National Editorial Board
Tadeusz Burczyński, Instytut Podstawowych Problemów Techniki PAN, Warszawa
Marek Danielewski, AGH, Kraków
Jacek Kitowski, AGH, Kraków
Michał Kleiber, Instytut Podstawowych Problemów Techniki PAN, Warszawa
Zygmunt Kolenda, AGH, Kraków
Jan Kusiak, AGH, Kraków
Wojciech Mitkowski, AGH, Kraków
Edward Nawarecki, AGH, Kraków
Ryszard Parkitny, Politechnika Częstochowska, Częstochowa
Maciej Pietrzyk, AGH, Kraków – Chief Editor
Jerzy Rojek, Instytut Podstawowych Problemów Techniki PAN, Warszawa
Tomasz Szmuc, AGH, Kraków
Ryszard Tadeusiewicz, AGH, Kraków

International Editorial Board
Carlos Agelet de Saracibar, ETS Ingenieros de Caminos, Canales y Puertos de Barcelona, Spain
Olivier Allix, LMT de l’ENS Cachan, France
Dorel Banabic, University of Cluj-Napoca, Romania
Jose Cesar de Sa, University of Porto, Portugal
Yvan Chastel, Renault, Technocentre Guyancourt, France
Jean-Loup Chenot, CEMEF, France
Leszek Demkowicz, University of Texas, USA
Andrzej Duda, Ensimag (INPG), LSR Laboratory, France
Jacob Fish, Columbia University, USA
Anne-Marie Habraken, University of Liege, Belgium
Peter D. Hodgson, Deakin University, Australia
Yong-Taek Im, Korea Advanced Institute of Science and Technology, Korea
Reiner Kopp, RWTH Aachen, Germany
Antti Korhonen, Aalto University, Helsinki, Finland
Jianguo Lin, Imperial College, UK
Kjell Mattiasson, Chalmers University of Technology, Sweden
Fabrizio Micari, University of Palermo, Italy
Evangelos Mitsoulis, National Technical University of Athens, Greece
Tomaz Rodic, University of Ljubljana, Slovenia
Jan Sokolowski, Universite Henri Poincare, Nancy, France
(on leave from Instytut Badań Systemowych PAN, Poland)
Roberto P. Tavares, Federal University of Minas Gerais, Brasil
A. Erman Tekkaya, ATILIM University, Turkey
Ton van den Boogaard, University of Twente, Enschede, The Netherlands
Roger N. Wright, RPI, Troy, USA
Jun Yanagimoto, University of Tokyo, Japan
Jeong Whan Yoon, Swinburne University of Technology, Australia

Copyright 2015 Wydawnictwo Naukowe AKAPIT
Printed in Poland
ISSN 1641-8581

Address of the Editor
Department of Applied Computer Science and Modelling
Akademia Górnio-Hutnicza
Mickiewicza 30, 30-059 Kraków, Poland
tel/fax: +48 12 617 2921
e-mail: maciej.pietrzyk@agh.edu.pl
www.cmms.agh.edu.pl
Secretariat: Anna Baran (barana@agh.edu.pl)
Production Editor: Karolina Grodecka (karolina@agh.edu.pl)
Web page administrator: Łukasz Rauch (lrauch@agh.edu.pl)

Address of the Publisher
Wydawnictwo Naukowe AKAPIT
Komandosów 11/29
30-334 Kraków, Poland
e-mail: wn@akapit.krakow.pl
www.akapit.krakow.pl
tel: 608 024 572

The Journal is financially supported by the Polish Ministry of Science and Higher Education

DEKLARATION CONCERNING MAIN VERSION
Since the journal „Computer Methods in Materials Science” is published in traditional (printed) version and electronic (on-line) version, the Publisher informs that the printed version is the main version.
General conditions for submitting the papers

Original articles within the scope of the Journal should be submitted through Internet:

www.cmms.agh.edu.pl

as pdf file. Papers have to be submitted in English. The file with the paper should contain:

Title
Authors – give names, affiliations and addresses of the author and all co-authors.
Abstract – it must state the main objectives, scope and findings of the work within approximately 300-500 words.
Keywords – few keywords are required for use in indexing.
Main text with figures
Acknowledgements (if necessary)
Appendix (if necessary, not advised)
References

Title and abstract (300-500) words in Polish. Only Polish Authors should submit abstract in Polish. The publisher will do translation of the title and the abstract into Polish for other Authors.

All contributions will be evaluated by two reviewers selected by the Editorial Board and the Authors will be notified about the decision.

When the paper is accepted, MS Word source file with the text as well as figures in separate files have to be sent. It can be done either through internet or, if the file is too big, on the CD together with the printed version.

Manuscript preparation

Text of the paper has to be written on the A4 format with 25 mm margins on all sides, using Times-Roman 12 pts font. Maximum length of the paper is 8 pages including figures and tables. Add page numbering at the bottom-centre of the page.

Equations should be justified to the left margin and equation number (in parenthesis) to the right margin:

\( F(x) = a_1 x^2 + b_2 y^2 + c_3 x y \)  

Figures, photographs and tables have to be included in the text. They should be referred to in the text by the whole word in lower case letters, for example function \( f(x) \) is presented in figure 1 and the results of calculations are given in table 1. Use the SI-system for units and all quantitative data. When the paper is accepted for publication, all figures have to be supplied electronically as *.bmp or *.jpg files. If requested, figures or photographs can be printed in colour for an additional cost of 40 euro per page paid by the authors.

List of references is located at the end of the text, in the alphabetic order. Reference to the list has to be made by the name of the authors and the year of publication, for example:

In the case of one author
... developed by Huber (1904).
... energetic criterion was applied (Huber, 1904).

In the case of two authors:
... on the basis of research performed by Thompson and Hamzeh (1992).
... energetic criterion was applied by Thompson and Hamzeh (1992)

In the case of more than two authors:
Bai et al. (1993) researched the effect of . . .

Parameters for the bearing steel are also available (Bai et al., 1993).

In the case of referring to more than one publication:
(Thompson & Hamzeh, 1992; Chenot & Bellet, 1992)

If one author of a group of authors published more than one cited paper in the same year, add letters a, b, . . after the year.

Examples of formatting references for books, journal publications, conference proceedings and PhD theses are presented below. Name of the publisher as well as names of all editors have to be added for edited books or conference proceedings. 1 cm indentation should be used.

Examples:


Huber, M. T., 1904, Właściwa praca odkształcenia jako miara wyteżenia materiału, Czasopismo Techniczne, 22, 38-81 (in Polish).


Copyright

For publication in the journal Computer Methods in Materials Science an article may not have been published previously anywhere else. Name of the author as well as names of all editors have to be added for edited books or conference proceedings. 1 cm indentation should be used.

Examples:


Huber, M. T., 1904, Właściwa praca odkształcenia jako miara wyteżenia materiału, Czasopismo Techniczne, 22, 38-81 (in Polish).


Correspondence

Halina Kusiak
Faculty of Metals Engineering and Industrial Computer Science
Akademia Górniczo-Hutnicza
Mickiewicza 30, 30-059 Krakow, Poland
e-mail: hkusiak@metal.agh.edu.pl

Subscription

All information about conditions of subscription can be found at: www.cmms.agh.edu.pl
# Contents

