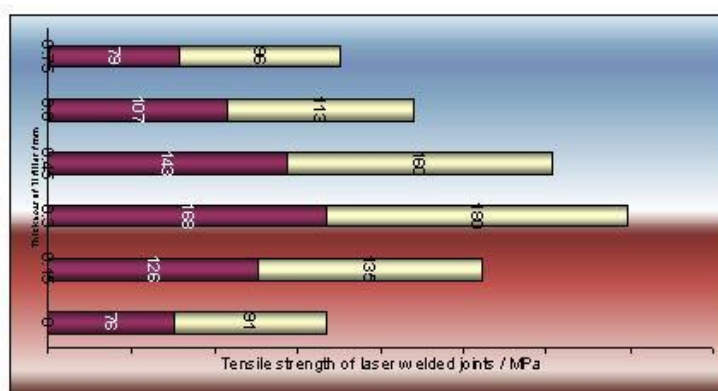


Figure 3. Tensile strength of laser welded joints with various Ti filler thicknesses

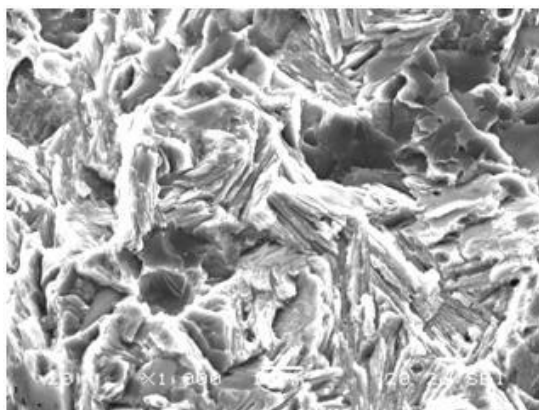


Figure 4. Fractograph of the laser welded joint without Ti filler

The microstructure of the in situ reinforced AMC with 0.3 mm thick Ti filler is shown in Figure 5. This figure shows a uniform distribution of in situ reinforcements, complete fusion and absence of Al_4C_3 . These features result in higher tensile strength (Figure 3) of the joint. The reinforcement particles are distributed more uniformly than in parent composite (cf.

Figures 1 and 5) and this improves the properties of welded joints as the newly-formed in situ reinforcement particles (Figure 5) replace the initial reinforcement particles (Figure 1). The dimples in the fracture surface (Figure 6) suggest that: (i) the newly-formed reinforcement particles have been perfectly wet [19-20]; and (ii) the harmful composite structure of the initial welding viz. reinforcement/Ti/reinforcement has been changed to reinforcement/matrix/reinforcement. XRD pattern of the fracture surface (Figure 7) of the weld joint does not reveal any harmful and brittle phases such as Al_4C_3 . According to the intensity spectra shown in Figure 7, the newly-formed reinforcement particle in the weld is identified as TiC.

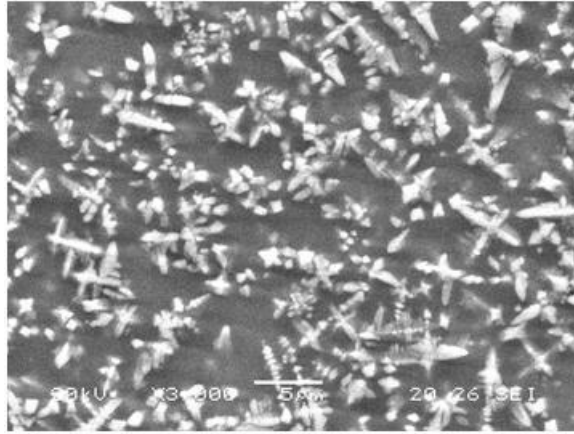


Figure 5. Microstructure of in situ reinforcement by laser welding with 0.3 mm thick Ti filler

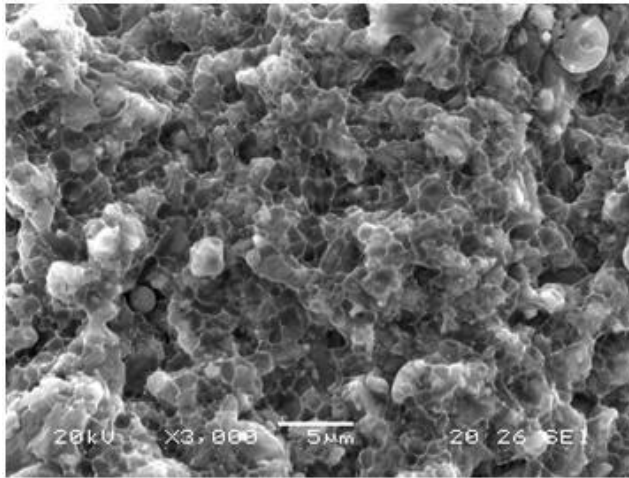


Figure 6. Fractograph of the laser welded joint with 0.3 mm thick Ti filler

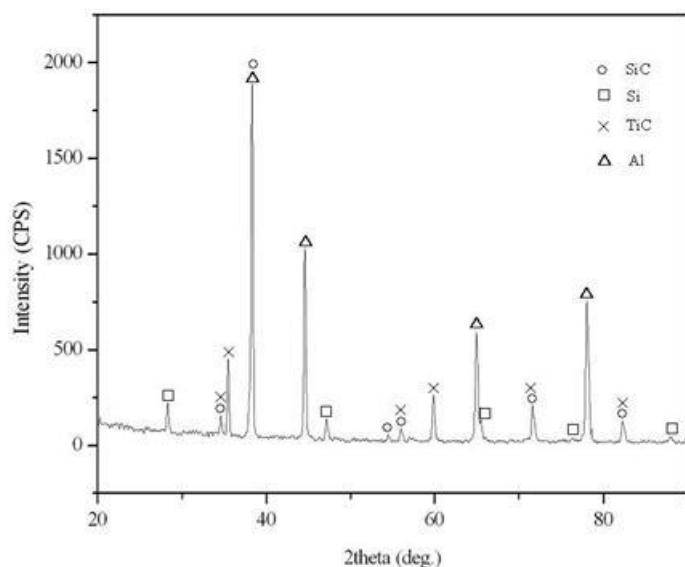


Figure 7. XRD pattern of the fracture surface for laser welding with 0.3 mm thick Ti filler



Figure 8. Macro-structure of the laser welded joint with 0.3 mm thick Ti filler

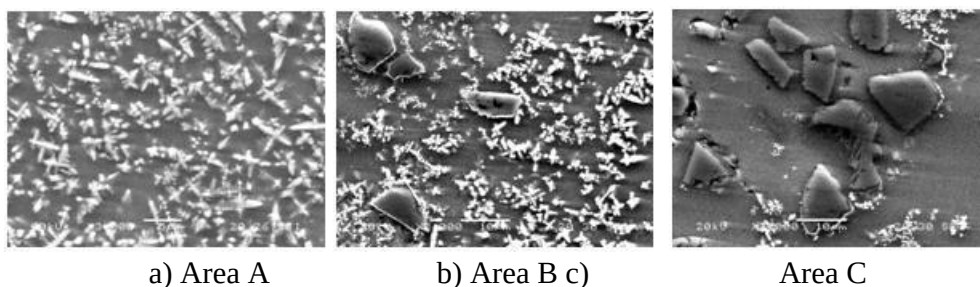


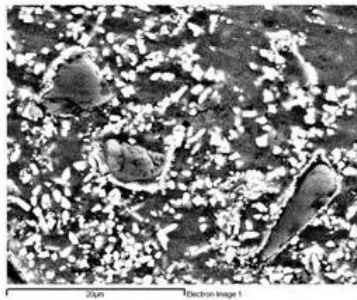
Figure 9. Microstructures of the different areas in the laser weld with 0.3 mm thick Ti filler

Figure 8 shows the macro-structure of welded joint with Ti filler. Basically, the weld consists of three main areas, namely: the in situ reinforcement area A, the two transitional areas B and C, and the reinforcement-denuded area D. Their individual microstructures are shown in Figure 9. The microstructures indicated that the initial reinforcement SiC particles were completely replaced by the newly-formed in situ reinforcement TiC particles that mainly resulted in the formation of the area A (Figure 9a). In area B, the newly-formed TiC particles and the SiC particles coexist (Figure 9b). In area C, little newly-formed TiC particles are found (Figure 9c). In area D, only SiC particles exist (Figure 1). It was found that Al_4C_3 has

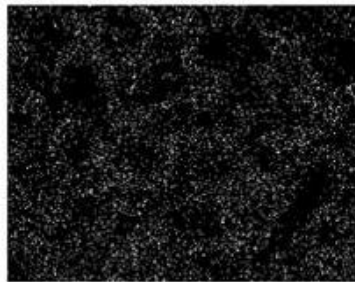
been effectively eliminated in the welded area. Hence, the properties of the welded joints improve markedly and their achievable tensile strength is up to 180 MPa (Figure 3) that is about 75 % of the strength of SiC_p/A356.

3.2. Element Distribution in the Transition Area

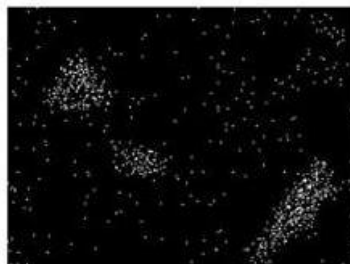
Figure 10 illustrates the element distribution of the area B in the weld as shown in Figure 8 and Figure 9b. It shows that the newly-formed in situ reinforcement particles surround the SiC particles which offers a high density area for the nucleation of in situ TiC. During welding, due to the temperature gradient and surface tension in the weld pool, convection can occur. Furthermore, under the effect of plasma, the weld pool will be stirred intensively. Consequently, the stirring effect in the weld pool by laser irradiation will promote the TiC formation (cf. Figures 10b and 10c) by the following reaction:



a) Micrograph of the area B

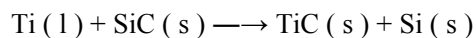


b) Ti element surface distribution



c) Si element surface distribution

Figure 10. Element distribution of B area in the weld



The free energy required to form TiC is much lower than that for Al_4C_3 when the reaction temperature is above 800 °C [22-24]. The affinity between Ti and C in the Nd:YAG laser welding is therefore greater than that of Al and C. The chemical reaction between Ti and SiC in the weld pool will take precedence over the reaction between Al and SiC and thus restrain the formation of the Al_4C_3 . Meanwhile, the Si formed during the reaction is distributed in the substrate under the stirring effect of the weld pool.

3.3. Influence of Ti Filler Thickness

The microstructures of in situ reinforcement with various thicknesses (δ) of Ti filler are shown in Figure 11 and the corresponding fractographs are shown in Figure 12. The amount of the in situ formed TiC is distinctly increased with the increase in the thickness of Ti filler. Test indicates that maximum strength of welded joints (Figure 5) is achieved at Ti filler thickness of 0.3 mm (Figures 3 and 6). This is because the TiC particles are uniformly distributed in the weld and the initial irregular (mostly hexagonal shape, Figure 1) reinforcement SiC particles in the weld are no longer observed (Figures 5 and 6). Moreover, Al_4C_3 formation is restrained (Figures 5 and 9a). At the thickness of Ti filler below 0.3 mm, due to the lack of titanium, TiC particles do not form sufficiently (Figure 12a) and a number of Al_4C_3 particles form in the weld. When the thickness of Ti filler is just beyond 0.3 mm, the properties of the joints tend to become poorer again (Figure 12b). This is because the laser input energy melts the Ti filler; as a result, the substrate can not be melted efficiently to form the TiC and the temperature of weld pool decreases to some extent. Therefore, the stirring effect in the weld pool decreases and results in coarse columnar crystals and fine equiaxed crystals (Figure 12b). When the thickness of Ti filler is further increased (Figure 12c), higher laser input energy is needed to melt the titanium. The temperature of weld pool decreases, the substrate does not melt efficiently, and the effective stirring effect between the titanium and substrate is restrained. Simultaneously, the percentage of liquid Ti in the weld pool also increases. Subsequently, the weld zone forms coarser columnar crystals, as displayed in the SEM micrograph of Figure 12c, after the resolidification of the melt.