**Spis treści**

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gerhard Hirt, Markus Bambach, Johannes Lohmar, Onur Güvenç, Thomas Henke, Gideon Schwich</td>
<td></td>
</tr>
<tr>
<td>MATERIALS MODELLING IN INDUSTRIAL BULK METAL FORMING PROCESSES AND PROCESS CHAINS</td>
<td>5</td>
</tr>
<tr>
<td>Bert Geijselaers, Ton Bor, Peter Hilkhujsen, Ton Van Den Boogaard</td>
<td>13</td>
</tr>
<tr>
<td>CYCLIC SHEAR BEHAVIOR OF AUSTENITIC STAINLESS STEEL SHEET</td>
<td></td>
</tr>
<tr>
<td>Alexander Nam, Uwe Prüfert, Michael Eiermann, Rudolf Kawalla</td>
<td>23</td>
</tr>
<tr>
<td>CONCEPT OF THERMAL MODELLING FOR HOT STRIP ROLLING OF MAGNESIUM</td>
<td></td>
</tr>
<tr>
<td>Łukasz Morawiński, Andrzej Kocańda</td>
<td>30</td>
</tr>
<tr>
<td>INFLUENCE OF TOOL GEOMETRY ON SURFACE CONDITION OF V-BENT ALUMINUM SHEET ....</td>
<td></td>
</tr>
<tr>
<td>Julita Winowiecka, Piotr Lacki</td>
<td>37</td>
</tr>
<tr>
<td>CALCULATION OF FORMING LIMIT CURVE FOR GRADE 2 TITANIUM USING MODIFIED SAMPLE Geometry</td>
<td></td>
</tr>
<tr>
<td>Elena Silvestre, Eneko Sáenz De Argandoña, Lander Galdos, Joseba Mendiguren</td>
<td></td>
</tr>
<tr>
<td>NUMERICAL SIMULATION OF THE ROLL LEVELLING OF DP1000 STEEL USING A NONLINEAR COMBINED HARDENING MATERIAL MODEL</td>
<td>44</td>
</tr>
<tr>
<td>Eneko Sáenz De Argandoña, Lander Galdos, Rafael Ortubay, Joseba Mendiguren, Xabier Agirretxe</td>
<td>51</td>
</tr>
<tr>
<td>W-TEMPER FORMING OF AA7075 ALUMINUM ALLOYS AS AN ALTERNATIVE TO THE WARM AND HOT STAMPING</td>
<td></td>
</tr>
<tr>
<td>Lander Galdos, Eneko Sáenz De Argandoña, Joseba Mendiguren, Rafael Ortubay, Xabier Agirretxe, José Miguel Martin</td>
<td>58</td>
</tr>
<tr>
<td>DETERMINATION OF HEAT TRANSFER COEFFICIENTS UNDER CLOSED LOOP CONTROLLED CONSTANT CONTACT PRESSURES</td>
<td></td>
</tr>
<tr>
<td>Marek St.Węglowski</td>
<td></td>
</tr>
<tr>
<td>ARTIFICIAL NEURAL NETWORKS AND RESPONSE SURFACE METHODOLOGY AS A TOOL FOR ANALYSIS THE SPINDLE TORQUE IN FSP PROCESS</td>
<td>65</td>
</tr>
<tr>
<td>Krzysztof Regulski, Gabriel Rojek, Dorota Wilk-Kołodziejczyk, Stanisława Kluska-Nawarecka, Edward Nawarecki, Grzegorz Dobrowolski</td>
<td>71</td>
</tr>
<tr>
<td>INTELLIGENT INTERFACE FOR DECISION SUPPORT SYSTEM IN METALLURGICAL DOMAIN ....</td>
<td></td>
</tr>
<tr>
<td>Gabriel Rojek, Krzysztof Regulski, Piotr Jarosz, Jacek Gabryel, Jan Kusiak</td>
<td>78</td>
</tr>
<tr>
<td>CONTROL OF LEAD REFINING PROCESS WITH THE USE OF CASE-BASED REASONING APPROACH</td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>Pages</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Extracting Knowledge From Integrated Experimental Data on the ADI</td>
<td>85</td>
</tr>
<tr>
<td>Manufacture</td>
<td></td>
</tr>
<tr>
<td>Optimization of Geometrical Structures of Modern Materials by Using</td>
<td>94</td>
</tr>
<tr>
<td>A Hybrid Evolution Strategy</td>
<td></td>
</tr>
<tr>
<td>Hybrid Parallel Evolutionary Algorithm in Optimization of 2D</td>
<td>103</td>
</tr>
<tr>
<td>Grapehene-Like Materials</td>
<td></td>
</tr>
<tr>
<td>Vibhu Trivedi, Manojkumar Ramteke</td>
<td></td>
</tr>
<tr>
<td>Multi-Objective Optimization of Phthalic Anhydride Catalytic Reactor</td>
<td>111</td>
</tr>
<tr>
<td>Using Genetic Algorithm with Simulated Binary Jumping Genes Operator</td>
<td></td>
</tr>
<tr>
<td>Sunil Kumar Jauhar, Millie Pant, Mahesh C. Nagar</td>
<td>118</td>
</tr>
<tr>
<td>Differential Evolution for Sustainable Supplier Selection in Pulp</td>
<td></td>
</tr>
<tr>
<td>and Paper Industry: A DEA Based Approach</td>
<td></td>
</tr>
<tr>
<td>Renu Tyagi, Millie Pant, Yuvraj Singh Negi</td>
<td>127</td>
</tr>
<tr>
<td>Optimizing the Performance of Electrically Poled Polymeric Films</td>
<td></td>
</tr>
<tr>
<td>Konrad Adamus, Piotr Lacki</td>
<td>137</td>
</tr>
<tr>
<td>The Numerical Analysis of a Titanium Sheets Welding Process and</td>
<td></td>
</tr>
<tr>
<td>Welding Joint Tensile Behavior</td>
<td></td>
</tr>
<tr>
<td>Anna Derlatka, Piotr Lacki</td>
<td></td>
</tr>
<tr>
<td>Numerical Investigation of Refill Friction Stir Spot Welding Joints</td>
<td>144</td>
</tr>
<tr>
<td>James D. Pollard, Andrew Watford, Martin Jackson, Bradley P. Wynne</td>
<td></td>
</tr>
<tr>
<td>The Modelling of Ring Tests at Elevated Temperatures for the</td>
<td></td>
</tr>
<tr>
<td>Determination of Friction in Ti-6Al-4V Forgings</td>
<td>150</td>
</tr>
<tr>
<td>Michał Gzyl, Andrzej Rosochowski, Lech Olejnik, Kamil Sikora,</td>
<td></td>
</tr>
<tr>
<td>Muhammad Jawad Qarni</td>
<td></td>
</tr>
<tr>
<td>Determination of Friction Factor by Ring Compression Testing and</td>
<td></td>
</tr>
<tr>
<td>FE Analysis</td>
<td></td>
</tr>
<tr>
<td>L. Chen, T. Clausmeyer, A.E. Tekkaya</td>
<td>156</td>
</tr>
<tr>
<td>Experimental Characterization and Numerical Modeling of the</td>
<td></td>
</tr>
<tr>
<td>Mechanical Behavior of Half Sandwich Laminate in the Context of</td>
<td></td>
</tr>
<tr>
<td>Blanking</td>
<td>162</td>
</tr>
<tr>
<td>Tomasz Nowak</td>
<td></td>
</tr>
<tr>
<td>Homogenization of Fiber Metal Laminate Structures Characterized by</td>
<td></td>
</tr>
<tr>
<td>Orthotropic and Elastic-Plastic Material Models</td>
<td>169</td>
</tr>
<tr>
<td>Bartek Wierzba, Patrycja Wierzba</td>
<td></td>
</tr>
<tr>
<td>Competition Between Kirkendall Shift and Frenkel Effect During 2D</td>
<td>176</td>
</tr>
<tr>
<td>Diffusion Process</td>
<td></td>
</tr>
<tr>
<td>Henryk Adrian, Przemysław Marynowski, Dariusz Jędrzejczy</td>
<td></td>
</tr>
<tr>
<td>Calculation of the Fe-Fe₃C Phase Equilibrium Diagram</td>
<td>179</td>
</tr>
<tr>
<td>Grzegorz Korpala, Rudolf Kawalla</td>
<td></td>
</tr>
<tr>
<td>Optimization and Application of GPU Calculations in Material</td>
<td>185</td>
</tr>
<tr>
<td>Science</td>
<td></td>
</tr>
<tr>
<td>Paweł Lipski, Maciej Paszyński</td>
<td></td>
</tr>
<tr>
<td>Multi-Frontal Parallel Direct Solver for One Dimensional</td>
<td>192</td>
</tr>
<tr>
<td>Isogeometric Collocation Method</td>
<td></td>
</tr>
<tr>
<td>Jan Sladek, Vladimir Sladek, Milan Zmindak, Slavomir Hrcek</td>
<td>198</td>
</tr>
<tr>
<td>Crack Analyses in Conducting and Non-Conducting Piezoelectric</td>
<td></td>
</tr>
<tr>
<td>Solids</td>
<td></td>
</tr>
<tr>
<td>Marta Oleksy, Witold Cecot</td>
<td>204</td>
</tr>
<tr>
<td>Application of the Fully Automatic HP-Adaptive FEM to Elastic-</td>
<td></td>
</tr>
<tr>
<td>Plastic Problems</td>
<td></td>
</tr>
</tbody>
</table>
Marcin Łoś, Maciej Paszyński, Lisandro Dalcin, Victor Calo
DEALING WITH PERIODIC BOUNDARY CONDITIONS FOR 1D, 2D AND 3D ISOGEOMETRIC
FINITE ELEMENT METHOD ............................................................................................................... 213
Jerzy Rojek, Szymon Nosewicz, Katarzyna Pietrzak, Marcin Chmielewski
EVALUATION OF MACROSCOPIC STRESSES IN DISCRETE ELEMENT MODELS OF
SINTERING PROCESSES .................................................................................................................. 219
Dimitri Claudel, Jean-Claude Gelin, Mohamed Sahli, Thierry Barrière
METHOD TO IDENTIFY RHEOLOGICAL CONSTITUTIVE MODEL ADAPTED FOR POWDER
INJECTION MOULDING PROCESS USING INVERSE METHOD .................................................... 226
Alexandre Royer, Jean-Claude Gelin, Thierry Barrière
CHARACTERIZATION BY INFRARED SPECTROSCOPY OF BINDER BASED ON
POLYETHYLENE GLYCOL AND INCONEL 718 FEEDSTOCK FOR POWDER INJECTION
MOULDING ........................................................................................................................................ 232
Zbigniew Gronostajski, Sławomir Polak, Bartosz Bartczak
MODELLING OF CLINCHING JOINT PULL-OUT TEST ..................................................................... 239
Mikhail Petrov, Yulian Philippov, Pavel Petrov
NUMERICAL RESEARCH ON BRAKE CALLIPER PISTON’S WEIGHT REDUCTION ...................... 244
Wojciech Szymański, Marzena Lech-Grega, Maciej Gawlik, Adam Kokosza, Adam Chochorowski
MEASUREMENT OF RESIDUAL STRESSES IN HOT-ROLLED STEEL SHEETS FOR LASER
CUTTING ......................................................................................................................................... 251
Chieh-Yang Chen, Yiming Li
DEVICE SIMULATION AND MULTI-OBJECTIVE GENETIC ALGORITHM-BASED
OPTIMIZATION OF GERMANIUM METAL-OXIDE-SEMICONDUCTOR STRUCTURE .................... 258
MATERIALS MODELLING IN INDUSTRIAL BULK METAL FORMING PROCESSES AND PROCESS CHAINS