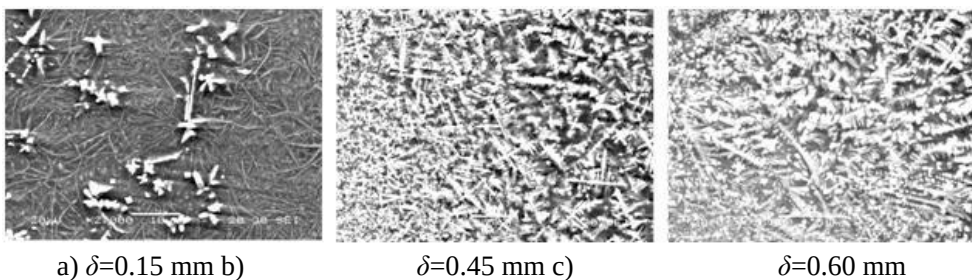


Figure 11. Microstructures of welded joints with various thicknesses of Ti filler (in A area)

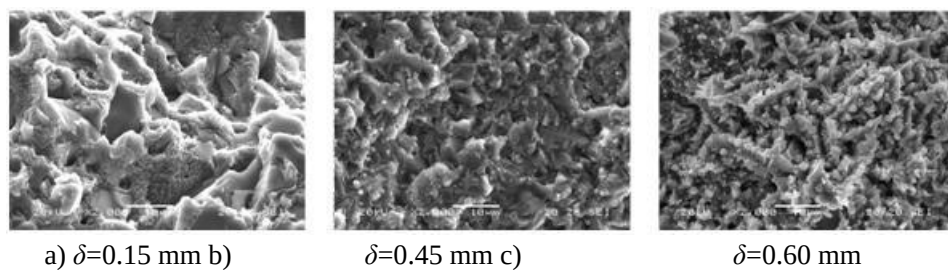


Figure 12. Fractographs of welded joints with various thicknesses of Ti filler (in A area)

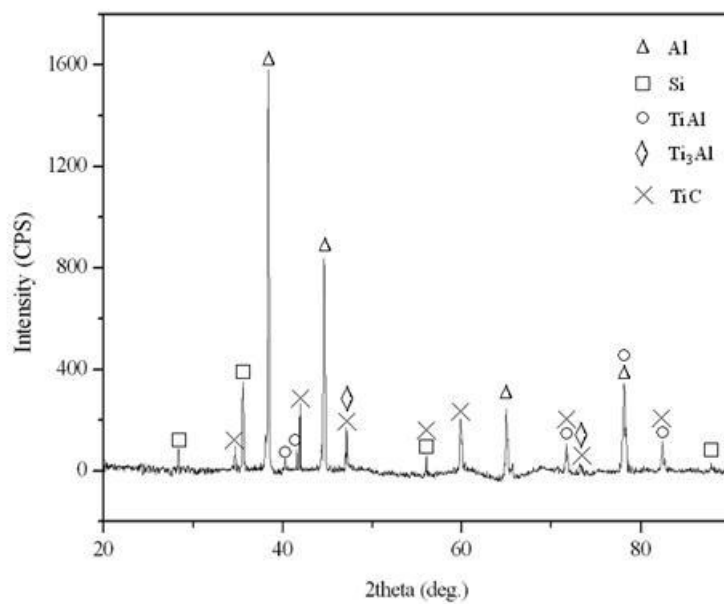


Figure 13. XRD pattern of fracture surface ($\delta=0.6$ mm)

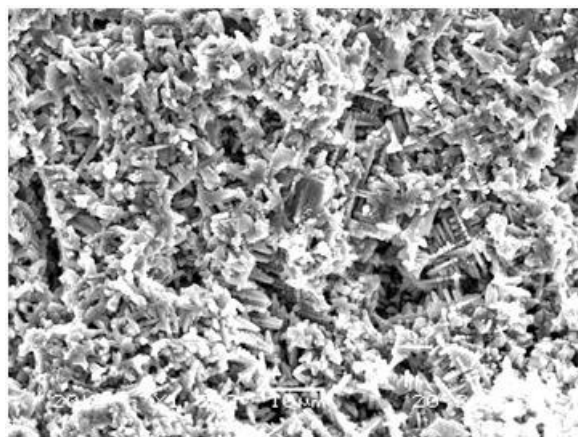


Figure 14. Columnar crystals in the laser weld with 0.6 mm thick Ti filler

From the Ti-AL binary phase diagram [25], it can be anticipated that increasing the content of Ti will lead to the formation of intermetallic compounds like TiAl and Ti_3Al , *etc.* during the Nd:YAG laser welding. As illustrated by the XRD pattern of the fracture surface of a laser weld joint with the thicker Ti filler (Figure 13), some brittle intermetallic compounds like TiAl and Ti_3Al have formed. Available literature [26] shows that TiAl and Ti_3Al are the harmful intermetallic compounds in the weld and tend to decrease the properties of welded joints. Such harmful effect may follow the chemical reaction of: $5\text{Ti}[\text{Al}(\text{l})] + 3\text{Al}[\text{l}] + \text{SiC}[\text{s}] \rightarrow \text{TiC}[\text{s}] + \text{Si}[\text{Al}(\text{l})] + \text{Al}[\text{l}] + (\text{TiAl} + \text{Ti}_3\text{Al})$. Hence, too thick of the Ti filler leads to: (i) the appearance of the large block of columnar crystals in the microstructure (Figure 14); and (ii) the newly-formed reinforcement TiC to be replaced by the melted/re-solidified Ti and subsequently only the melted/re-solidified Ti existed in the weld. Results (Figures 3, 9 and 11) indicate that there exists an optimal thickness of Ti filler in the individually set parameters in the Nd:YAG laser welding of $\text{SiC}_p/\text{A356}$. With the optimal thickness of Ti filler, the initial SiC particles distributed in the AMC will offer a highly dense nucleus area for the in situ TiC nucleation. This will effectively suppress the formation of intermetallic compounds like TiAl and Ti_3Al in the weld. Ultimately this creates favorable conditions to provide relatively superior properties of the welded joints compared to that of the conventional laser welding.

3.4. TEM of the Interface between in Situ Formed TiC and Matrix

The interface between in situ formed TiC and the matrix was analyzed by the TEM micrograph displayed in Figure 15. It shows a clear interface between the newly-formed TiC and the matrix. This suggests the occurrence of prominent in situ reaction to integrate the reinforcement particle with matrix (cf. Figures 6 and 15), and the high probability of successfully transferring load from the matrix to TiC and vice versa. It also gives indication that the aluminum matrix composite $\text{SiC}_p/\text{A356}$ will be welded satisfactorily by Nd:YAG laser.

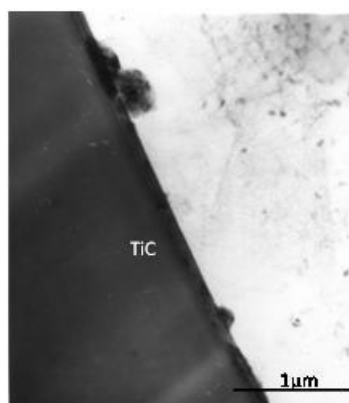


Figure 15. TEM of interface between in situ TiC reinforcement and the matrix for laser welding with 0.3 mm thick Ti filler.

4. FEM COMPUTATIONAL MODEL AND SIMULATION

4.1. Equations

Using energy balance, a differential equation can be obtained for the steady temperature distribution in a homogeneous isotropic medium, that is [27, 28]

$$\frac{\partial}{\partial x} \left(K_x \frac{\partial \theta}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_y \frac{\partial \theta}{\partial y} \right) + \frac{\partial}{\partial z} \left(K_z \frac{\partial \theta}{\partial z} \right) = -q^B \quad (1)$$

Where the boundary conditions are $\theta|_{s_1} = \theta_e$, $K_s \frac{\partial \theta}{\partial x}|_{s_2} = q^s$

For

$$\pi = \int_V \left\{ K_x \left(\frac{\partial \theta}{\partial x} \right)^2 + K_y \left(\frac{\partial \theta}{\partial y} \right)^2 + K_z \left(\frac{\partial \theta}{\partial z} \right)^2 \right\} dV - \int_V \theta q^s dV - \int_{s_2} \theta^s q^s d_s \quad (2)$$

After Eq. 2 is discrete for the element, and according to $\delta \pi = \sum_{e=1}^n \delta \pi_e = 0$, it will be obtained

$$\overline{K^K} \overline{\theta} = \overline{\theta_s} + \overline{\theta_B} - \overline{C} \overline{\theta} + \overline{K^c} (\overline{\theta_c} - \overline{\theta^s}) + \overline{K^r} (\overline{\theta_r} - \overline{\theta^s}) \quad (3)$$

where S : isothermal boundary, B : the heat input, c : the conductive and r : the irradiative.

4.2. Hypothesis and Mesh

Based on the situations during the laser welding and mainly focused on the temperature distribution, it is supposed that the laser resource is considered as a Gaussian distribution. Also, on the basis of specimen size wire-cut, the calculating size is set as 25 mm (x) $\times$ 20 mm (y) $\times$ 3 mm (z), the schematic of its finite element (FE) mesh is shown in Figure 16. Moreover, Ti filler is considered as a section of the substrate with the different properties to ignore the effect of gap between the Ti filler and the substrate.

4.3. Temperature Distribution

The simulated results are shown in Figure 17, Figure 19 to Figure 22. It shows that the temperature without Ti filler is same as the traditional laser welding. Simultaneously, due to the heat input into the substrates directly, without the additional heat resource for melting Ti filler, the peak of temperature (heat input) is relatively higher to form the weld (Figures 2 and

17). As a result, increasing the heat input into the substrate will decrease the tensile strength of the welded joint and wide the heat affected zone (HAZ) resulted in lower properties in the succedent practical applications (Figures 3 and 18). Furthermore, a large amount of coarser acicular Al_4C_3 distributes in the fracture surface as shown in Figure 18 which decreases the tensile strength of the welded joints seriously.