GERHARD HIRT*, MARKUS BAMBAKH, JOHANNES LOHMAR, ONUFRIY GÜVENÇ, THOMAS HENKE, GIDEON SCHWICH

Institute of Metal Forming, RWTH Aachen University, Intzestrasse 10, D-52056 Aachen, Germany

*Corresponding author: hirt@ibf.rwth-aachen.de

Abstract

Bulk metal forming processes range from processes with a single deformation step such as certain closed-die forging operations to processes with many subsequent stages such as hot rolling, ring rolling or open die forging. Modelling of these manufacturing processes requires both precise process models as well as adequate material models. Microstructure evolution by recrystallization is decisive in all of these processes since the microstructure determines the flow stress and hence the forming forces but it also influences the product properties. In this context, the propagation of variations in the processing conditions and in the material behavior are of special importance and methods for the quantification of uncertainties and their effect on model predictions are required. Such questions can be approached using models of different complexity on various scales as shown in the following examples: In closed die forging of a gear wheel from 25MoCr4 alloy the complex geometry requires a Finite Element process model which in this case is combined with a JMAK type material model. In plate rolling a simplified process model can be applied successfully. Based on the slab theory, which is enhanced for spatial resolution of shear strain using a meta model derived by FEM, this model can simulate even longer roll pass schedules within seconds and offers the possibility to combine it with numerical optimization techniques. Recrystallization of a high-manganese steel in interpass times between hot rolling passes is an example where models with spatial resolution (CP-FEM and phase field) are combined on the micro-scale to predict the recrystallization kinetics based on physically meaningful variables such as grain boundary mobility. In ring rolling the process model must include the closed-loop control system of the rolling machine to achieve a realistic prediction of the process kinematics. Feedback control loops for up to eight kinematic degrees of freedom (velocities and positions of all radial, axial and guiding rolls) have been defined using virtual sensors integrated in the simulation. Offline coupling with microstructure simulation is used to predict the final grain size and determine under which conditions static recrystallization occurs during the rolling sequence.

Keywords: microstructure evolution, bulk metal forming, recrystallization, uncertainty quantification

1. INTRODUCTION

The importance of numerical simulation has increased considerably in recent years, especially in hot working. A primary driver for the widespread use of numerical simulation are economic and environmental requirements such as increasing cost pressure, shorter development cycles, and the increasing need to reduce material consumption and preserve resources. On the macroscopic scale, finite element models prevail in hot working. They provide both spatial and temporal resolution of field variables such as temperature, velocity, strain, strain rate and stress, but they also yield global properties like forming forces. Hot working operations such as open-die forging, flat or ring rolling impose multiple deformation passes on the workpiece, which are separated by interpass times or deliberate heating cycles. The microstructure undergoes tremendous changes in such processes. To analyse, design or improve hot working processes using finite element
COMPUTER METHODS IN MATERIALS SCIENCE

Grained microstructure is obtained in the gear wheel can be accomplished only if a homogeneous, fine-grained microstructure is obtained in the gear wheel. Heat treatment, high strength and fatigue resistance, mechanical properties are determined largely by the toughness and fatigue resistance. Although the final face hardness, static strength, root bearing capacity, to provide mechanical properties such as high surface hardness, required for ring rolling is presented.

Phase-field models. In addition, the coupling of ing a sequentially coupled crystal-plasticity and spatially resolved microstructure evolution model.

Spatial resolution of the microstructure have been developed to allow for a deeper insight into the local phenomena governing microstructure evolution during and after hot working. To allow for predictive simulations of processes which crucially depend on feed-back control in reality, coupling of process models with the control algorithms of the real process is necessary, which further increases the complexity of metal forming simulations. This paper details applications of JMAK models and spatially resolved microstructure evolution models to selected hot working processes such as plate or ring rolling, as well recent work on modeling static recrystallization using a sequentially coupled crystal-plasticity and phase-field models. In addition, the coupling of closed-loop feed-back control with an FEM model for ring rolling is presented.

2. MODELING GRAIN SIZE EVOLUTION IN GEAR WHEEL FORGING INCLUDING UNCERTAINTY

Hot-forged gear wheels are widely used, e.g., in drive systems of heavy-duty equipment. They have to provide mechanical properties such as high surface hardness, static strength, root bearing capacity, toughness and fatigue resistance. Although the final mechanical properties are determined largely by the heat treatment, high strength and fatigue resistance can be accomplished only if a homogeneous, fine-grained microstructure is obtained in the gear wheel after hot forging. When only the hot forging process is considered, the process chain essentially consists of heating, forming and cooling. The hot forging process is affected by uncertainties in the material behavior, the process boundary conditions and the processing conditions. For robust design of the gear wheel forging process, it is necessary to quantify all uncertainties contained in the model input and trace them to the model predictions such as the final grain size. Uncertainty propagation for the gear wheel forging was discussed in detail in a previous publication by Henke et al. (2013). It was shown that a certain resampling method can be used to reveal the probability distribution of the final average grain size after hot forging and cooling, thus providing information about the uncertainty of the model predictions instead of providing only a single deterministic value. The basic model used by Henke et al. (2013) was a standard Finite Element model for the forging process which was coupled to a JMAK microstructure evolution model.

In this paper, the results published previously are enriched with more recent results. The main mechanisms creating an unwanted microstructure are abnormal grain growth and incomplete dynamic recrystallization. The occurrence of these phenomena is largely controlled by the initial temperature and the tool velocity. The billet temperature changes by heat transfer and dissipation, which are taken into account in the model. At lower forging temperatures, a finer grain size can be obtained but the risk of tool fracture increases. Varying the initial billet temperature and the tool velocity thus allows for identifying the process window shown in figure 1b which is bounded by abnormal grain growth at higher temperatures and incomplete DRX and possible tool fracture towards lower temperatures. The process simulation in figure 1a yields the maximum temperature after dissipation, which is used to create the map in figure 1b, as well as information on the completion of DRX and the occurrence of abnormal grain growth. An elastic simulation of the tool load reveals the possibility of tool fracture. The approach makes it possible to determine safe process conditions (grey dot in figure 1b), which yield the desired part with a homogeneous microstructure. The model validation with experimental results shown in figure 1c reveals a good conformance of model predictions and experimental results.
3. FAST MATERIALS AND PROCESS MODELLING FOR PLATE ROLLING

Plate rolling is a flexible bulk metal forming operation where a slab with up to 900mm in thickness is rolled down to the desired thickness in ~5 to 40 passes. Some of the passes are used to control the final width of the product. In principle, plate rolling can be modelled by a full-scale FEM model, which is very time consuming for pass schedules with many passes. Alternatively, the “slab method” (see e.g. Sims (1954)) for the fast prediction of rolling forces used in this paper can be used. It is based on the assumption that the velocity in rolling direction only depends on the current position in the roll gap but is constant in thickness direction. Sims also proposed to introduce a geometric correction term that compensates the model inaccuracies introduced by the simplified assumptions. For hot rolling it is crucial to integrate a thermal description of the rolling process and additionally consider the microstructural changes. Still neglected even in recent fast rolling models is the fact that friction and the roll gap geometry produce shear strains that can influence the flow stress and the microstructure evolution. Kawalla and Schmidtchen (2013) addressed this issue using a so called “Layer Model”. Another approach recently proposed by Seuren et al. (2014) is to extract the shear strain profiles from FEM simulations and integrate them into a fast rolling model based on the slab method. A comparison of the equivalent strain profiles calculated by the FEM and the fast rolling model for a thick and a thin slab is shown in figure 2a. The resulting through-thickness profiles differ for thick and thin slabs due to roll gap geometry, but the agreement between FEM and the fast rolling model is excellent. The additional information about the deformation gained from the shear strain model can be combined with a local temperature distribution of the roll stock that is obtained via an FDM model. Hence, it is possible to predict the microstructure in through thickness direction over all roll passes. In figure 2b it is clearly visible that the different forming conditions over the thickness of the slab have a notable influence on the roll stock’s microstructure. For an extended description of the fast rolling model capabilities for microstructure prediction as well as the modelling see Lohmar et al. (2014b).

All models for bulk metal forming depend on material parameters to match the materials behaviour during and after forming. To determine these parameters conventionally lab-scale material characterization tests are used to obtain flow curves and recrystallization kinetics. Those results are then used to fit the parameters for the material currently under
Fig. 2. Comparison of the equivalent strain in the roll gap obtained from FE-simulation and the fast rolling model for thick and thin plate (a), grain size evolution over the roll schedule and in through-thickness direction of the roll stock (b).

consideration directly. Szeliga et al. (2002) suggested replacing the direct evaluation by an inverse analysis of the basic laboratory tests using FE technique to increase accuracy. Still using the conventional approach for a multitude of different materials e.g. for over 200 steel grades typically found in plate rolling is infeasible due to the high experimental and financial effort. Here inverse techniques can also help as industrial field data, i.e. roll forces, torques and temperatures which are typically collected and archived in a rolling mill, can serve as reference data to calibrate the material model parameters and hence material testing becomes obsolete.

In this concept the deviation between the roll force prediction by the fast rolling model using some arbitrary parameters and the actual measurement can serve as indicator to evaluate the precision of the material model parameters. Reducing the roll force deviation of several slabs at a time by optimizing the parameters thus directly corresponds to improving the material model. Technically speaking the initial material model parameters are fed into an optimization loop as depicted in figure 3a. This loop is interrupted once the deviation of calculation and measurement drops below a cut off value and the optimal parameters are returned. Using this concept covered in detail in publications by Lohmar et al. (2014a, 2014c) only the computational costs rise if the number of considered materials is increased. By combining the inverse approach and the fast rolling model, the results shown in figure 3b can be obtained. Here, a cluster of different micro-alloyed low carbon steel grades with in total 88,000 roll passes was used as input. The resulting optimal parameter set is nevertheless able to retrieve the material model’s hardening and softening behaviour with high accuracy (the mean deviation between predicted and measured forces is < 7%). The completion of the
optimization run with its ~30 iterations on a state of the art workstation took little more than 21 hours.

4. MICROSTRUCTURE MODELING WITH SPATIAL RESOLUTION

In the examples shown so far, only JMAK models where used which do not take into account the spatial arrangement of grains and the orientation and neighborhoods of individual grains. Models for microstructure evolution with spatial resolution include cellular automata, Monte-Carlo-Potts, phase field models and vertex models. There is a vast amount of literature available dealing with the methods detailed above and their application to the simulation of dynamic and static recrystallization after hot working as well as SRX after cold working. The restrictions of the above-mentioned models for the simulation of a single deformation step and concurrent DRX or subsequent SRX have been investigated in depth in the past. Phase field modeling of SRX after hot deformation was recently analyzed by Güvenç et al. (2013, 2014), who used a crystal plasticity (CP)-FEM model to generate the input for the phase field simulation. The initial and deformed state of an RVE were simulated using the CP-FEM framework DAMASK developed at the MPIE Düsseldorf, Germany. The results of the CP-FEM are used as input to the phase field simulation to analyse the resulting evolution of the microstructure during SRX.

Although close, the grain morphology from the 2D simulation shown in figure 4 does not reproduce experimental findings precisely. If the simulation is performed with the goal that both the SRX kinetics and the grain morphology match the experiments, it is due to the lack of nucleation sites in 2D compared to 3D that this task cannot be accomplished. The SRX kinetics obtained from the phase-field simulation are also shown in figure 4. The experimentally obtained kinetics can be easily approximated by JMAK kinetics, but the phase field simulation gives only a rough approximation of the kinetics. Due to the large computation time of the phase field simulation even in the 2D case, there is no possibility to determine the parameters of the phase field model such as mobilities inversely. At present, it has to be stated that a full 3D simulation would be necessary to obtain more accurate results. Simulations by Zhu and Militzer (2012) with 3D phase field simulations were reported to match well with simulated results, showing that a 3D approach is indeed more suitable. However, the simulations were only performed for SRX after cold rolling, and solutions for recrystallization during and after hot working are still needed.

5. FEEDBACK CONTROL AND MICROSTRUCTURE MODELING IN RING ROLLING

In FE simulation of ring rolling processes, the model must include the machine’s closed-loop control system to achieve a realistic prediction of the process kinematics. The process holds at least eight degrees of freedom (figure 5). These degrees of freedom are adaptively adjusted by closed-loop control systems according to preselected control strategies and actual process variables (sensor values).

Until recently, in contrast to the actual process, most simulations of ring rolling processes were carried out based on kinematic data of all degrees of freedom.
freedom recorded from experiments. Simulations in the process layout stage are conducted using different simplifying approaches which can be found in literature cf. Li et al. (2008), Forouzan et al. (2003) and Wang et al. (2010). Since most of these approaches do not use the real ring rolling mill’s characteristics, deviations in predicting the process kinematics by simulation may occur. For an accurate process simulation in the layout stage, Jenkouk et al. (2012) integrated all sensors and actuators used in real ring rolling mills in an Abaqus/explicit FE model (figure 5). Also, parts of the technological controller CARWIN, which was provided by SMS Meer as a pre-compiled object file (black box to the user), were integrated. Due to the fact that the industrial control algorithms are unknown to the user, they cannot be adjusted for the development of new ring rolling strategies. Hence, as an alternative concept, a new control function was developed, cf. Hirt et al. (2014). In a self-designed controller, the initial parameters are entered in a Matlab GUI which is similar to the real ring rolling mill’s interface. The tool motions are controlled based on the parameters and the machine limits in regard to actual roll power, torque and force during the entire simulation.

However, the accuracy of FE simulations strongly depends on the input data. Especially the evolution of the workpiece’s microstructure (grain size and recrystallized fraction) as well as the proper approximation of boundary conditions is of high importance. Postprocessing of tracked points with microstructure simulation can be used to predict the workpiece’s final grain size and determine under which conditions recrystallization occurs during rolling. Henke et al. (2013) investigated the influences of dynamic and static recrystallization of an AISI 304 steel ring on rolling forces and geometry. Due to the comparatively slow SRX behavior of AISI 304, in this case static softening does only occur to a small extent between the deformation steps throughout the process. Therefore, only minor influences on rolling forces were found. Hence, the use of tabular flow stress values, determined in cylindrical compression tests leads to good correlation between experiment and FEA in calculated microstructure, rolling forces and material flow, respectively geometry (Henke et al., 2013).

Still, the approximation of boundary conditions is of high importance. Schwich et al. (2014) investigated the influence of boundary conditions on roll forces and geometrical values as well as on microstructure. Influences of transfer time from furnace to the ring rolling mill, emissivity, heat transfer and friction on the outer diameter, the maximum force, the force after 20 seconds and workpiece temperatures after rolling and a 300 second cooling period were examined. Therefore, a simulation study using a design of experiments (DOE) approach was conducted. Forming histories of one tracked point per simulation where compared (figure 6a) and a wide span of resulting forces was found (figure 6b). The transfer time was found to be the most influencing parameter whereas friction seemed to have minor influence on geometry as well as forces (figure 6c). Furthermore, influences on recrystallized fraction and grain size were examined. Throughout the pro-
6. CONCLUSION

The results shown in this paper can be grouped into four kinds of coupled models: (i) coupled process and JMAK microstructure models (gear wheel), (ii) a fast rolling model coupled to optimization (plate rolling), (iii) a sequentially coupled model for deformation and static recrystallization on the micro-level and (iv) a process-control coupling for ring rolling. The examples show that the complexity of simulations of bulk metal forming processes increases, and the success of the simulations does not only depend on the accuracy of individual models such as the process and the material model, but also on their interaction and interplay with further models along the periphery of the process.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support of the Deutsche Forschungsgemeinschaft (DFG) within the Cluster of Excellence “Integrative Production Technology for High Wage Countries” and the Collaborative Research Center SFB 761 “Steel – ab initio”. They are also grateful to “AG der Dillinger Hüttenwerke” and “SMS Meer” for the cooperation and support.

REFERENCES


MODELOWANIE MATERIALÓW
W PRZEMYSLOWYCH, OBJĘTOŚCIOWYCH
PROCESACH PRZERÓBKI PLASTYCZNEJ
I ŁAŃCUCHACH PRODUKCYJNYCH

Streszczenie

Objętościowe procesy przeróbki plastycznej to zarówno procesy składające się z jednej operacji odkształcania, na przykład kucia na matrycy zamkniętej, jak i procesy składające się z kilku kolejno po sobie następujących operacji, na przykład proces walcowania na gorąco, proces walcowania pierścieni czy kucia na matrycy otwartej. Modeleowanie numeryczne tego rodzaju procesów wymaga dokładnego ich opisu, w tym zastosowania odpowiednich modeli materiałowych. Należy podkreślić, że zmiany mikrostruktury podczas rekrysalizacji są w tym wypadku ważne, ponieważ mikrostruktura bezpośrednio przełada się na wartość naprężenie uplastyczniającego i siły w procesie, a także decyduje o własnościach produktu. Ewentualne zmiany warunków procesu mają wpływ na zachowanie się materiału i ich uwzględnienie jest konieczne. Wymaga to metod pozwalających na jakościową ocenę niewspółmiernych oraz ich wpływu na wyniki modelowania. W pracy przedstawiono procesy, dla których zastosowano modele o różnych złożonościach i różnych skalach. Pierwszym przykładem jest kucie koła zębatego ze stopu 25MoCr4 w zamkniętej matrycy. Skomplikowany kształt kutej części wymagał użycia metody elementów skończonych i zastosowania modelu materiału uwzględniającego przemiany fazowe (JMMAK). Drugim przykładem jest walcowanie blach, które można modelować w sposób uproszczony, wykorzystując rozwiązanie różniczkowego równania równowagi i uwzględniające naprężenie ściśające dzięki zastosowaniu meta modelu opracowanego na podstawie obliczeń MES. Tak zdefiniowanym modelem symulowano w krótkim czasie proces walcowania nawet długo, co pozwoliło na wykorzystanie modelu w zadaniach optymalizacyjnych. Zjawisko rekrysalizacji dla stali wysoko-manganowych zachodzące pomiędzy przepustami walcowania na gorąco jest przykładem, gdzie model CF PEM, czy pola faz, są połączone w skali mikro. Dzięki temu możliwe było przewidywanie kinetyki rekrysalizacji z wykorzystaniem zmienności o charakterze fizycznym, takich jak mobilność granicy ziarna. Z kolei dla walcowania pierścieni, aby otrzymać zgodną z rzeczywistością kinematykę procesu, model musi uwzględniać zamknięty system sterowania walcarką. Zdefiniowano osiem kinematycznych stopni swobody dla tego systemu (prędkości oraz położenia wszystkich rolów promieniowych, osiowych oraz pozycjonujących) poprzez zastosowanie wirtualnych sensorów zintegrowanych z modelem procesu. Przewidywanie końcowej wielkości ziarna oraz wyznaczenie warunków, w których zachodzi statyczna rekrysalizacja podczas sekwencji walcowania, otrzymano poprzez połączenie wyników modelu z wynikami symulacji mikrostruktury.