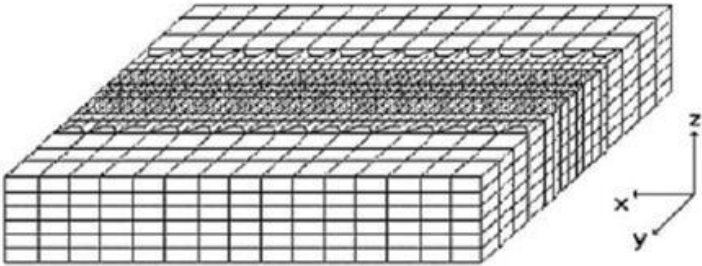


Figure 16. FE mesh for 3D numerical analysis

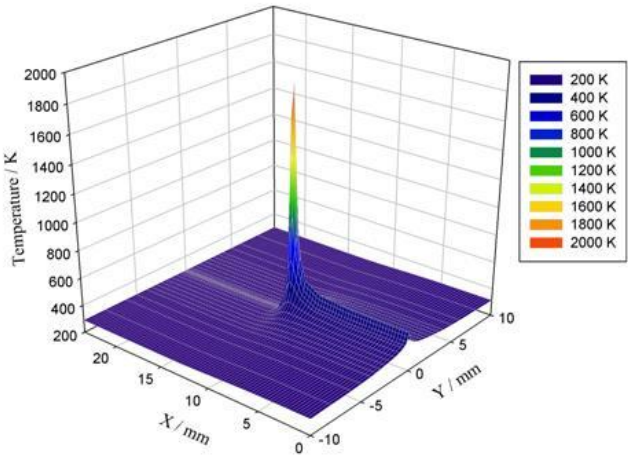


Figure 17. Temperature distribution without Ti filler

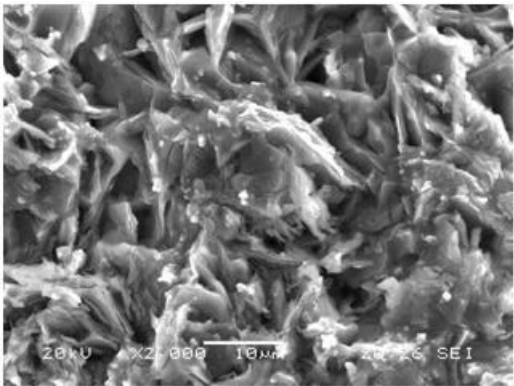


Figure 18. Fractograph of the laser welded joint without Ti filler

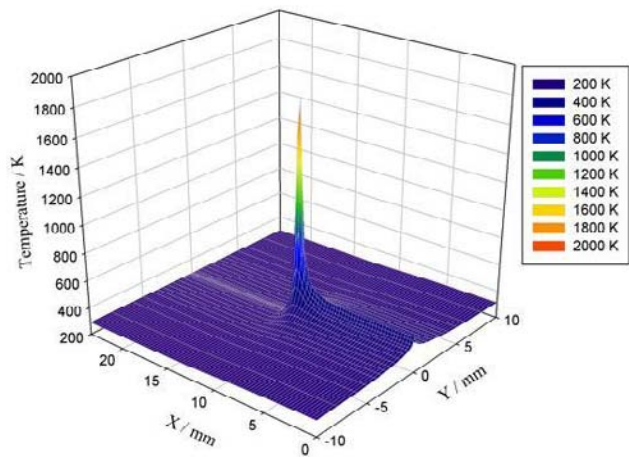


Figure 19. Temperature distribution with Ti filler

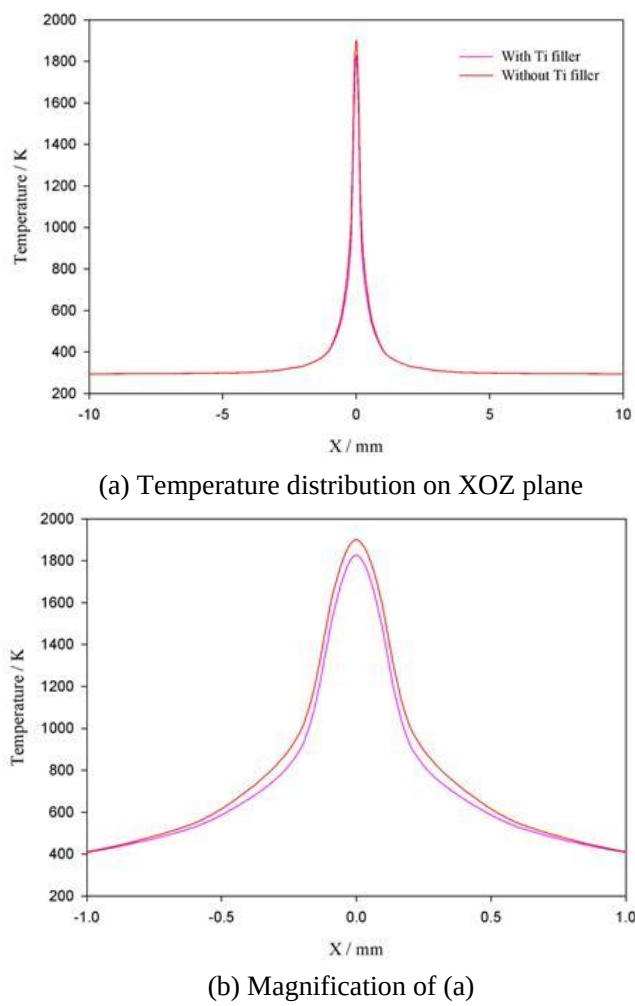
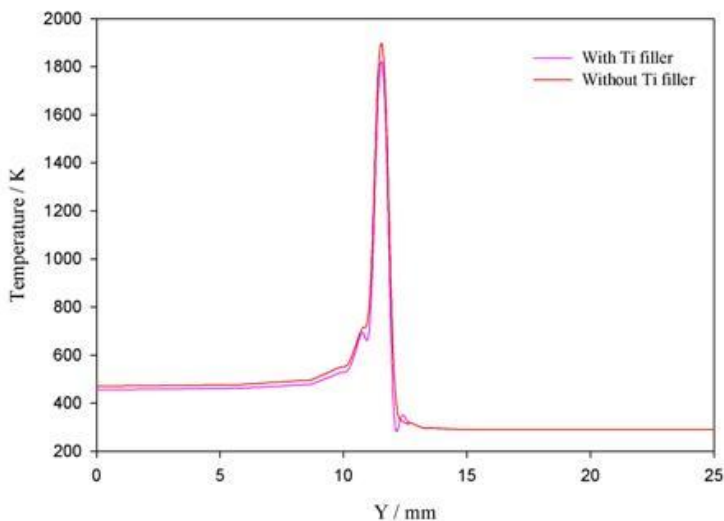
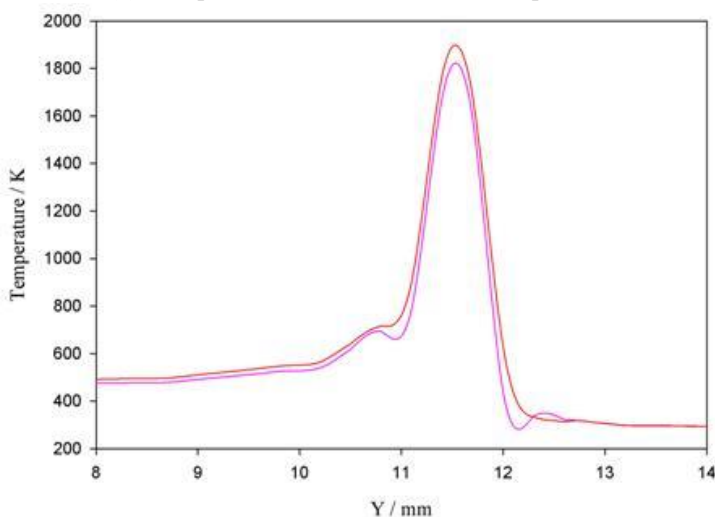


Figure 20. Temperature distribution of central heating on XOZ plane

Figure 19 shows the temperature field of laser welding SiCp/A356 with Ti filler. Considering the Ti melting and in situ reaction in the welding pool as an endothermic reaction, the welding temperature decreases and will be lower than that of laser welding directly (cf. Figures 17 and 19), and its temperature field is distributed more smoothly with in situ reaction than that of laser welding without Ti filler as shown in Figure 20. Also, the width of HAZ is decreased to some extent (Figure 20b). Furthermore, it shows that according to the real effect of laser beam diameter, the thickness of Ti filler is about 0.3 mm will be optimal for in situ welding which conformed to the experimental results as shown in Figure 3.



(a) Temperature distribution on YOZ plane



(b) Magnification of (a)

Figure 21. Temperature distribution of central line on YOZ plane

In addition, the effect of Ti on the temperature distribution on the central line is shown in Figure 21. It illustrates that the peak of the temperature is changed distinctly. Because of the sandwiched Ti between the substrates and in situ endothermic reaction, the temperature of

substrate ahead of laser resource is lower than that of without Ti filler. Moreover, the temperature at the succedent distance is increased or accumulated a little bit due to the different conductive coefficient between Ti and substrate. On the other side, its corresponding trend of the temperature behind the laser resource (resolidification) is same as that of without Ti filler except for a peak appearance induced by more serious exothermic potential during the crystallization.

Figure 22 shows the temperature distribution when Ti filler is thick. The peak of temperature is decreased obviously and leads to the welding failure.

Figure 23 shows the microstructure of laser welded joint with thick Ti filler and its corresponding energy dispersive X-ray spectroscopy (EDX) results. It can be observed that a large number of columnar Ti crystallization is distributed in the weld. From Figures 22 and 23, it elucidates that with the increase of Ti thickness, the heat input into the substrate is decreased and most of energy is used for melting Ti led to the insufficient in situ reaction and stirring in the welding pool resulted in lower properties of welded joints.

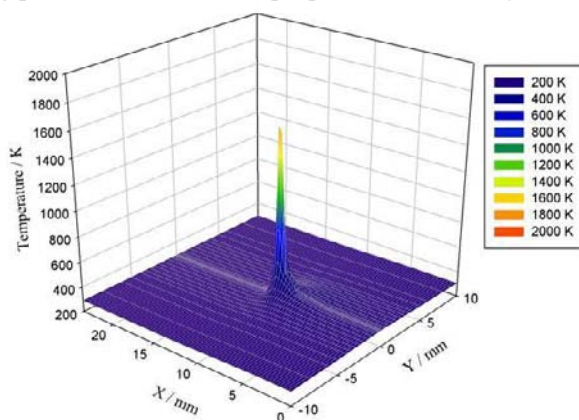


Figure 22. Temperature distribution with thick Ti filler

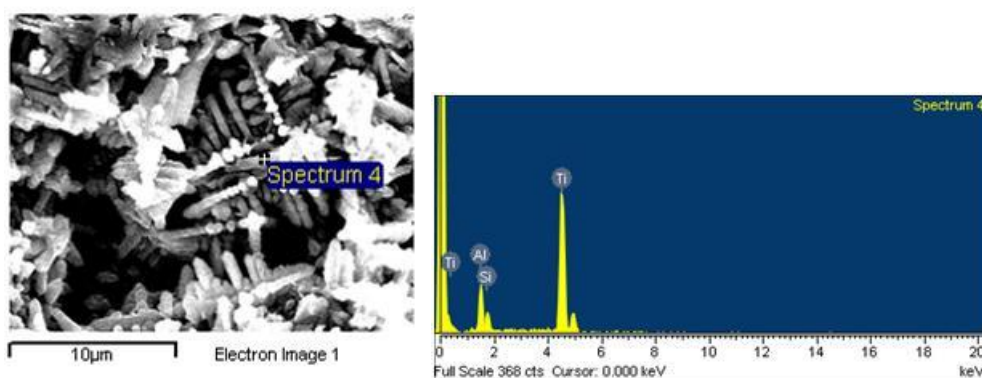


Figure 23. Microstructure and EDX of laser weld with thick Ti filler

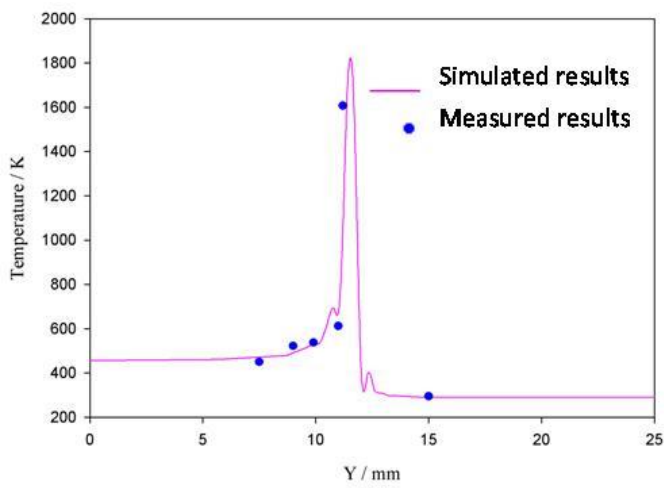


Figure 24. Surface temperature distribution in the processing center

Furthermore, in order to verify the temperature field, noncontact thermometer (model AZ9881) was used to measure the spot temperature on-line. The measured temperature results are shown in Figure 24. It shows that the measured results agree well with the simulated results.

5. FATIGUE TEST

The fatigue test was carried on the Cameron-Plint TE67 wear test rig, where the maximum heating temperature is 500 °C with 2 min preservation, and the heating speed is 20 °C/min. Subsequently, samples are put into the 25 °C water. After cooling, samples are cleaned by acetone and alcohol, and dried by the drier. Finally, samples are observed by optical microscopy. The results are listed in Table 2. It shows that with the fatigue property of laser welding with in situ reaction is superior to that of laser welding directly.

Table 2. Fatigue results with/without in situ reaction

Sample	Cycles (1 unit = 50 cycles)															
	1	2	3	4	5	6	...	14	15	16	17	18	19	20	21	22
A	+	+	+	+	+	+	...	+	+	+	+	+	+	+	-	
B	+	+	+	+	+	+	...	+	-							

*A: With Ti filler

*B: Without Ti filler

CONCLUSION

Titanium as a filler metal in Nd:YAG laser welding of SiC_p/A356 provides beneficial in situ reinforcement effect. Simultaneously, the newly-formed reinforcement TiC particles

distribute uniformly in the weld that assists AMC welding. Moreover, Al_4C_3 formation is effectively restrained in the Nd:YAG laser welding of $\text{SiC}_p/\text{A356}$ with Ti filler. Simulated results illustrates that in addition to the lower heat-input into the substrate because of Ti melting, in situ reaction as an endothermic reaction decreases the heat-input further, and its temperature field distributes more smoothly with in situ reaction than that of laser welding directly. Also, the succedent fatigue test shows the antifatigue property of welded joints with in situ reaction is superior to that of traditional laser welding.

ACKNOWLEDGMENT

The work is supported by a RGC general research fund (GRF) (Grant No.:9041503.) and a Strategic Research Grant (SRG) from City University of Hong Kong (Grant No.: 7002582.)

REFERENCES

- [1] Nair, SV; Tien, JK; Bates, RC. SiC-Reinforced Aluminum Metal Matrix Composites. *Int. Met. Rev.*, 1985, 30(6), 275-290.
- [2] Gupta, M; Srivatsan, TS. Interrelationship between Matrix Microhardness and Ultimate Tensile Strength of Discontinuous Particulate-Reinforced Aluminum Alloy. *Mater. Lett.*, 2001, 51(10), 255-261.
- [3] Shen, YL; Chawla, N. On the Correlation Between Hardness and Tensile Strength in Particle Reinforced Metal Matrix Composites. *Mater. Sci. Eng. A*, 2001, A297, 44-47.
- [4] Gomez de Salazar, JM; Barrena, MI. Dissimilar Fusion Welding of AA7020/MMC Reinforced with Al_2O_3 Particles: Microstructure and Mechanical Properties. *Mater. Sci. Eng. A*, 2003, A352, 162-168.
- [5] Loyd, DJ. Particle-Reinforced Aluminum and Magnesium Matrix Composites. *Int. Mater. Rev.*, 1994, 39(1), 1-23.
- [6] Lienert, TJ; Brandon, ED; Lippold, JC. Laser and Electron Beam Welding of SiCp Reinforced Aluminum A-356 Metal Matrix Composite. *Scripta Metall. Mater.*, 1993, 11(28), 1341-1346.
- [7] Bushby, RS; Scott, VD. Liquid Phase Bonding of Aluminum and Aluminum/Nicalon Composite Using Interlayers of Cu-Ag Alloy. *Mater. Sci. Technol.*, 1995, 11, 643-649.
- [8] Askew, JR; Wilde, JF; Khan, TI. TLP Bonding of 2124 Aluminum Metal Matrix Composite. *Mater. Sci. Technol.*, 1998, 14(5), 920-924.
- [9] Ulrich, K. Tests with regard to the Resistance Spot Welding of Particle – Reinforced Aluminum Matrix Composites. *Weld. Cutt.*, 1999, 51(1), 9-12.
- [10] American Welding Society: ‘*Welding handbook*’; 1996, Miami, FL, American Welding Society.
- [11] Cam, G; Kocak, M. Progress in Joining of Advanced Materials. *Int. Mater. Rev.*, 1998, 43(1), 1-44.
- [12] Wert, JA. Microstructures of Friction Stir Weld Joints between an Aluminum-Base Metal Matrix Composite and a Monolithic Aluminum Alloy. *Scr. Mater.*, 2003, 49(6), 607-612.