Received: October 21, 2014
Received in a revised form: October 21, 2014
Accepted: November 18, 2014
CYCLIC SHEAR BEHAVIOR OF AUSTENITIC STAINLESS STEEL SHEET

BERT GEIJSELAERS*, TON BOR, PETER HILKHULISEN, TON VAN DEN BOOGAARD

Universiteit Twente, Engineering Technology,
POBox 217, 7500AE Enschede, Netherlands
*Corresponding author: h.j.m.geijselaers@utwente.nl

Abstract

An austenitic stainless steel has been subjected to large amplitude strain paths containing a strain reversal. During the tests, apart from the stress and the strain also magnetic induction was measured to monitor the transformation of austenite to martensite. From the in-situ magnetic induction measurements an estimate of the stress partitioning among the phases is determined.

When the strain path reversal is applied at low strains, a classical Bauschinger effect is observed. When the strain reversal is applied at higher strains, a higher flow stress is measured after the reversal compared to the flow stress before reversal. Also a stagnation of the transformation is observed, meaning that a higher strain as well as a higher stress than before the strain path change is required to restart the transformation after reversal.

The observed behavior can be explained by a model in which for the martensitic transformation a stress induced transformation model is used. The constitutive behavior of both the austenite phase and the martensite is described by a Chaboche model to account for the Bauschinger effect. In the model mean-field homogenization of the material behavior of the individual phases is employed to obtain a constitutive behavior of the two-phase composite. The overall applied stress, the stress in the martensite phase and the observed transformation behavior during cyclic shear are very well reproduced by the model simulations.

Key words: metastable austenite, deformation induced martensite, constitutive model

1. INTRODUCTION

Transformation of retained austenite under mechanical loading is especially prominent in austenitic stainless steel. Under the right circumstances, the metastable austenite transforms to martensite under mechanical loading. For recent experimental studies see for example Lebedev and Kosarchuk (2000), Nagy et al. (2004), Post et al. (2008).

Austenitic stainless steels have a broad range of applications. In general, they have high corrosion resistance, high cryogenic toughness, high work hardening rate, high hot strength, high ductility, high hardness, an attractive appearance and low maintenance. The delayed cracking of stainless steel products is in general attributed to the presence of martensite combined with residual stress (Berrahmoune et al., 2006). For the prediction of martensite fraction and residual stresses it is important to have reliable models.

Olson and Cohen (1975) formulated a kinetic model which explains the martensite formation from ε-phase nucleation on shear band intersections during plastic deformation Venables (1962). This strain induced kinetic model for martensitic phase transformation has been combined by Stringfellow et al. (1992) with a mean-field homogenization model to obtain overall visco-plastic behavior from the constitutive behavior of the individual phases. Also the influence of the stress state and transformation plasticity were added. Further extensions have been provided by Tomita and Iwamoto (1995) for strain rate dependence and by Diani and Parks (1998) for crys-
tual plasticity. Han et al. (2004) added stress dependence by evaluating the mechanical driving force on individual martensite variants. This enabled them to calculate the texture of the resulting martensite.

An alternative theory for mechanically induced martensite formation was proposed by Tamura (1982). In his model the driving force of the applied stress is considered as the reason for the transformation. See also Perdahcioğlu et al. (2008a). When the thermodynamic driving force as defined by Patel and Cohen (1953) exceeds a threshold value, the transformation will start. Applications of stress induced transformation models suitable for macro scale simulations have been presented by Hallberg et al. (2007) and Perdahcioğlu and Geijselaers (2012) for austenitic steel and by Lani et al. (2007), Delannay et al. (2008) and Kubler et al. (2011) for TRIP steel.

For accurate prediction of the state of the material after forming, it is important that the non-proportional deformation behaviour is captured correctly. Very few studies of the large amplitude cyclic and non-proportional response of metastable austenitic stainless steel are available in literature. An extensive experimental program, including tension-compression tests, was conducted by Spencer et al. (2009) on austenitic steel. They report a strong Bauschinger effect in the austenite stress-strain response. Results from cyclic shear tests and tensile tests followed by shear tests were presented by Gallée et al. (2007). They formulated a model based on Stringfellow et al. (1992). Hamasaki et al. (2014) showed that observations during large amplitude cyclic tension-compression tests cannot be captured by the strain induced transformation model.

In this paper we report on cyclic shear tests, which have been conducted on a low Carbon 12Cr9Ni4Mo austenitic stainless steel. During the testing the martensite transformation was monitored by a magnetic induction sensor.

A constitutive model of austenitic steel which undergoes a mechanically induced transformation will be presented, where the martensitic transformation is modeled as a stress-driven process similar to the model of Tamura (1982). This transformation model is then combined with a mean-field formulation for description of the constitutive behavior of the two-phase composite.

2. EXPERIMENTS

The material used in the tests is 12Cr9Ni4Mo austenitic stainless steel. Its nominal composition is given in table 1. Specimens were cut from 0.5 mm thick sheet as described in Perdahcioğlu et al. (2008b) for deformation in shear, which was applied at a rate of approximately 0.001 s⁻¹. The strain was measured real-time on the material surface using a camera and dot-tracking software. Dots were applied to the specimen surface before the test and the corresponding positions were recorded with a frequency of approximately 10 s⁻¹. The data was averaged and post-processed to find the 2-dimensional deformation tensor $\mathbf{F}$ in the material from which the shear strain $\gamma_{xy}$ is calculated.

<table>
<thead>
<tr>
<th>C+N</th>
<th>Cr</th>
<th>Ni</th>
<th>Mo</th>
<th>Cu</th>
<th>Ti</th>
<th>Al</th>
<th>Si</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0.05</td>
<td>12.0</td>
<td>9.1</td>
<td>4.0</td>
<td>2.0</td>
<td>0.9</td>
<td>0.4</td>
<td>&lt;0.5</td>
</tr>
</tbody>
</table>

During the cyclic shear tests the magnetic induction was measured to monitor the course of the martensitic transformation. Post et al. (2008) give calibration data for this specific sensor. For this paper however the raw sensor readings will be of more interest than the actual martensite volume fractions. The magnetic induction value is subject to the Villari effect, it depends on the applied stress. This has been shown for tensile stresses by for example Post et al. (2008) and Maréchal et al. (2012). It also appears when a shear stress is applied. Moreover, the effect of the shear stress is symmetric with respect to zero stress. This offers the possibility to determine the strain and stress at which, during the strain reversal, a zero shear stress in the martensite is reached. In this way the partitioning of the stress between both phases can be estimated.

3. EXPERIMENTAL RESULTS

The measured shear stress vs. shear strain data are shown in figure 1 and the absolute values of the stresses vs. cumulative strains are plotted in figure 2. It is clearly seen, that after strain reversal reyielding starts at a distinctly lower stress than was attained before strain reversal. This indicates that the material behavior of the austenite has a strong Bauschinger effect, which agrees with the findings of Spencer et al. (2009). The tests with considerable transformation before strain reversal show that soon after reyielding a stress level is reached, which exceeds the stress level before reversal.
Fig. 1. Shear stress versus shear strain during cyclic shear tests.

Fig. 2. Absolute stress versus cumulative strain during cyclic shear tests.

Fig. 3. Magnetic induction versus cumulative shear strain during cyclic shear tests.

In figure 3 the magnetic induction is plotted as a function of total accumulated strain. After strain reversal considerably more strain needs to be applied for the transformation to restart. Hamasaki et al. (2014) reported a similar stagnation of martensite transformation after strain reversal. In test R04 no martensite was formed before strain reversal. Yet, more plastic strain is needed to obtain a similar amount of martensite as in a monotonic test (M).

In figure 4 the magnetic induction is plotted against the cumulative absolute shear stress. The curve for test R04 now closely follows the monotonic test result. This indicates that the stress rather than the strain is driving the transformation. When more martensite has formed before strain reversal, again considerably more stress is required for martensite formation than before the strain path change. The reason for this is that the hard martensite already present in the material will carry a higher portion of the applied stress than the soft austenite. This effect is enhanced by the large Bauschinger effect which is present in the austenite. More stress must be applied to the phase mixture to raise the stress in the austenite to a level where transformation is induced again. This will be confirmed by the model calculations in section 5.

3.1. Stress partitioning

In figure 5 a detailed view of the induction voltage during stress reversal of test R20 is shown. It can be clearly seen, that the induction signal gradually rises when the stress drops from its maximum value of approximately 330 MPa. Maréchal et al. (2012) used this effect to calibrate the induction signal for the applied stress in tensile tests. With that result they estimated the value of the stress in the martensite fraction. They did this by removing the applied stress and measuring the value of the signal also at zero load.
Fig. 5. Magnetic induction versus stress during cyclic shear tests. Detail of test R20 in figure 4.

Table 2. Applied external stress and estimated martensite stress.

<table>
<thead>
<tr>
<th>test</th>
<th>applied strain</th>
<th>applied stress</th>
<th>martensite stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>R09</td>
<td>9.4%</td>
<td>265 MPa</td>
<td>280 MPa</td>
</tr>
<tr>
<td>R16</td>
<td>16.3%</td>
<td>300 MPa</td>
<td>360 MPa</td>
</tr>
<tr>
<td>R20</td>
<td>20.5%</td>
<td>330 MPa</td>
<td>410 MPa</td>
</tr>
<tr>
<td>xtra</td>
<td>24.0%</td>
<td>355 MPa</td>
<td>425 MPa</td>
</tr>
<tr>
<td>R27</td>
<td>27.0%</td>
<td>410 MPa</td>
<td>500 MPa</td>
</tr>
</tbody>
</table>

However, from our measurements it is apparent, that the induction signal keeps rising even after the applied stress has reached zero level and becomes negative. This indicates, that the stress in the martensite is actually higher than the overall applied stress. The actual level of the martensite stress can be estimated by determining the (negative) applied stress at which the induction signal reaches its maximum. Assuming both phases behave elastically during this phase of the stress reversal, the martensite stress can be estimated as the difference between the overall applied stress at reversal and the stress at maximum signal. For test R20 this is estimated as \( \tau_a \approx 330 + 80 = 410 \) MPa.

Note that this measurement is only possible in a shear test as the Villari effect is symmetric with respect to positive and negative values of the shear stress. No such symmetry exists with respect to tensile and compressive stresses. The martensite stresses estimated in this way are summarized in table 2. The ‘xtra’ entry is from a separate test shown in figure 7.

4. CONSTITUTIVE MODEL

The martensitic transformation is modeled as a stress-driven process similar to the model of Tamura (1982). It depends on the stress resolved in the austenite phase and it is determined as a function of the additional mechanical driving force supplied to the material as formulated by Patel and Cohen (1953). The model uses the Mean-Field homogenization method, which is based on the evolution of the average values of the field variables, stress and strain, in the constituting phases and the interactions between these average values. In this way it is possible to distinguish the stress in the phases from the overall applied stress. A detailed description of the complete model can be found in Perdahcioglu and Geijselaers (2012). A resume of it will be given in this section.

4.1. Martensite transformation model

The martensitic transformation involves a diffusion-less change of crystal structure. This was analyzed by Wechsler et al. (1953) and Bowles and MacKenzie (1954) starting from the postulate of an invariant plane (habit plane) as interface between the martensite and the parent austenite, where \( \mathbf{n} \) is the normal to the habit plane. The deformation applied to the normal is described by the vector \( \mathbf{m} \). Due to lattice symmetry 24 different transformation systems \((\mathbf{n}, \mathbf{m})\) can be identified.

When a stress \( \sigma \) acts, while the transformation evolves, it supplies additional mechanical driving force \( U \) for the transformation (Patel & Cohen, 1953)

\[
U = \sigma_j \cdot (\mathbf{m} \otimes \mathbf{n}) = \sigma_j \cdot \frac{1}{2} (\mathbf{m} \otimes \mathbf{n} + \mathbf{n} \otimes \mathbf{m}) \quad (1)
\]

Here, \( \sigma_j \) is the Cauchy stress in the austenite phase. In a polycrystalline material there are always some grains optimally oriented with respect to the local stress to maximize the mechanical driving force. Then, the maximum value of \( U \) is found as:

\[
U_{\text{max}} = \sum_j \sigma_j \lambda_j \quad (2)
\]

where \( \lambda_j \) are the eigenvalues of the symmetric transformation deformation tensor \((\mathbf{m} \otimes \mathbf{n} + \mathbf{n} \otimes \mathbf{m})/2\) and \( \sigma_{ij} \) are the principal values of the local austenite stress tensor, both sorted in descending order (Geijselaers & Perdahcioglu, 2009).

The values of \( \lambda \) are material parameters, which are based on measured data such as transformation dilatation or crystal lattice constants. By XRD measurement the lattice parameters of both the aus-
tensite and the resulting bcc phases can be determined. With these data 24 transformation variants (n;m-pairs) can be calculated with respect to the austenite lattice along the procedures outlined by Wechsler et al. (1953) or Bowles and MacKenzie (1954). The resulting eigenvalues are given in table 3.

When the maximum supplied driving force \( U_{\text{max}} \) exceeds the required critical driving force \( \Delta G^* \) then according to Tamura (1982) the transformation will start. The required critical driving force \( \Delta G^* \) is the lumped value of a collection of separate energy terms such as chemical driving force, elastic energy, plastic dissipation, surface energy and interface mismatch energy.

### Table 3. Parameters for martensite transformation kinetics.

<table>
<thead>
<tr>
<th>( \lambda_1 )</th>
<th>( \lambda_2 )</th>
<th>( \lambda_3 )</th>
<th>( \Delta G^* )</th>
<th>( n )</th>
<th>( q )</th>
<th>( r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.124</td>
<td>0.0</td>
<td>-0.104</td>
<td>56.5 MPa</td>
<td>2</td>
<td>0.5</td>
<td>2</td>
</tr>
</tbody>
</table>

The amount of martensite formed \( f_a^* \) can be expressed as a monotonic function of \( U_{\text{max}} \) (Perdahcioglu and Geijselaers, 2012):

\[
f_a^* = F(U_{\text{max}}) = 1 - \left( \frac{U_{\text{max}} - \Delta G^*}{q\Delta G^*} \right)^n
\]

where \( n, r, q \) and \( q \) are parameters that determine the shape of the transformation curve. The values are obtained from fitting to the experimental results and are also given in table 3.

#### 4.2. Mean-field modeling

The mean-field homogenization method is based on the evolution of the average values of the field variables in sub domains and the interactions between these average values. The overall stress \( \sigma \) and strain \( \varepsilon \) are related to the average stresses \( \sigma_y \) and strains \( \varepsilon_y \) in the austenite and \( \sigma_{a'} \) and \( \varepsilon_{a'} \) in the martensite by:

\[
\begin{align*}
\sigma &= (1 - f_{a'})\sigma_y + f_{a'}\sigma_{a'} \\
\varepsilon &= (1 - f_{a'})\varepsilon_y + f_{a'}\varepsilon_{a'}
\end{align*}
\]

It is assumed that the macroscopic stress-strain relation that is determined for an individual phase is also valid as an average stress average strain relation for that phase within the compound:

\[
\sigma_y = C_{y} : D_y, \quad \sigma_{a'} = C_{a'} : D_{a'}
\]

where \( D_{a'}, \sigma_{a'} \) is the average strain rate in the respective phase and \( C_{a'} \) is the consistent fourth order elastoplastic tangent of the phase. The constitutive model used here is Chaboche (1986) kinematic hardening. The data used in the model are summarized in appendix A.

To close the set of equations the relation between average phase strain rates \( D \) and the overall strain rate \( \dot{D} \) has to be specified through fourth order strain concentration tensors \( A_{y,a} \) (Hill, 1965):

\[
\begin{align*}
\dot{D}_y &= A_{y}: D; \quad \dot{D}_{a'} = A_{a'}: D \\
\end{align*}
\]

which, by virtue of equation (4), are subject to:

\[
(1-f_{a'}) A_{y} + f_{a'} A_{a'} = I
\]

where \( I \) is the symmetric fourth order unit tensor.

Different schemes have been formulated using specific definitions of \( A \). Here we use an approximation to the self consistent scheme by Lielens et al. (1998), the so called double inclusion scheme, see also Doghri and Ouaar (2003) or Perdahcioglu and Geijselaers (2010). It is derived by interpolating between two variants of the Mori and Tanaka (1973) scheme with the roles of both phases as matrix and inclusion interchanged. For the Mori-Tanaka model with martensite as inclusion in austenite and the other way around we can write:

\[
\begin{align*}
D_{a'} &= H_{a'} : D_y \\
D_y &= H_y : D_{a'} \rightarrow D_{a'} = H^{-1}_{a'} : D_y
\end{align*}
\]

where \( H_i \) is the 'local' strain concentration tensor for the strain in the inclusion \( i \) with respect to that in the matrix \( m \). \( H_i \) is calculated as:

\[
H_i = (1 - S_m : (1 - C_{m}^{-1} : C_i))^{-1}
\]

where \( S_m \) is the Eshelby tensor of the matrix:

\[
S_m = \frac{3\kappa_m}{3\kappa_m + 4\mu_m} I + \frac{6\kappa_m + 2\mu_m}{5 \cdot 3\kappa_m + 4\mu_m} I^d
\]

where \( I^1 \) and \( I^d \) are the fourth order volumetric and deviatoric unit tensors. The bulk modulus \( \kappa \) and the shear modulus \( \mu \) are found from an isotropic projection of the elasto-plastic constitutive tensor:

\[
3\kappa_m = C_m :: I^1; \quad 10\mu_m = C_m :: I^d
\]

From the interpolated local concentration tensor \( H_{a'}^{-1} \):

\[
H_{a'}^{-1} = (f_{a'} H_{a'}^{-1} + (1 - f_{a'}) H_y)^{-1}
\]
with the help of equation (7), the strain concentration tensors with respect to the global strain are calculated

\[ A_\gamma = ((1 - f_{a'})I + f_{a'}H_{a'})^{-1} \]
\[ A_{a'} = H_{a'} : A_\gamma \]  

(13)

4.3. Mean-field model with phase transformation

To obtain a homogenized stress-strain relation the deformation rate is partitioned into an elastic rate \( D^e \), a plastic deformation rate \( D^p \) and a transformation plasticity rate \( D^t \) (Kubler et al., 2011), of which the elastic plus plastic rate is partitioned among the phases:

\[ D_{\gamma,a'} = D^e_{\gamma,a'} + D^p_{\gamma,a'} = A_{\gamma,a'} : (D - D^p) \]  