- [13] Fernandez, GJ; Murr, LE. Characterization of Tool Wear and Weld Optimization in the Friction Stir Welding of Cast Aluminum A359 + 20 % SiC Metal Matrix Composite. *Mater. Charact.*, 2004, 52(1), 65-75.
- [14] Hsu, CJ; Kao, PW; Ho, NJ. Ultrafine - Grained Al-Al₂Cu Composite Produced In-Situ by Friction Stir Processing. *Scr. Mater.*, 2005, 53(3), 341-345.
- [15] Marzoli, LM; von Strombeck, A; dos Santos, JF; Gambaro, C; Volpone, LM. Friction Stir Welding of an AA6061/Al₂O₃/20p Reinforced Alloy. *Compos. Sci. Technol.*, 2006, 66(2), 363-371.
- [16] Guo, W; Hua, M; Law, HW; Ho, JKL. Liquid-Phase Impact Diffusion Welding of SiCp/6061Al and Its Mechanism. *Materials Science and Engineering, A*, 2008, 490, (1-2), 427-437.
- [17] Guo, W; Hua, M; Ho, JKL. Study on Liquid-Phase-Impact Diffusion Welding SiCp/ZL101. *Compos. Sci. Technol.*, 2007, 67(6), 1041-1046.
- [18] Hua, M; Guo, W; Law, HW; Ho, JKL. Half-Transient Liquid Phase Diffusion Welding: An Approach for Diffusion Welding of SiCp/A356 with Cu Interlayer. *Int. J. Adv. Manuf. Technol.*, 2008, 37, (5-6), 504-512.
- [19] Ochiai, S. *Mechanical Properties of Metallic Composites*. New York: Marcel Dekker, 1994.
- [20] Guagliano, M; Aliabadi, MH. *Fracture and Damage of Composites*. Southampton, Boston: WIT, 2006.
- [21] Guo, KW. Influence of In Situ Reaction on the Microstructure of SiCp/AlSi7Mg Welded by Nd:YAG Laser with Ti Filler. *J. Materials Engineering and Performance*, 2010, 19, 52-58.
- [22] Porter, DA; Easterling, KE. *Phase Transformations in Metals and Alloys*, 2nd. Cheltenham: Nelson Thornes, 2001.
- [23] Riedel, R. *Handbook of Ceramic Hard Materials*. New York: Wiley-VCH, Weinheim, 2000.
- [24] Boyer, R; Welsch, G; Collings, EW. *Materials Properties Handbook: Titanium Alloys*. Materials Park, Ohio: ASM International, 1994.
- [25] Davis, JR. *ASM Specialty Handbook - Aluminum and Aluminum Alloys*. Materials Park, Ohio: ASM International, 1993, 557.
- [26] Mall, S; Nicholas, T. *Titanium Matrix Composites - Mechanical Behavior*. Lancaster, Pa.: Technomic Pub. Co. Inc., 1998.
- [27] Callen, HB. *Thermodynamics and an Introduction on Thermostatistics*. 2nd ed., Wiley New York 1985.
- [28] Kondepudi, D; Prigogine, I. *Modern Thermodynamics: From Heat Engines to Dissipative Structures*, Wiley New York 1998.

Chapter 9

RESIDUAL STRESS EVOLUTION IN WELDED JOINTS SUBJECT TO FOUR-POINT BENDING FATIGUE LOAD

M. De Giorgi^{*}, R. Nobile and V. Dattoma

Dipartimento di Ingegneria dell'Innovazione, Università
del Salento, Via per Arnesano – 73100 Lecce.

ABSTRACT

Residual stresses, introduced into a component by manufacturing processes, significantly affect the fatigue behaviour of the component. External load application produces an alteration in the initial residual stress distribution, so it is reasonable to suppose that residual stress field into a component subject to a cyclic load presents an evolution during the total life. In this work, the authors analysed the evolution that the residual stress field, pre-existing in a butt-welded joint, suffers following the application of cyclic load. The comparison between two residual stress measurements, carried out on the same joint before and after the cyclic load application, allowed to obtain interesting information about the residual stress evolution. It was found that in particular condition, unlike the general opinion, a cyclic load application produces an increasing in the residual stress level rather than a relaxation. This phenomenon is to take well in account in order to avoid unexpected failure in components subjected to a fatigue load.

Keywords: Residual stress, fatigue, four-point bending load, mechanical relaxation.

1. INTRODUCTION

Welding is actually the most used joining technique in every engineering fields, substituting advantageously bolted and riveted joints in ship, pipe, pressure vessel and aeronautical or nuclear applications. Development of modular construction methodologies in

^{*} Corresponding author: Email: Marta De Giorgi; e-mail: marta.degiorgi@unisalento.it.

the field of building and plant engineering allowed overcoming a large number of problems due to execution of welding during assembly.

Large parts of welded components are used in structures subjected to variable loads, determining more or less heavily the strength capacity. Therefore, welding structures are often affected by fatigue phenomena, as it is evident considering the various kinds of loads that generally affect these structures:

- 1) moving loads, having increased entity and frequency, interest normally bridge, ship and crane structures;
- 2) fluctuating pressure, originated by frequent transient operation in plants, acts on pressure vessel, pipe and containers;
- 3) thermal strain, due to stop and start procedure of manufacturing installations, interests process machinery for heat or cold treatment of material;
- 4) vibration in rotating machine and random overloads are finally always possible.

Therefore, it is not surprising that about 90% of engineering component failure can be brought back to fatigue. Welded joint fatigue is complicated by the fact that high residual stress generally exist before external load application, especially at the weld toe. Even if advanced welding technologies are used, thermal cycle associated to welding process introduces several alterations in the material that reduce fatigue strength. During welding, in fact, the component is subjected to severe thermal cycle that produces a highly not uniform temperature distribution. Until temperature remains at high level, a coupled and self-equilibrated thermal and plastic strain field is present; thermal strain is progressively reduced with temperature, while it remains an incompatible strain field, induced by shape variation associated to solidification process, metallurgical changes and plasticizations. Progressive reduction of thermal strain introduces a not equilibrated condition in the material, especially for highest and irregular temperature reduction in the welding component. At room temperature, finally, welded joint will be interested by residual stress state, misalignments and distortions that will influence the in-service structural behaviour.

A pre-existent residual stress state modifies applied nominal mean stress in a substantial way, even if mean stress does not correspond equally to residual stress field, except the case of stress lower than yielding stress. Therefore, the influence of residual stress on fatigue behaviour of welded joints is not easy and widely discussed [1-3].

Residual stresses that usually interest welded joints are often invoked to justify experimental fatigue test result, but their effect represents a debated question, since other factors have a not negligible effect on welded joint fatigue. Geometrical effect, surface irregularities and metallurgical changes in welded zone could hide residual stress influence. Moreover, fatigue crack propagation is higher or lower if a tensile or compressive stress state is encountered. In such case, global effect of residual stress can be negligible. Finally, residual stress field can change with the application of cyclic loads [4-15].

Based on this last consideration, this study considers the interaction existing between fatigue and residual stress; in particular, the evolution of residual stress existing in a butt-joint is followed during the application of an external cyclic load. A first residual stress measurement is carried out to determine initial residual stress field in several butt-welded joints; these joints are then divided into two groups and subjected to two different constant amplitude stress, which correspond to a fatigue life of respectively $0.35 \cdot 10^6$ and $2.8 \cdot 10^6$

cycles. Finally, a second residual stress measurement is carried out to evaluate change induced by fatigue loads.

The comparison of residual stress state before and after fatigue load application is highly complicated from an experimental point of view. Experimental data are affected by a large scatter and uncertainty and physical interpretation is very difficult. Nevertheless, several useful indications are obtained by this kind of experimental measurement.

2. MATERIAL AND METHODS

A total number of 16 butt-welded joints were tested, with the aim to evaluate the interaction residual stress-fatigue. Specimens were obtained by two MIG welded plates having dimensions 800x150mm and two different thicknesses (8 and 20mm) made of structural steel Fe430. The Fe430 is a hot-rolled structural steel of the Italian Standard CNR-UNI 10011 simply identified by its Ultimate Tensile Strength and widely used in mechanical structures. Tensile test was carried out on the base material in order to determine the real value of the yield strength, which resulted equal to 300 N/mm². Figure 1 reports the initial portion of the tensile σ - ϵ curve.

After welding, a milling process is used to remove a 2-mm thick layer of each plate in order to eliminate discontinuity and stress concentration effects caused by the weld seam. Finally, each plate was cut to obtain ten transversal welded joints that were 80 mm wide (Figure 2).

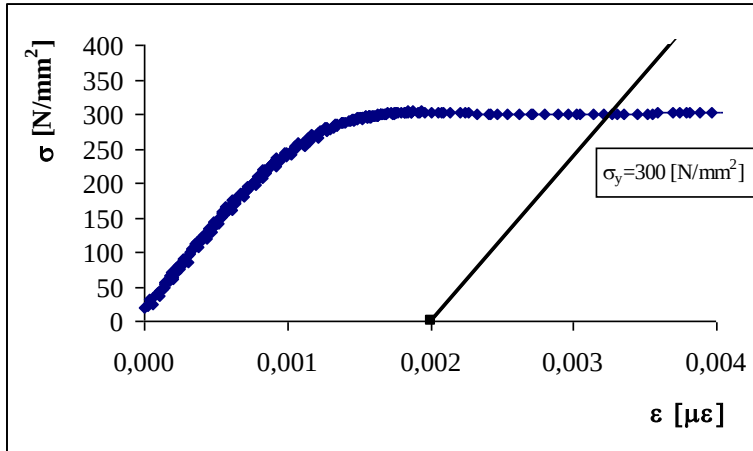


Figure 1. σ - ϵ curve for Fe430 base material.

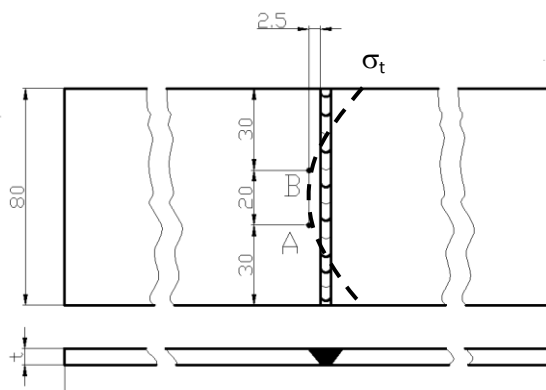


Figure 2. Geometry of the joints

Such joints were subjected to the following experimental procedure:

- Step 1. First measurement of the residual stress field in the point A (Figure 2) to evaluate the initial pre-stress field;
- Step 2. Fatigue load application using four-point bending modality according the experimental plan exposed in detail in the following;
- Step 3. Second measurement of the residual stress field in the point B (Figure 2) to evaluate the final pre-stress field.

Residual stress measurements were carried out by means of the hole-drilling method. Since this methodology is a semi-destructive technique, measurements points A and B had to be different. Considering the ideal transversal residual stress profile (dashed line in Figure 2), measurement points were chosen in symmetric positions respect to longitudinal axis of the specimen. In this way, it was possible to suppose that the initial residual stress level was the same in the points A and B. The hole-drilling method was implemented according to ASTM E 837-01 standards. The diameter and depth of the hole were 1.6 mm and 2 mm, respectively, and subdivided in 40 steps. A vertical motion of 0.05 mm/min and a HBM strain gauge rosette named 1.5/120RY61S were used. Since such incremental hole-drilling method allowed the measurement of non-uniform residual stresses in the thickness, the residual stress dependence against depth was calculated using the power series method [16, 17]. The correction of the residual stresses that exceed one half of the yield stress was carried out based on literature [18].

Fatigue load was applied using the resonant testing machine RUMUL Testronic 50kN and four-point bend loading mode was used in order to reduce the effect of the joint mismatch on fatigue behaviour. A load ratio equal to $R = \sigma_{\min}/\sigma_{\max} = 0.1$ was used.

The specimens were chosen in such a manner that they were affected by different transversal residual stress levels. This was necessary in order to evaluate their effects on fatigue life. The transversal residual stress was considered as the most relevant because it was in the same direction as the applied load (Figure 3).

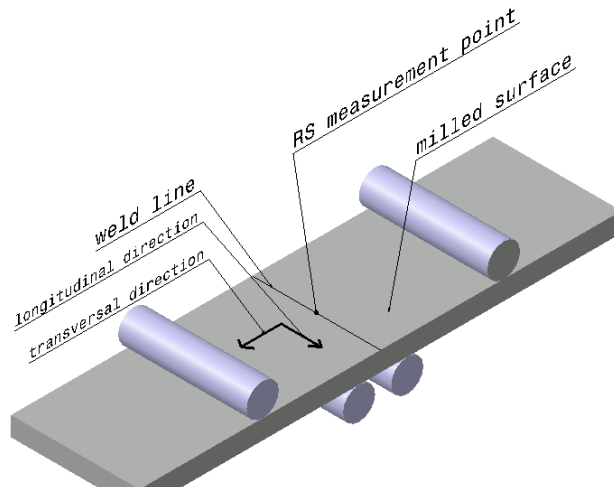


Figure 3. Fatigue test set-up.

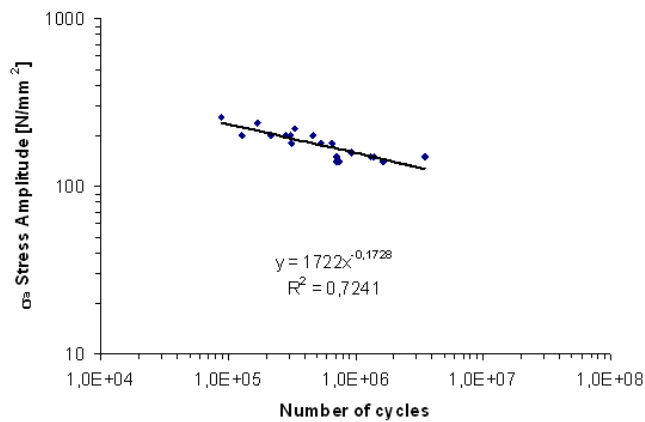


Figure 4. Fatigue curve for 8 mm thick transversal joints.

Based on fatigue curves obtained in a previous work [19] on the same joints and reported in Figures 4-5, two load amplitudes were chosen corresponding to fatigue life equal to 0.35 e 2.8 mln of cycles. In this way, the effect of the load level was evaluated.

In order to evaluate the hardening level caused by welding, micro-hardness measurements were performed near the weld seam along a transversal line at depth of 2 mm. The Vickers micro-hardness profile is reported in Figure 6a. Micro-hardness measurements in points far from the weld seam allowed obtaining the hardness HV value of the base material equal to 160 kg/mm². Dividing this value by the yield stress value of the base material, it obtained a correlation factor useful to calculate the yield stress (Figure 6b) of the welded material based on micro-hardness profile. The maximum value of the yield stress was found to be 400 N/mm² in proximity of the weld seam. This value will be very useful to interpret the residual stress data as exposed in the next section.

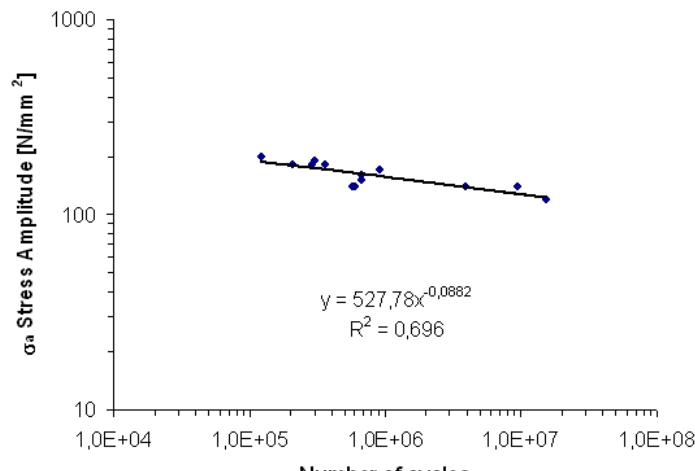


Figure 5. Fatigue curve for 20 mm thick transversal joints.

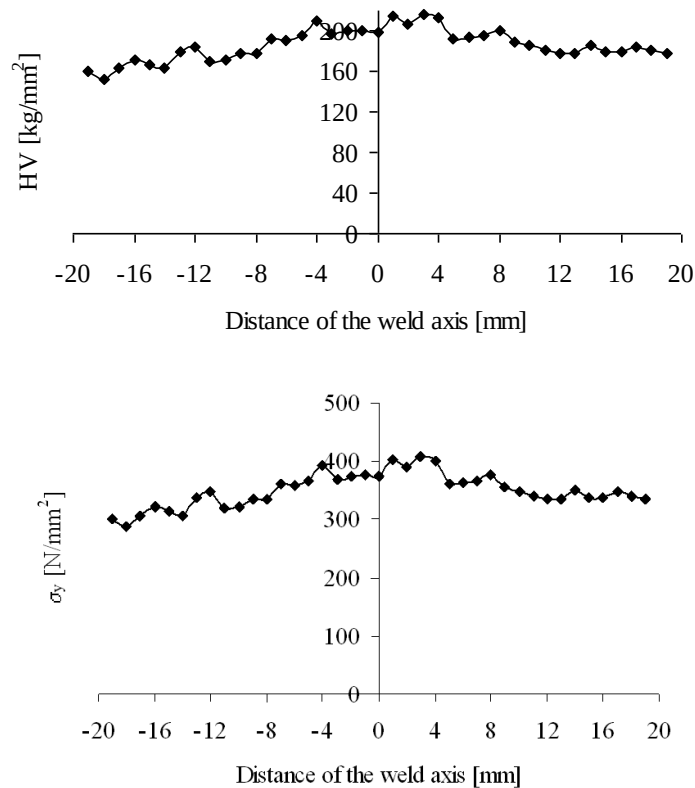


Figure 6. Micro-hardness HV a) and yield stress b) profile along a transversal line to the weld axis.

2.1. Experimental Plan

Fatigue test plan was defined according to the following steps:

1. subdivision of the specimens having same thickness in two sets each of them composed by four specimens having different initial residual stress level; in this way also the effect of the initial residual stress value on the relaxation process will be evaluated;
2. application of amplitude load corresponding to $0.35 \cdot 10^6$ cycles in the Wöhler curve for number of cycles equal to 1%, 5%, 10% and 20% of $0.35 \cdot 10^6$ cycles at two sets of specimens (one for each thickness);
3. application of amplitude load corresponding to $2.8 \cdot 10^6$ cycles in the Wöhler curve for number of cycles equal to 1%, 5%, 10% and 20% of $2.8 \cdot 10^6$ cycles at the remaining sets of specimens (one for each thickness); it could be noticed that this load amplitude corresponded to the fatigue limit.

Following the scheduled load cycles application, the second residual stress measurement was performed on each specimen. The final residual stress value was different from the initial value because of the applied load cycles, supposing negligible the measurement errors. The load program is reported in detail in Table 1.

Table 1. Experimental plan and residual stress values at 0.2 mm depth.

Joint		Load amplitude [N/mm ²]	N _i	Initial transversal residual stress [N/mm ²]	Final transversal residual stress [N/mm ²]	Initial longitudinal residual stress [N/mm ²]	Final longitudinal residual stress [N/mm ²]	Initial Von-Mises residual stress [N/mm ²]	Final Von-Mises residual stress [N/mm ²]
8 mm	1	140 $N_f=2.8 \cdot 10^6$	1% = 28000	-76	-3	-104	-65	93	63
	2		5% = 140000	51	5	-68	-167	103	170
	3		10% = 280000	36	89	-159	98	180	94
	4		20% = 560000	-3	31	-152	-31	150	54
	5	200 $N_f=0.35 \cdot 10^6$	1% = 3500	-34	29	-198	-57	183	76
	6		5% = 17500	16	27	-33	-78	43	94
	7		10% = 35000	12	12	-119	-10	125	19
	8		20% = 70000	-11	-8	-93	-99	88	95
20 mm	1	142 $N_f=2.8 \cdot 10^6$	1% = 28000	4.6	-10	-51	-89	53	84
	2		5% = 140000	34.2	-98	-132	-25	151	88
	3		10% = 280000	-96.8	45	-162	-74	141	104
	4		20% = 560000	-5.6	-128	-109	-131	106	129
	5	171 $N_f=0.35 \cdot 10^6$	1% = 3500	36.9	-19	-184	-117	204	109
	6		5% = 17500	4.4	-20	-111	-21	113	20
	7		10% = 35000	-15.5	-25	-170	-111	163	101
	8		20% = 70000	-198.2	-54	-185	-168	192	148

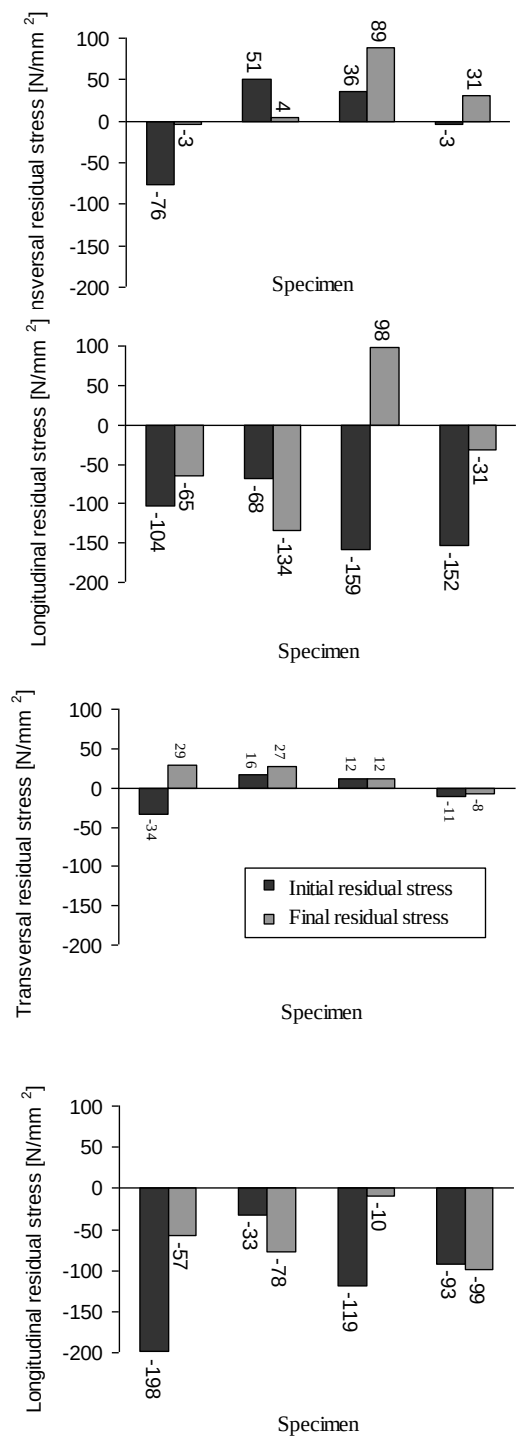


Figure 7. Comparison between residual stress before and after application of load having amplitude 140 N/mm^2 a) and 200 N/mm^2 b) for 8 mm thick joints.

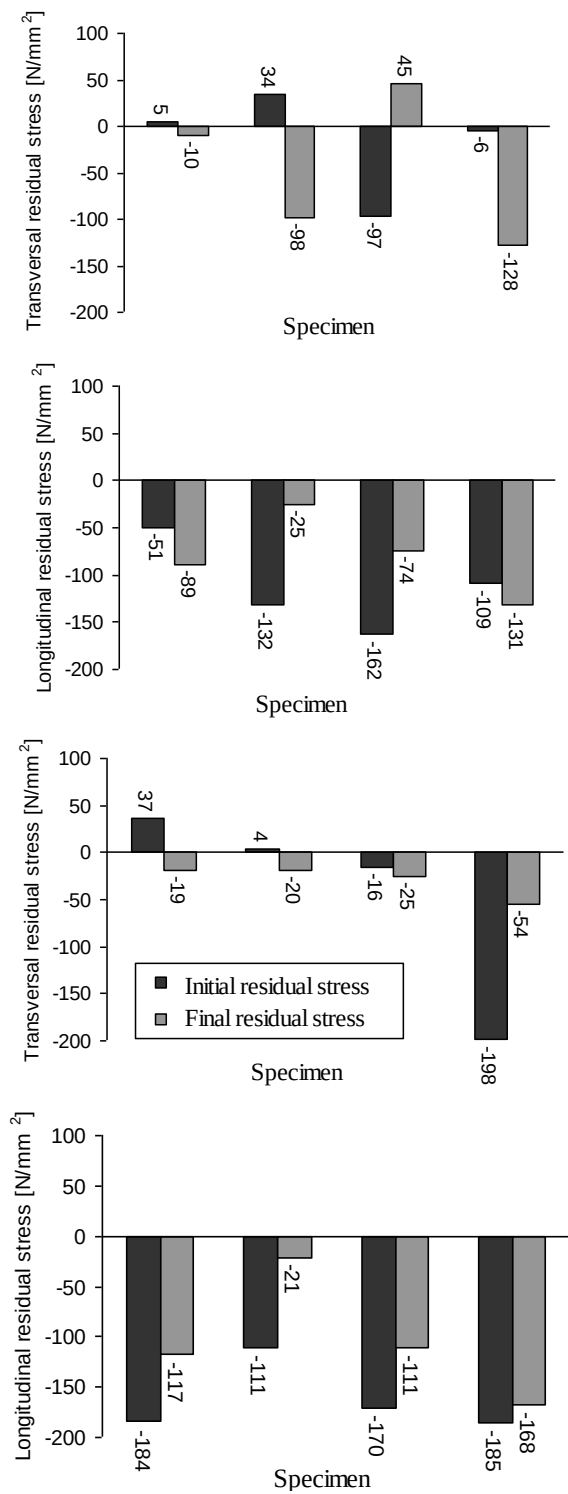


Figure 8. Comparison between residual stress before and after application of load having amplitude 142 N/mm² a) and 171 N/mm² b) for 20 mm thick joints.

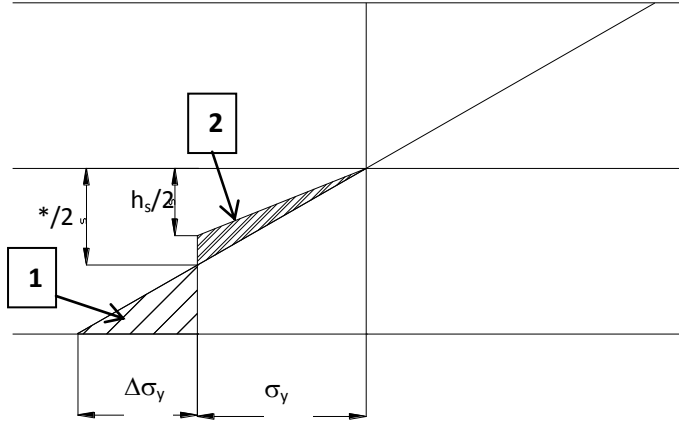


Figure 9. Ideal profile of the bending stress in the cross-section.

3. RESULTS AND DISCUSSION

The initial and final residual stresses, expressed as longitudinal, transversal and Von Mises stress for each specimen, measured at 0.2 mm depth are reported in Table 1. Transversal and longitudinal residual stress values before and after load application are also reported as histogram, in the Figures 7 and 8, for a more immediate comparison.

The analysis of the histogram did not show any common behaviour of the specimens neither any evident effect of the applied load on residual stress modification. The observation of the phenomenon is particularly complicated by the fact that the residual stress measurements present a high variability in proximity of weld cord. Moreover, the initial transversal residual stresses were quite low, except for few cases. The analysis of the stress field resulting from the superposition of initial residual stress and applied load resulted more fruitful. In this case, it was essential to consider the plasticization mechanism due to external load application. At this aim, it was considered a simple but efficient analytical plasticization model that described what occurred when yield stress was exceed.