(14)

Differentiation of the stress as defined in equation (4) yields:

\[ \dot{\sigma} = (1 - f_{a'})\dot{\sigma}_\gamma + f_{a'}\dot{\sigma}_{a'} + \dot{f}_{a'}(\sigma_{a'} - \sigma_\gamma) \]  

(15)

The consequence of the last term in the right hand side of this equation would be that newly formed martensite gets the same stress as the already formed martensite. A more realistic assumption is to assign to pristine martensite the stress of the parent austenite and add this as a dilution term to the martensite stress rate:

\[ \dot{\sigma}_{a'} = C_{a'} : D_{a'} + \frac{\dot{f}_{a'}}{f_{a'}}(\sigma_\gamma - \sigma_{a'}) \]  

(16)

Substitution of equation (14) into (16) and of the result into equation (15) yields:

\[ \dot{\sigma} = ((1 - f_{a'})C_{\gamma} : A_{\gamma} + f_{a'}C_{a'} : A_{a'}) : (D - D^p) \]  

(17)

The transformation plasticity depends on the transformation rate:

\[ D^t = T \dot{f}_{a'} \]  

(18)

where \( T \) is the second order transformation plasticity tensor, which can be expressed as a function of \( U_{\text{max}} \) (Perdahcoğlu & Geijserelaers, 2012). An implicit equation for the transformation rate is found by differentiation of equation (3) and substitution of equations (14) and (18):

\[ \dot{f}_{a'} = F^r \frac{\partial U_{\text{max}}^{\max}}{\partial \sigma_\gamma} : C_{\gamma} : A_{\gamma} : (D - T \dot{f}_{a'}) \]  

(19)

Solving for \( \dot{f}_{a'} \) yields an explicit expression:

\[ \dot{f}_{a'} = \frac{F^r \frac{\partial U_{\text{max}}^{\max}}{\partial \sigma_\gamma} : C_{\gamma} : A_{\gamma} : D}{1 + F^r \frac{\partial U_{\text{max}}^{\max}}{\partial \sigma_\gamma} : C_{\gamma} : A_{\gamma} : T} \]  

(20)

After combining (20) with (18) and (14) and substitution into (17), the stress-strain response for the homogenized material including transformation plasticity is obtained as:

\[ \sigma = \sum_{i=\gamma,a'} f_i C_i : A_i : \left( \frac{F^r \frac{\partial U_{\text{max}}^{\max}}{\partial \sigma_\gamma} : C_{\gamma} : A_{\gamma} : D}{1 + F^r \frac{\partial U_{\text{max}}^{\max}}{\partial \sigma_\gamma} : C_{\gamma} : A_{\gamma} : T} \right) \]  

(21)

This material tangent describes the constitutive behavior of the austenitic steel including the phase transformation and transformation plasticity. Also the tension-compression asymmetry of the mechanical response is included. The driving force for transformation differs between tension and compression due to the difference of the positive and negative eigenvalues of the transformation strain.

5. SIMULATION RESULTS

The parameters for the cyclic stress strain behaviour of both austenite and martensite as used in the simulations are summarized in table 4. The austenite is modeled with pronounced Bauschinger effect, whereas this is kept to a modest level in martensite. The parameters describing the transformation are given in table 3. All parameters have been optimized to fit the simulation results to the measurements.

5.1. Cyclic stress-strain and transformation response

In figure 6 the simulated stress-strain response is plotted, together with the measured behavior. The correspondence is very good over the whole range of strains in the monotonic behavior as well as in the reversed shear response. This also holds for the ‘unfinished’ cyclic test as shown in figure 7.

In figure 8 the phase fraction as a function of cumulative strain is shown. Comparison with figure 3 shows that the main characteristics are largely reproduced. In test R04 very little martensite was pro-
duced before strain reversal, yet after reversal extra strain needs to be applied to obtain a comparable amount of martensite. In all tests where martensite was formed before strain reversal, a stagnation of the transformation is observed after reversal. After strain reversal additional strain must be applied to restart the transformation. This is caused by the Bauschinger effect of the austenite stress-strain response. More strain is required to raise the stress in the austenite to the level before strain reversal. The stress in the austenite determines the transformation. In figure 9 the phase fraction as a function of applied stress is shown. Also here comparison with figure 4 shows that the applied stress needed for transformation in test R04 closely follows the stress-transformation curve of the monotonic test. In all tests where martensite was formed before strain reversal, after strain reversal more stress than before reversal needs to be applied to restart the transformation. The reason for this is that the applied stress is partitioned among both phases. The hard martensite will tend to carry more stress than the soft austenite. The stress in the austenite is considerably lower than the externally applied stress.

![Fig. 6. Stress versus strain during cyclic shear tests, comparison of measurements and simulations (dashed).](image)

![Fig. 7. Stress versus strain during 'xtra' cyclic test, comparison of measurements and simulations (dashed).](image)

![Fig. 8. Simulated martensite phase fraction versus absolute cumulative strain during cyclic shear.](image)

![Fig. 9. Simulated martensite phase fraction versus absolute stress during cyclic shear.](image)

5.2. Stress partitioning

On account of equation (16), in the calculations the first formed martensite has an average stress equal to that of austenite. The higher stiffness of the martensite compared to that of the austenite causes stress concentration in the martensite, the average stress in the martensite quickly rises. This is shown in figure 10. The calculated stress in the martensite compares well with the values of the stress in the martensitic phase determined from the analysis of the magnetic induction signal as explained in section 3 and summarized in table 2.
6. CONCLUSIONS

A 12Cr9Ni4Mo austenitic stainless steel has been subjected to cyclic shear tests. This steel transforms to martensite when subjected to mechanical working. During the tests, apart from the stress and strain also the magnetic induction has been measured to monitor the course of the transformation. From the magnetic induction signal after strain reversal the partitioning of the shear stress among both phases can be estimated.

To represent the material behavior of this steel a model can be used in which i) the martensite transformation is modeled as a stress-induced transformation, where the stress in the austenite phase is assumed to drive the transformation, ii) the stress-strain response of the austenite is characterized by a strong Bauschinger effect and iii) both phases, austenite and martensite are modeled individually and are combined through a mean-field homogenization method.

This way an excellent fit of the stress-strain response is obtained. This applies to the stress during a monotonic deformation as well as to the stress level after strain reversal. The model also reproduces accurately the stress partitioning among the phases as estimated from the magnetic induction signal. The simulated transformation as function of both stress and strain displays the main characteristics that are found in the measurements.

ACKNOWLEDGEMENT

This research was carried out under project number M63.1.09373 in the framework of the Research Program of the Materials innovation institute M2i (www.m2i.nl).

APPENDIX A. MATERIAL PARAMETERS FOR KINEMATIC HARDENING

Both austenite and martensite are modeled as kinematic hardening with Chaboche back stress evolution and a Voce law for the yield surface radius.

The elasto-plastic behavior of phase i can be described by the yield condition:

\[
\frac{3}{2} (S^i - X^i) : (S^i - X^i) - R^i(p_i) = 0
\]  

(22)

where \( S^i \) is the deviatoric stress and \( X^i \) the back stress in the phase, \( R^i \) is the radius of the yield surface, which depends on the equivalent plastic strain \( p_i \). Armstrong and Frederick (1966) formulated an evolution equation for the back stress, which was extended by Chaboche (1986) to:

\[
\dot{X}^i = \sum_j X^j = \sum_j (h^j D^j - c^j X^j p_i) 
\]  

(23)

A Voce-type law is used for the yield surface radius \( R^i(p_i) \):

\[
R^i(p_i) = R^i_0 + \sum_j \Delta R^j (1 - e^{-\rho^j p_i})
\]  

(24)

The parameters used in the simulations are summarized in table 4.

<table>
<thead>
<tr>
<th>Material properties for austenite and martensite.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R^i_0 ) (MPa)</td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td>331</td>
</tr>
<tr>
<td>( \Delta R^j ) (MPa)</td>
</tr>
<tr>
<td>-75</td>
</tr>
<tr>
<td>-180</td>
</tr>
<tr>
<td>( r^i_{ij} )</td>
</tr>
<tr>
<td>1.3</td>
</tr>
<tr>
<td>( h^i_{ij} ) (GPa)</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>0.27</td>
</tr>
<tr>
<td>( c^i_{ij} )</td>
</tr>
<tr>
<td>1.3</td>
</tr>
</tbody>
</table>

REFERENCES

Stress Analysis, Mat Sci Eng A-Struct, 438-440, 262–266.


ZACHOWANIE SIĘ BLACH Z AUSTENITYCZNEJ STALI NIERDZEWNEJ PODDAWANYCH CYKLICZNYM OBCIĄŻENIOM ŚCINAJĄCYM

Streszczenie

Próbki z austenitycznej stali nierdzewnej poddawano odkształceniu o dużej amplitudzie i przeciwnych zwrotach. W przeprowadzonych doświadczeniach, poza pomiarem naprężenia i odkształcenia, mierzono również indukcję magnetyczną do monitorowania przemian austenit w martensyt. Dzięki bezpośrednim pomiarom indukcji magnetycznej możliwe było oszacowanie wielkości naprężenia w poszczególnych fazach materiału.