Referring to a generic rectangular cross-section subjected to a bending load and supposing a bi-linear material behaviour, it is possible to determine qualitatively and quantitatively what is the actual stress distribution in the cross-section. Denoting as $\Delta\sigma_y = \sigma_{\max,e} - \sigma_y$ the stress amount that exceed the yield stress if the material is perfectly elastic, it is feasible to calculate the height h_s of the unyielded central portion of the section (Figure 9).

Imposing that the areas of the triangles 1 and 2 are equal, it is possible to obtain following relations from simple geometric considerations:

$$h^* = \frac{h \times \sigma_y}{\Delta\sigma_y + \sigma_y} \quad (1)$$

$$h_s = \left(1 - \frac{\Delta\sigma_y}{\sigma_y} \right) \times h \quad (2)$$

where h is the specimen thickness, h^* and h_s are indicated in Figure 9, σ_y is the material yield stress and $\Delta\sigma_y$ is the difference between the maximum stress (in perfectly elastic regime) and yield stress.

The derived model, even if so simple, describe with good accuracy what happens in uniaxial stress state. In the present case, however, a biaxial stress state must be considered because of the presence of relevant longitudinal residual stresses. For this reason, the previous model can be considered valid if the stress $\sigma_{\max,e}$ used to calculate $\Delta\sigma_y$ is the Von Mises equivalent stress, applying it at the stress state resulting from superposition of the initial residual stress and the maximum applied stress reached during the fatigue test of each specimen.

In particular, it is considered that the acting transversal residual stress σ_{tr} was the sum of the initial transversal residual stress and the maximum bending stress. Von Mises equivalent stress corresponded to the simultaneous presence of the initial longitudinal residual stress and the transversal stress σ_{tr} . In this way, it is determined the $\Delta\sigma_y$ near the weld seam, where $\sigma_y = 400$ MPa, used in the relations (1) and (2) to calculate h^* and h_s and then the percentage of the plasticised cross section in each specimen, as reported in Table 2.

Table 2. Introduced and initial Von Mises residual stress for each specimen at 0.2 mm depth and percentage of plasticised cross-section.

Specimen			TR _{VM_initial}	$\Delta\sigma_y$	% of plasticization	TR _{VM_final} - TR _{VM_initial}
8 mm	140 N/mm ²	1	93	-99	-25	-30
		2	103	0,5	0	66
		3	180	48	12	-86
		4	150	6	1	-97
	200 N/mm ²	5	183	137	34	-107
		6	43	78	19	51
		7	125	126	31	-106
		8	88	87	22	7
20 mm	142 N/mm ²	1	53	-51	-12	31
		2	151	31	8	-64
		3	141	-69	-17	-37
		4	106	-23	-6	23
	171 N/mm ²	5	204	133	33	-96
		6	113	50	12	-93
		7	163	73	18	-62
		8	192	-82	-20	-43

Table 2 reports also the initial Von Mises residual stresses values and their variation caused by fatigue cycles application. Negative values of plasticization percentage did not have a physical meaning, but, due to the calculation modality, indicated simply that the material of the cross section did not have reached the yield surface.

The data in Table 2 allowed leading interesting considerations, facilitated by proper diagram. First diagram reports the Von Mises residual stress variation versus the plasticisation percentage (Figure 10).

For the negative values of percentage, or rather for absence of plasticization, the residual stress variation returned in an interval of $\pm 40 \text{ N/mm}^2$, comparable with the residual stress variability and measurement error.

In presence of plasticization, large part of specimen presented a significant reduction of residual stress that resulted in the range $60 \div 110 \text{ N/mm}^2$. On the contrary, three specimens 8 mm thick presented a residual stress increment rather than reduction, even if they reached the yield surface.

This behaviour could seem anomalous, but through a deeper analysis it was possible to observe that these three specimens presented the minimum initial residual stress level, excluding the specimens that did not reach the yielding conditions. To confirm this observation, it is useful to observe the diagram in Figure 11, where the trend of the Von Mises residual stress variation is reported against the initial Von Mises residual stress. For lower initial residual stress, the pre-stress field increased; for higher initial residual stress, a significant reduction of the residual stress occurred. Practically, when the initial residual stress field was low, the stress state approached yield surface, determining the increase of the initial residual stress. On the contrary, when the initial residual stress was high, the significant plasticization caused by the load application relaxed the initial residual stress. It was also evident the presence of a threshold value of the initial residual stress beyond that residual stress relaxation occurred: this value was about 100 N/mm^2 .

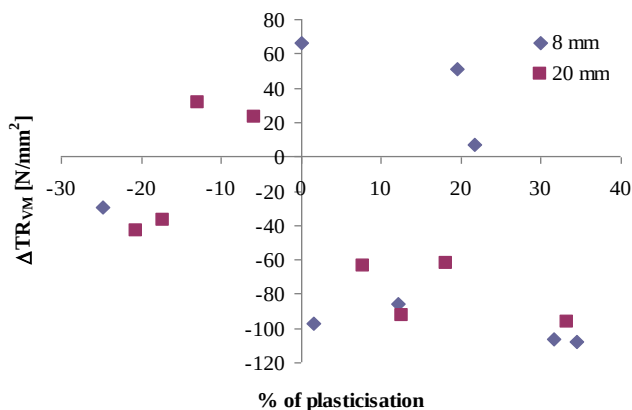


Figure 10. Von Mises residual stress variation versus the plasticisation percentage.

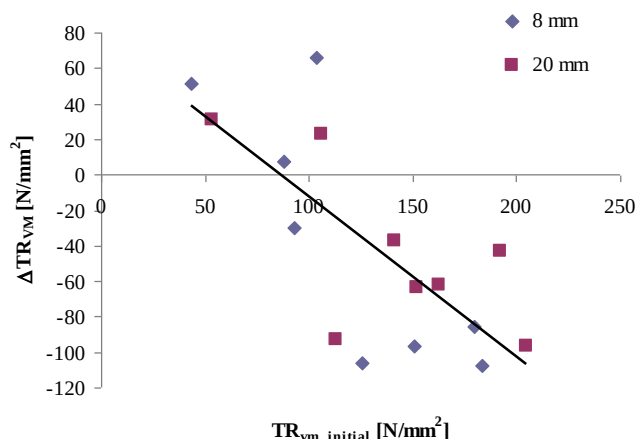


Figure 11. Von Mises residual stress variation as function of initial Von Mises residual stress.

4. CONCLUSIONS

In this work, the analysis of the interaction residual stress-fatigue behaviour of butt-welded joints has been carried out through the comparison between two pre-stress fields before and after the application of a four-point bend cyclic load. In this way, some interesting informations have been obtained. In particular, it has been found that they exist particular conditions where, unlike commonly asserted, the cyclic load application causes the increase of the residual stress rather than their relaxation.

Analysing the obtained data, it can be concluded that, in absence of plasticization in the cross-section of the specimen, the residual stress relaxation can be negligible since it results quite equal to the measurement error. On the contrary, in presence of plasticization, it results that, for low initial residual stress, the pre-stress field increases, while for higher initial residual stress, they relax significantly.

REFERENCES

- [1] Gurney, TR. *Fatigue of welded structures*, Cambridge University Press, 1979.
- [2] Masubuchi, K. Analysis of welded structure, *International Series on Materials Science and Technology*, Vol 33, Pergamon Press.
- [3] Masubuchi, K. *Residual stresses and distortion*, ASM Handbook, Vol. 6, 1992.
- [4] Bergstrom, J; Ericsson, T. Proceedings, Second International Conference on Shot Peening, ICS P-2, American Shot Peening Society, Paramus, NJ, 1984, 241-248.
- [5] Blom, AF. Spectrum fatigue behaviour of welded joints, *Int. J. Fatigue*, 1995, 17, 485-491.
- [6] Dattoma, V; De Giorgi, M; Nobile, R. *Numerical evaluation of residual stress relaxation by cyclic load*, *J. Strain Anal*, 2004 39, 663-672.
- [7] Lopez Martinez, L; Lin, R; Want, D; Blom, AF. *Investigation of residual stresses in as-welded and TIG-dressed specimens subjected to static/spectrum loading*. In:

- Proceedings of the North European Engineering and Science Conference (NESC): Welded High-Strength Steel Structures, Stockholm, Sweden (Edited by AF. Blom), EMAS Publishing, London, UK, 1997.
- [8] Khanna, SK; He, C; Agrawal, HN. Residual stress measurement in spot welds and the effect of fatigue loading on redistribution of stresses using high sensitivity moire interferometry. *J. Engng. Mater. Technol*, 2001, 123, 132-138.
- [9] Iida, K; Yamamoto, S; Takanashi, M. *Residual stress relaxation by reversed loading*. Welding in the World/Le Soudage dans le Monde, 1997, 39, 138-144.
- [10] Iida, K; Takanashi, M. *Relaxation of welding residual stresses by reversed and repeated loadings*. Welding in the World/Le Soudage dans le Monde, 1998, 41, 314-327.
- [11] Takanashi, M; Kamata, K; Kunihiro, I. Relaxation behavior of welding residual stresses by fatigue loading in smooth longitudinal butt welded joints. *Welding World*, 2000, 44, 28-34.
- [12] Nitschke-Pagel, Th; Wohlfahrt, H. Residual stress relaxation in welded high strength steels under different loading conditions. In: *Proceedings of the 6th International Conference on Residual Stresses*, ICRS-6, Oxford, UK, 2000, 1495-1502.
- [13] Lachmann, C; Nitschke-Pagel, Th; Wohlfahrt, H. Characterisation of residual stress relaxation in fatigue loaded welded joints by x-ray diffraction and barkhausen noise method. In: *ECRS 5, Proceedings of the 5th European Conference on Residual Stresses*, Delft-Noordwijkerhout, the Netherlands, Mater. Sci. Forum, 2000, 347, 374-379.
- [14] Han, ST; Lee, Shin, B. Residual stress relaxation of welded steel components under cyclic load. *Mater. Technol*, 2002, 73, 414-420.
- [15] Casavola, C; Dattoma, V; De Giorgi, M; Nobile, R; Pappalettere, C. Experimental Analysis of the Residual Stresses Relaxation of Butt-Welded Joints Subjected to Cyclic Load, 4th *Int.Conf. on Fracture Damage Mechanics*, 12-14 July 2005, Mallorca, Spain.
- [16] Kelsey, RA. Measuring non-uniform residual stress by the hole drilling method, *Proceedings SESA*, 1956, Vol. 14, n. 1, 181-194.
- [17] Vangi, D. Data management for the evaluation of residual stress by the incremental hole-drilling method, *ASME Journal of Engineering Materials and Technology*, 1994, Vol. 116, 561-566.
- [18] Beghini, M; Bertini, L; Raffaelli, P. Numerical analysis of plasticity effects in the hole-drilling residual stress measurement, *Journal of Testing and Evaluation*, 1994, v. 22, n 6, Nov. 522-529.
- [19] Dattoma, V; De Giorgi, M; Nobile, R. Some considerations about fatigue failure of components affected by residual stress, *Journal of Mechanical Science and Technology*, 2010, 24(2). 453~460, DOI 10.1007/s12206-009-1208-4.

INDEX

A

absorption, 135, 392, 394, 395, 399, 400, 401, 402, 404, 405, 406, 408, 412
accessibility, 115
accuracy, 49, 84, 86, 164, 167, 172, 228, 244, 257, 258, 289, 292, 298, 306, 405, 448
acetone, 419, 434
acquisition of knowledge, 226
acrylonitrile, 406
additives, 394, 395, 399, 400, 408, 412
adjustment, 52, 76, 121
advantage effect, 420
advantages, xiii, 2, 104, 112, 116, 118, 166, 244, 391, 408, 413
aerospace, 124, 244, 334, 418
aggregation, 217, 219
algorithm, 35, 83, 193, 215, 220, 221, 222, 225, 226, 228, 231, 232, 233
aluminium, 282
amorphous polymers, 394
amplitude, 137, 377, 439, 443, 444, 445, 446
anisotropy, 286
annealing, 220, 221, 222, 246, 258, 266, 274, 275, 280
ANOVA, 194, 195, 201, 203, 204, 209, 210, 212, 214, 224
antifatigue, xiii, 417, 434
artificial intelligence, 192
Artificial Neural Networks, 184
assessment, xii, 99, 177, 190, 285, 286, 354, 357, 358, 359, 360, 362, 364, 369, 370, 371, 373, 374, 375, 378, 379, 380, 381, 382, 383, 384, 385, 386, 388, 399
assessment procedures, xii, 285, 286, 354, 357, 359, 360, 364, 388
astigmatism, 28, 61, 62

asymmetry, 39, 66, 76
atmospheric pressure, 116
atoms, 132, 136, 137, 261, 275
Austria, 246
automation, x, 111, 194, 195, 227, 409
automobiles, xii, 391

B

background noise, 61
beams, ix, x, 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 24, 27, 28, 29, 33, 35, 41, 42, 67, 74, 77, 83, 87, 92, 97, 103, 104, 106, 108, 109, 112, 113, 114, 115, 118, 131, 183
Belarus, 122
bending, 326, 327, 333, 334, 368, 375, 376, 437, 440, 447, 448
bias, 100, 106, 107, 273
biocompatibility, 412
biomedical applications, 416
Boltzmann constant, 68
boundary conditions, 83, 85, 86, 132, 200, 288, 321, 322, 324, 326, 335, 336, 347, 354, 376, 428
boundary value problem, 288
bounds, 287, 289, 370
buildings, 187
Bulgaria, 1, 109, 111, 121, 182, 416
butadiene, 406

C

C++, 222
calibration, 41
capillary, 137
carbon, 41, 400, 401, 414
case study, 190
casting, 6

categorization, 368
 cathode materials, 70
 ceramic, 418
 charge density, 84, 85, 93, 94, 96, 97
 chemical bonds, 7
 chemical etching, 148
 China, 245, 246
 City, 417, 435
 class, 112, 290, 343, 382
 cleaning, 115, 198, 419
 clusters, 131
 CO₂, 191, 394, 396, 413
 coatings, 116
 coding, 227
 collisions, 8, 12, 14, 16, 30, 67, 137
 color, 148, 394, 400, 401, 407, 416
 compatibility, 395, 406, 408
 compensation, ix, 2, 3, 14, 15, 17, 21, 22, 60, 137
 complexity, x, 111, 185, 290, 373
 composition, 7, 81, 124, 125, 129, 136, 149, 150, 187, 188, 189, 273, 399, 400, 412, 418
 compression, 24, 269, 292, 295, 296, 298, 386
 computation, 99
 computer simulation, x, 2, 80, 83, 85, 91, 92, 97, 103, 108, 115, 121
 computer simulations, 91
 conduction, 81, 190, 395, 396, 403, 404, 405
 conductivity, 17, 81, 118, 139, 143, 150, 155, 163, 187, 258, 259, 267, 279, 280, 284, 398, 399
 configuration, xi, 3, 18, 34, 68, 77, 86, 93, 97, 99, 103, 104, 124, 211, 213, 247, 248, 251, 252, 254, 258, 262, 265, 267, 268, 271, 276, 292, 294, 305, 307, 310, 311, 317, 318, 326, 335, 343, 366, 368, 396, 407
 configurations, 249, 250, 262, 263, 286, 311, 318, 370, 396
 conflict, xi, 186, 226, 231, 232, 233, 244
 conflict resolution, 226, 231, 232, 233, 244
 conservation, 16, 30, 93, 94
 contaminant, 131, 408
 contamination, 188, 258, 412
 contour, 55, 97, 99, 104, 105, 151, 153, 155, 161, 172, 173, 177, 179, 397, 398, 402, 403, 406, 409
 contradiction, 322
 convergence, 11, 75, 78, 79, 80, 84, 290, 303
 conviction, 115
 cooling, 121, 186, 188, 189, 190, 191, 221, 222, 256, 282, 395, 434
 copper, 123, 124, 128, 129, 131, 132, 150, 282
 correlation, 31, 33, 57, 140, 172, 442
 correlation coefficient, 33
 corrosion, xii, 165, 357, 358, 359, 388, 418

cost, xii, 115, 166, 187, 189, 190, 193, 195, 244, 391, 395, 397, 409, 412
 Coulomb interaction, 30
 covering, xii, 253, 270, 357, 358, 372, 388
 creep, xii, 357, 358, 359, 368, 373, 374, 388
 critical value, 3, 138, 319, 341
 cross-validation, 169
 crystal structure, 118, 119
 crystalline, 70, 100, 401, 409, 413, 414
 crystallinity, 399
 crystallization, 83, 136, 432
 crystals, 100, 425, 427
 current limit, 15
 cycles, 144, 166, 190, 377, 380, 434, 439, 442, 443, 449

D

damages, iv
 data distribution, 156
 data processing, 33
 data set, 226
 data structure, 227
 database, 211, 226
 datasets, 167
 decomposition, 398, 403
 defects, 113, 137, 174, 181, 194, 275, 358, 373, 379, 385
 deformation, 251, 256, 258, 265, 266, 273, 274, 275, 277, 278, 279, 280, 281, 286, 314, 354, 368
 degradation, 358, 368
 deposition, 112, 187, 403
 deposition rate, 187
 deviation, 27, 48, 49, 150, 164, 167, 174
 diagnosis, 166
 dialysis, 410
 diaphragm, 67, 78, 80
 dielectric constant, 15, 62
 dielectric permittivity, 61, 86
 differential equations, 26, 27, 94, 96
 diffraction, 418, 451
 diffusion, 120, 257, 258, 262, 274, 275, 276, 282, 283, 395, 401, 403
 diffusion process, 275
 diffusivity, 139, 163, 276, 283
 diode laser, 394, 395, 397, 407, 412, 413, 414, 416
 diodes, 83, 96
 discontinuity, 287, 289, 290, 292, 294, 296, 300, 301, 302, 303, 304, 307, 309, 310, 311, 312, 313, 318, 319, 321, 323, 324, 325, 326, 328, 329, 330, 331, 335, 337, 339, 340, 341, 343, 346, 347, 348, 350, 351, 352, 353, 355, 439
 discriminant analysis, 174

dispersion, 128, 137, 138
 displacement, 126, 361, 368
 distortion, x, 28, 115, 185, 186, 188, 189, 190, 191, 193, 198, 201, 203, 207, 209, 212, 213, 214, 215, 216, 220, 223, 224, 225, 236, 240, 241, 242, 243
 distortions, xi, 30, 116, 185, 186, 187, 189, 190, 438
 distribution function, ix, 1, 8, 9, 11, 31
 divergence, 2, 11, 30, 46, 60, 61, 62, 63, 84, 104, 106, 108, 113
 ductility, 188, 274, 405
 dyes, 406, 415

E

Efficiency, 414, 416
 efficiency level, 155
 elastomers, 403, 409, 413, 415
 electric field, 3, 15, 18, 20, 23, 24, 71, 83, 84, 90, 100
 electrical fields, 89
 electricity, 15
 electrodes, 16, 67, 68, 72, 76, 77, 78, 80, 86, 87, 93, 114, 121
 electromagnetic, 76, 93, 113, 121
 electromagnetic field, 93, 113
 electromagnetic fields, 113
 electron beam lithography, 7
 electrons, ix, 1, 2, 3, 4, 7, 8, 11, 12, 14, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 27, 29, 30, 34, 41, 44, 45, 48, 67, 68, 71, 72, 73, 74, 76, 80, 84, 86, 91, 92, 93, 94, 95, 96, 98, 100, 101, 103, 104, 105, 119, 131, 136, 137, 139
 emission, 11, 15, 24, 34, 67, 68, 69, 70, 71, 72, 74, 78, 80, 82, 83, 89, 90, 92, 96, 102, 104, 401
 emitters, 14, 70, 71, 80, 81, 92
 encapsulation, 412
 endothermic, xiii, 417, 431, 432, 434
 endurance, 380
 energy consumption, 396
 energy density, 4, 113, 137, 400, 402, 403, 407
 energy efficiency, 396
 engineering, xii, 116, 118, 166, 172, 180, 181, 187, 189, 192, 193, 225, 245, 285, 354, 357, 358, 362, 373, 388, 438
 environmental conditions, xii, 187, 357, 358, 404
 environmental degradation, 368
 equality, 24, 139
 equilibrium, 3, 4, 101, 131, 186, 320, 322
 equipment, ix, 1, 108, 114, 116, 120, 177, 187, 409
 etching, 148, 165
 ethanol, 407
 ethyl alcohol, 419
 EU, 358, 388

evacuation, 121
 evaporation, 6, 7, 70, 82, 112, 131, 132, 136, 137
 examinations, 129
 exclusion, 86
 execution, 190, 211, 231, 438
 experiences, 266, 278, 279, 282
 experimental condition, 124, 125, 126, 128, 147
 experimental design, 172
 expert systems, 66, 181, 192, 193, 211, 212, 216
 exploitation, 82, 83
 exploration, 81
 exposure, 266, 277, 280, 281, 398
 extinction, 394
 extraction, 67, 213
 extrusion, 266, 272, 273, 278, 279

F

fabrication, xi, xii, 7, 186, 187, 369, 381, 391, 410
 FAD, 359, 360, 361, 362, 369, 375
 fault diagnosis, 166
 FDA, 379, 380, 382, 383
 feedback, 195
 FEM, 191, 418
 fiber, 395, 397, 399, 400, 404, 407, 412
 fibers, 321, 418
 filament, 80, 81, 82, 83, 98, 100
 fillers, 399, 400
 films, 112, 121, 258, 396, 413
 financial support, 284
 finite element method, 289, 292, 418
 fluctuations, 102
 fluid, 404
 formula, 46, 49, 57, 84, 140, 160, 229, 291, 373, 374
 France, 121, 245
 free energy, 425
 free volume, 123
 free will, 232
 frequencies, 259
 friction, xi, 124, 247, 250, 251, 252, 253, 254, 255, 257, 258, 259, 261, 262, 267, 268, 269, 273, 274, 275, 276, 277, 280, 281, 284, 292, 295, 296, 370, 392
 function values, 194
 fusion, 124, 138, 142, 153, 155, 159, 178, 188, 196, 248, 381, 392, 398, 403, 421
 fuzzy sets, xi, 186, 213, 216, 217, 218, 220, 221, 222, 224, 226, 227, 228, 229, 230, 232, 233, 234, 235, 236, 240, 244

G

general knowledge, 193
 geometrical parameters, 364
 geometrical properties, 31
 Germany, 15, 121, 123, 129, 388, 389
 grades, 402, 404, 407
 grain boundaries, 275
 grounding, 39
 growth mechanism, 373
 growth rate, 384
 guidance, 123, 363, 372, 373
 guidelines, 359, 379

H

Hamiltonian, 94
 hardness, 119, 120, 128, 148, 165, 188, 281, 283, 442, 443
 harmony, 279
 heat capacity, 145, 150
 heat conductivity, 267, 279, 284
 heat loss, 81, 251, 258, 398, 403
 heat transfer, x, 111, 112, 113, 115, 132, 138, 139, 143, 155, 162, 163, 401, 403, 405
 heat treatment, 138, 193, 199, 274, 275, 370, 373, 394
 height, 73, 123, 135, 136, 364, 447
 heteroscedasticity, 172
 high density polyethylene, 404
 histogram, 378, 447
 homogeneity, 29
 Hong Kong, 417, 435
 housing, 91, 410, 411
 human brain, 212
 human intelligence, 192
 hydrostatic stress, 321

I

ideal, 11, 12, 21, 27, 29, 81, 162, 212, 233, 256, 267, 273, 276, 279, 440
 illumination, 136
 image, 12, 25, 27, 28, 29, 35, 39, 42, 43, 44, 49, 66
 images, 26, 27, 28, 29, 45, 148
 immersion, 103
 impact strength, 190
 impulses, 8, 12, 35, 42
 induction, 20, 190
 inequality, 293, 310, 321, 331, 333, 341
 information processing, 193

initiation, 189, 358, 360, 361, 362, 377
 insertion, 48
 Instron, 125, 126
 integration, 10, 16, 19, 165, 304, 305, 322, 323, 324, 325, 338, 340, 341, 349, 409
 intelligence, 192
 interface, xi, 128, 134, 137, 186, 192, 213, 227, 228, 235, 236, 240, 241, 242, 243, 244, 262, 266, 274, 275, 292, 294, 326, 335, 355, 364, 392, 396, 400, 401, 405, 406, 407, 409, 420, 427, 428
 intermetallic compounds, 427
 intermetallics, 250, 275, 276
 intervention, 244
 ion bombardment, 71, 76, 98
 ionization, 67, 394
 ions, ix, 2, 3, 4, 17, 21, 22, 23, 24, 137
 irradiation, 7, 138, 182, 402, 404, 406, 407, 413, 424
 isolation, 81, 121
 isotherms, 162
 Italy, 246
 iteration, 97, 99, 221, 222

J

joints, xii, xiii, 108, 115, 124, 131, 190, 285, 286, 334, 354, 355, 363, 370, 376, 385, 386, 388, 389, 391, 396, 406, 407, 408, 417, 418, 419, 420, 421, 424, 425, 426, 427, 429, 432, 434, 438, 439, 440, 442, 445, 446, 450, 451

K

knowledge acquisition, 192, 193

L

laminar, 11, 16, 22, 23
 laser ablation, 113
 laser beam welding, 414, 416
 laser radiation, 394
 lasers, 112, 113, 115, 394, 395, 396, 401, 406, 412, 413
 learning, xi, 165, 186, 192, 193, 195, 226, 235
 lens, 12, 25, 26, 27, 28, 29, 44, 45, 46, 49, 53, 54, 59, 60, 61, 63, 76, 78, 91, 103, 146, 147, 148, 150, 159, 160, 169, 393, 404
 lifetime, 82, 83, 212
 ligament, 369
 light scattering, 400, 401
 light transmission, 400, 416
 lithography, 6, 7, 29

lying, 22, 28, 29, 104, 201, 203, 204, 209, 211, 360

M

machine learning, xi, 186, 226
 machinery, 438
 magnetic field, 15, 18, 20, 21, 22, 23, 24, 26, 27, 61, 62, 67, 79, 86, 94, 131
 majority, 166, 370, 412
 Malaysia, 247, 284
 management, 452
 manufacture, 187, 412
 manufacturing, xi, xiii, 67, 112, 186, 192, 193, 194, 226, 245, 398, 399, 437, 438
 Markov chain, 221
 material sciences, 39
 material surface, 393
 matrix, xiii, 35, 43, 66, 93, 94, 190, 399, 400, 401, 412, 417, 418, 419, 420, 421, 427, 428
 mechanical properties, xii, 118, 188, 198, 258, 276, 277, 279, 281, 357, 358, 404
 mechanical stress, 370, 373, 408
 mechanical testing, 281
 media, 191
 melt, 113, 116, 131, 138, 145, 188, 250, 261, 392, 393, 395, 401, 403, 405, 406, 425
 melting, xiii, 6, 7, 80, 97, 112, 113, 115, 119, 131, 136, 138, 139, 142, 143, 150, 155, 183, 188, 248, 249, 250, 256, 257, 258, 273, 274, 275, 358, 394, 395, 399, 417, 429, 431, 432, 434
 melting temperature, 119, 142, 143, 150, 155, 249, 257, 258, 273, 274, 395
 melts, 135, 396, 398, 406, 425
 membership, 193, 218, 219
 memory, 166, 211, 213
 metallurgy, 7
 metals, xi, 7, 67, 70, 71, 81, 112, 116, 118, 124, 131, 140, 247, 248, 250, 255, 256, 257, 258, 266, 272, 273, 275, 276, 277, 280, 282, 283, 363, 394
 meter, 125
 methodology, 150, 164, 166, 173, 174, 190, 193, 221, 228, 229, 233, 364, 384, 415, 440
 Miami, 415, 435
 microscope, 74, 128, 129, 420
 microscopy, 434
 microstructure, 128, 188, 248, 257, 263, 265, 266, 267, 268, 273, 275, 276, 277, 280, 281, 282, 283, 368, 373, 379, 418, 419, 420, 421, 427, 432
 microstructures, 263, 275, 277, 414, 423, 425
 mixing, 87, 124, 128, 136, 256, 259, 266, 272, 283
 mobile phone, 409
 modeling, x, 78, 108, 166, 174, 181, 182, 185, 191, 194

modification, 97, 112, 114, 294, 387, 447
 modules, 213, 227, 397
 modulus, 361, 374
 moisture, 399, 404, 415
 moisture content, 399, 404
 molecules, 67, 132, 137
 momentum, 93, 94, 104
 monitoring, ix, 1, 107, 226
 Moscow, 109, 183, 285, 356
 MTS, 419
 multiplication, 104
 multiplier, 355

N

National Research Council, 213, 246
 Nd, vii, xiii, 394, 395, 401, 402, 406, 412, 417, 418, 419, 420, 425, 427, 434, 436
 neglect, 374
 Netherlands, 246, 388, 416, 451
 neural network, x, 112, 166, 167, 169, 173, 179, 193
 Neural Network Model, 165
 neural networks, x, 112, 166, 169, 179, 193
 nodes, 83, 86, 96, 232
 noise, 35, 61, 165, 451
 normal distribution, ix, 1, 31, 45, 48, 55, 57
 nucleation, 278, 424, 427
 nucleus, 427
 numerical analysis, 187, 429

O

oil, 121, 187, 190, 244, 258
 one dimension, 39
 opacity, 401
 operating parameters, 146
 optical fiber, 397
 optical microscopy, 434
 optical properties, 107, 392, 395, 399, 400, 401, 402, 405, 409, 412, 414
 optical systems, 14, 16, 18, 29, 30, 66, 74, 83
 optimization method, 150
 oscillation, 114, 115, 137
 oscillations, 15, 22, 24, 136
 overlay, 35
 oxygen, 187

P

packaging, 396, 406, 409, 410
 Pakistan, 185, 246

parallel, 16, 77, 86, 96, 124, 166, 195, 255, 267, 291, 292, 295, 298, 307, 313, 347, 371
 parallel implementation, 166
 parallelism, 29
 Pareto, 176, 177, 178, 181
 Pareto optimal, 176
 pattern recognition, 166, 226
 permission, iv, 393, 397, 398, 399, 408, 411, 412
 permit, 68, 123, 135
 permittivity, 61, 86
 phase diagram, 427
 phase transformation, 191, 257, 275, 283
 phase transitions, 145, 162
 photographs, 118
 physical and mechanical properties, 118, 258
 physical properties, 81, 97, 115, 135, 267, 276, 283, 399, 404
 physics, 15, 67
 pigmentation, 414
 pigments, 400, 413
 plastic deformation, 186, 256, 258, 266, 274, 275, 278, 281, 286, 314
 plasticity, 286, 319, 360, 362, 366, 369, 373, 452
 plasticization, 447, 448, 449, 450
 plasticizer, 405
 plastics, 392, 395, 396, 399, 400, 402, 404, 408, 409, 410, 411, 412, 413, 414, 415, 416
 platform, 123, 211, 419
 POEs, 403
 Poisson equation, 15
 Poland, 121
 polarity, 198
 polyamides, 414, 416
 polycarbonate, 402, 406, 407, 408
 polyimide, 416
 polymer, 396, 399, 400, 401, 404, 405, 408, 409, 412
 polymer chains, 404
 polymer matrix, 399, 400, 401, 412
 polymerization, 7
 polymers, 394, 398, 400, 401, 402, 408, 410, 413, 414, 415, 416
 polymethylmethacrylate, 406
 polynomial functions, 371
 polypropylene, 403, 406, 407, 413, 414
 polystyrene, 402, 407
 predictor variables, 190, 196, 198, 201, 203, 204, 205, 207, 209, 210, 211, 213, 214, 216, 217, 224, 232, 235
 prevention, 11
 probability, 5, 8, 30, 31, 34, 105, 136, 174, 221, 427
 probe, xi, 65, 247
 problem solving, 192, 195, 211
 process control, x, 111

product performance, 2
 productivity, 115, 190, 392
 prognosis, 114, 140, 180
 programming, 190, 192, 227
 project, 107
 propagation, 10, 62, 67, 108, 166, 167, 189, 358, 377, 378, 379, 384, 385, 388, 439
 proportionality, 27
 prototype, 194
 pumps, 116, 120, 123

Q

quality control, 66
 quality improvement, x, 2, 112, 115, 171, 172, 174, 177, 180, 182, 184
 quartz, 134

R

radial distribution, 5, 8, 21, 29, 132, 146, 182
 radiation, 6, 74, 80, 81, 83, 112, 116, 120, 191, 392, 393, 394, 395, 397, 400, 408, 413
 radius, 5, 11, 14, 17, 18, 19, 20, 23, 27, 61, 62, 63, 91, 92, 96, 104, 107, 139, 252, 253, 335, 343, 372
 ray-tracing, 99
 reaction temperature, 425
 reality, 11, 216, 279
 reasoning, xi, 185, 192, 193, 211, 216, 362
 recognition, 166, 226
 recommendations, iv, 225
 reconstruction, 36, 39, 40, 65, 107
 recrystallization, 256, 258, 274, 275, 278, 281, 282
 recycling, xii, 391
 redistribution, 373, 451
 reflectivity, 393
 refractive index, 401
 refractive indices, 399
 regression, 53, 150, 155, 161, 162, 172, 173, 174, 194, 403
 regression analysis, 174, 194, 403
 regression equation, 53, 162
 regression model, 150, 155, 161, 172, 173, 174
 reinforcement, xiii, 188, 399, 400, 417, 418, 420, 421, 422, 423, 424, 425, 427, 428, 434
 rejection, 392
 relaxation, xiii, 4, 97, 373, 374, 437, 443, 449, 450, 451
 relaxation process, 443
 replacement, 212, 392, 395
 replication, 182, 190
 resins, 413, 416

resistance, 81, 188, 283, 358, 359, 360, 362, 378, 380, 383, 392, 418
 resolution, 5, 29, 61, 89, 226, 231, 232, 233, 244
 room temperature, 125, 140, 142, 370, 371, 374, 438
 root-mean-square, 32
 rotations, 276
 roughness, 71, 136, 267
 Russia, 108, 109, 124, 285

S

saturation, 68, 71, 94, 96, 100
 scaling, 30
 scatter, 394, 439
 scattering, 41, 132, 137, 400, 401, 420
 self-consistency, 67
 sensitivity, 5, 11, 151, 187, 191, 451
 sensors, 33, 398, 410
 service life, 98
 shape, 28, 32, 80, 81, 87, 89, 100, 105, 113, 118, 121, 132, 135, 136, 137, 138, 139, 145, 161, 186, 188, 253, 259, 267, 270, 271, 279, 280, 290, 329, 350, 376, 386, 401, 404, 425, 438
 shear, 261, 289, 290, 292, 295, 296, 330, 346, 403
 shear strength, 403
 shipbuilding, 244
 ships, 187, 389
 shortage, 253, 255, 257
 shrinkage, 189
 signals, 39, 44, 61
 signs, 34, 36, 43
 simulation, ix, x, xiii, 2, 80, 83, 85, 86, 89, 91, 92, 97, 103, 108, 115, 121, 190, 191, 192, 417
 software, 93, 193, 201
 solid phase, 136
 solid solutions, 275
 solid state, 145, 256, 257, 261, 262, 273, 282
 solidification, 136, 137, 258, 276, 398, 438
 solvents, 406
 space charge distribution, 83, 86
 Spain, 357, 452
 specific heat, 81, 139, 142, 150, 163, 399
 specific knowledge, 193
 specifications, 227
 spectroscopy, 432
 speed of light, 30, 104, 113
 stabilizers, 399, 400
 standard deviation, 31, 42, 45, 48, 49, 66, 167
 standardization, ix, 1, 2, 66, 108
 statistics, 138
 steel, xi, 107, 123, 124, 136, 141, 143, 146, 147, 149, 150, 155, 165, 174, 178, 181, 182, 186, 190, 191, 194, 195, 198, 201, 203, 204, 205, 209, 211, 225,

242, 244, 282, 362, 378, 380, 381, 383, 386, 389, 439, 451
 storage, 66, 187, 227
 streams, 9
 stress fields, 186, 450
 stress intensity factor, 360, 361, 366, 369, 376
 structural changes, 182
 styrene, 402, 406
 subgroups, 161
 substitution, 320, 329, 338
 surface area, 148, 253, 259, 262
 surface modification, 114
 surface tension, 132, 135, 136, 137, 424
 Sweden, 451
 symmetry, 24, 29, 86, 290, 291, 294, 296, 298, 309, 319, 321, 323, 326, 330, 335, 345, 346, 347
 synchronization, 279

T

tantalum, 71, 81
 telecommunications, 74
 TEM, 418, 420, 427, 428
 temperature dependence, 139
 tensile strength, xiii, 190, 198, 281, 363, 367, 400, 401, 404, 406, 407, 417, 418, 420, 421, 424, 429
 tension, 132, 135, 136, 137, 191, 268, 269, 286, 292, 321, 364, 424
 testing, 42, 125, 126, 128, 150, 167, 195, 198, 222, 226, 281, 419, 441
 textbooks, 286
 textiles, 406, 409
 thermal deformation, 118
 thermal energy, 142, 155
 thermal expansion, 187, 276, 405, 413
 thermal properties, xi, 247, 250
 thermal treatment, 138
 thermometer, 434
 thermoplastics, xiii, 391, 392, 402, 405, 406, 413, 414, 415, 416
 thin films, 112, 121
 titanium, 401, 418, 419, 425
 total energy, 105, 112
 tracks, 86, 93
 trade-off, xi, 185, 186, 190, 214, 228
 training, 166, 167, 168, 169, 170, 171
 trajectory, ix, 2, 3, 11, 19, 21, 22, 23, 24, 25, 27, 78, 85, 87, 91, 92, 93, 97, 99, 104, 108
 transformation, ix, 1, 34, 35, 39, 40, 45, 66, 91, 93, 94, 97, 113, 191, 257, 275, 283, 320
 transformations, 35, 43, 66, 113, 191

transmission, xii, xiii, 391, 392, 393, 394, 395, 396,
399, 400, 401, 403, 404, 405, 406, 407, 408, 409,
410, 411, 412, 413, 414, 415, 416
Transmission Electron Microscopy, 419
Transmission Electron Microscopy (TEM), 419
transparency, 408
transparent medium, 396
transport, 30, 49, 67, 103, 113, 135, 136, 265, 276,
279, 280, 283, 406
transportation, 67, 108, 113, 266, 270, 272, 273, 276,
277, 278, 281
tungsten, 42, 76, 80, 81, 92, 98, 100, 116, 121, 124,
199
turbulent flows, 145
two-dimensional space, 30

U

UK, 246, 388, 389, 451
Ukraine, 121
uniform, 5, 7, 12, 23, 25, 26, 36, 72, 80, 81, 91, 97,
100, 136, 273, 359, 370, 372, 374, 398, 404, 421,
438, 441, 452
updating, 227, 244

V

vacuum, 7, 15, 17, 22, 29, 30, 60, 67, 68, 69, 72, 74,
86, 98, 113, 115, 116, 118, 120, 121, 123, 124,
132, 137, 145, 147, 148
validation, 167, 169, 170, 373
vapor, 132, 135, 137, 138, 143, 155

variations, x, 8, 102, 112, 114, 124, 135, 136, 138,
145, 151, 171, 172, 181, 189, 217, 243, 370, 384
vector, 8, 15, 18, 139, 300, 301, 311, 313, 326, 337,
347, 350
versatility, 187, 392
vessels, 187
vibration, 392, 404, 407, 409, 438
Vickers hardness, 128
virtual work, 294, 296, 307, 309

W

warehouses, 187
wavelengths, 394, 402
wear, 276, 409, 418, 434
weight ratio, xii, 391
windows, 114, 120
working conditions, 115, 131
working memory, 211, 213

X

X-axis, 46, 48, 49, 52
X-ray, 6, 74, 132, 418, 420, 432
X-ray analysis, 6, 420
X-ray diffraction, 418
X-ray diffraction (XRD), 418
XRD, 418, 421, 423, 426, 427

Y

Y-axis, 46, 48, 49, 51, 52