Przy małych odkształceniach, dla przeciwnego zwołu, obserwowany jest klasyczny efekt Bauschingera. Gdy dokonywana jest zmiana zwołu przy większych odkształceniach, mierzone naprężenie jest większe niż przed zmianą. Obserwowana jest...
również stagnacja przemiany fazowej, to znaczy, że potrzebne jest wyższe odkształcanie i naprężenie niż przed zmianą ścieżki odkształcania aby ponownie zaczęła się przemiana po zmianie. Zaoberwowane zachowanie można opisać modelem, w którym dla przemiany martensytycznej ma zastosowanie model naprężenia indukowanego przemianą.

Zachowanie konstytutywne austenitu i martensytu zostało opisane modelem Chaboche’a uwzględniającym efekt Bauschinger’a. W modelu wykorzystano metodę homogenizacji dla opisu materiału i faz, otrzymując konstytutywny model zachowania się dwu-fazowego kompozytu. Przyłożone naprężenie, naprężenie w fazie martensytu i obserwowane zachowanie się podczas przemiany w trakcie cyklicznego ścignania zostało bardzo dobrze odtworzone w symulacjach modelu.

Received: November 10, 2014
Received in a revised form: November 17, 2014
Accepted: November 17, 2014
CONCEPT OF THERMAL MODELLING FOR HOT STRIP ROLLING OF MAGNESIUM

ALEXANDER NAM1*, UWE PRÜFERT2, MICHAEL EIERMANN2, RUDOLF KAWALLA1

1 Institute of Metal Forming, Technische Universität Bergakademie Freiberg
Bernhard von Cotta Straße 4, Freiberg, Germany
2 Institute of Numerical Analysis and Optimization, Technische Universität Bergakademie Freiberg
Akademiestraße 9, Freiberg, Germany
*Corresponding author: alexander.nam@imf.tu-freiberg.de

Abstract

The paper introduces a concept for the reversing hot rolling of magnesium strip consisting of reheating of twin roll casted rough strip, carrying of coil from a furnace to coiler, setting down on mandrel, uncoiling, hot rolling and coiling of the magnesium strip. In the first stage the thermal models should be established. These models shall serve in the subsequent stages in order to support by developing of a roll gap model as well as model for forecasting the microstructure and material properties of the magnesium strip. Moreover, the paper represents the state of the concept development of the thermal models. Up to now, the following technological steps are implemented: reheating of the coil, transport of coil from furnace to the coiler, setting down on the mandrel and uncoiling of coil. The modelling of these models is worked out in three dimensions and are calculated numerically by using the Finite Element Method.

Key words: hot strip magnesium rolling, heating, numerical modelling

1. INTRODUCTION

The application of magnesium alloys becomes more important in various industrial fields over the past years. This is due to the fact, that magnesium alloys have high functional properties by contrast with the most frequent applicable materials as steel and aluminum. Magnesium stands out by means of its low density (1740 kg/m³), which is 35% lighter than aluminum and 78% lighter than steel. The production of semi-finished products from magnesium alloys is usually carried out at higher temperatures. It causes through its hexagonal lattice structure (Kawalla et al., 2006). In this paper the technological chain of magnesium strip production at the institute of metal forming, TU-Freiberg is considered. The details of the technological chain will be considered subsequently in this paper. Moreover, this paper introduces a modelling concept for reversing hot strip rolling of magnesium alloy AZ31. We present the current state of this project. In the last section we show the results of simulation.

2. TECHNOLOGICAL CHAIN FOR HOT STRIP ROLLING OF MAGNESIUM ALLOY AZ31

The Institute of Metal Forming have been carrying out the research on the production of industrially applied strip/sheet from magnesium alloys since 2002. Since then, it has been achieved substantial progress in the developed technology. Therefore, the current technological route can be subdivided into 3 steps:
1. Twin-Roll (TR) casting (manufacturing of rough TR casted strip or TR casted coil)
2. reheating of TR casted coil (getting of rolling temperature)
3. reversing hot strip rolling up to appropriate thickness

Figure 1 shows the whole technological chain at the Institute of Metal Forming.

3. THERMAL MODELLING CONCEPT

The thermal modelling concept for hot strip rolling of magnesium alloys consists of technological steps starting from reheating of TR casted coil up to manufacturing of semi-finished products through hot reversing rolling. Therefore, the reversing hot rolling step of magnesium alloy AZ31 can be subdivided in following sub-steps:
1. reheating of Twin-Roll casted coil in furnace,
2. transport of reheated TR casted coil from furnace to coiler of reversing four-high mill,
3. putting of coil on the mandrel of the coiler,
4. unrolling of coil,
5. reversing hot strip rolling (roll gap),
6. roll-up of coil.

Every sub-step is modelled individually separated from each other. However, the results of single sub-steps are connected with each other. Moreover, a coil-model was modelled, which is used within modelling. The description of coil-model and single steps will be further given. The current state of work are the sub-steps including a coil unrolling, which are being describing subsequently in this paper. This issue, i.e. the modelling of coil unrolling, which can be found in the bibliography, has been treated in a number of different ways by various authors (Troyani, 1996; Troyani & Montano, 1999). In all these works the solution was achieved using geometrically adaptive finite element in 2D, 8-node isoparametric elements to adapt the curve shape changing of elements. In our case we have used a different approach to deal with this aspect of the problem. We perform the modelling of coil unrolling on the one hand in 3D and in other hand we execute the geometrical modelling of coil unrolling and the calculation of heat transfer sequentially. The details of implementation will be given below.

3.1. Governing equations

The transient heat transfer calculations are based on the convective-diffusion equation (Ozisik, 1993)

\[ c_p \rho \frac{\partial T}{\partial t} = \nabla \cdot (k \nabla T) \]  

(1)

where \( T(K) \) is temperature, the matrix \( k = \{ k_{ij} \} \), \( i,j = 1, \ldots, 3 \) \((\text{W/(m K)})\) contains the thermal conductivity according to the spatial directions, \( \rho \) (kg/m\(^3\)) is the density, and \( c_p \) is the specific heat capacity. The solution of equation (1) has to satisfy the boundary conditions given below. In the present model, Neumann’s type of boundary conditions are used for coil heating/cooling. Moreover, the surface of the hollow cylinder is divided into single segments \( \Gamma_n \), whereby the boundary conditions distinguish from boundary segment to boundary segment. We have

1. the inner mantle surface \( \Gamma_1 \),
2. the outer mantle and the top of the hollow cylinder \( \Gamma_2 \), and
3. the bottom of the hollow cylinder (symmetry) \( \Gamma_3 \).

The boundary conditions related to these segments are

\[ \vec{n} \cdot (k \nabla T) = -h_1(T - T_{\text{f}}) \quad \text{on } \Gamma_1, \]  

(2)

\[ \vec{n} \cdot (k \nabla T) = -h_2(T - T_{\text{f}}) \quad \text{on } \Gamma_2, \]  

(3)

\[ \vec{n} \cdot (k \nabla T) = 0 \quad \text{on } \Gamma_3, \]  

(4)

where \( T_{\text{f}} \) is the temperature of the air into furnace, and hence of the coils boundary; \( h_1, k = \{ 1, 2 \} \) is the heat transfer coefficient and \( \vec{n} \) is the outer normal to the boundary surface. A radiative heat transfer boundary condition is not considered in the present development. The initial temperature is given by \( T(t_0) = T_{\text{start}} \) in the beginning of calculation.

The idea of the finite element method (FEM) is to bring equation (1) into the variational or weak form by multiplying equation (1) with a test function \( v \in V \), where \( V \) is the space of test function, e.g. \( v \in C^1(\Omega) \). The weak form of a differential equation
is a weight-integral statement that is equivalent to both the governing differential equation as well as certain types of boundary conditions. We obtain the variational equation

$$\int_{\Omega} \rho c_p \frac{\partial T(t,x)}{\partial t} v(x) dx =$$

$$\int_{\Omega} \nabla \cdot \kappa(T(t,x)) \nabla T(t,x) v(x) dx$$  \hspace{1cm} (5)

for all \( v \in V \).

After applying Green’s formula to equation (5) and inserting the boundary conditions we obtain

$$\int_{\Omega} \rho c_p \frac{\partial T(t,x)}{\partial t} v(x) dx =$$

$$- \int_{\Gamma} h_1(T(t,x) - T_{\Gamma_1}(t)) v(x) ds$$

$$- \int_{\Gamma} h_2(T(t,x) - T_{\Gamma_2}(t)) v ds$$

$$- \int_{\Omega} k(T(t,x) \nabla T(t,x) \nabla v(x) dx,$$  \hspace{1cm} (6)

where the integral over the symmetry boundary \( \Gamma_3 \) disappears. After discretization of the domain \( \Omega \), the boundary \( \Gamma \), and of all functions by finite elements, we finally obtain the discrete problem

$$D \frac{\partial T}{\partial t} + (K(T) + H)T = G,$$  \hspace{1cm} (7)

where \( K \) is stiffness matrix, \( D \) is the mass matrix, and the matrix \( H = H_1 + H_2 \) corresponds to the functions \( h_k, k = \{1, 2\} \) and \( G = G_1 + G_2 \) is the vector corresponding to the given data \( T_{\Gamma_k}, k = \{1, 2\} \), in the boundary integrals. By the dependence of the stiffness matrix from the temperature this is a non-linear problem. For the details of the used FEM approach we refer to Section 2 in (Prüfert, 2014).

### 3.2. Modeling of the coil

The domain \( \Omega \subset \mathbb{R}^3 \) is modelled as a hollow cylinder with inner radius \( r_{\text{inner}} \), outer radius \( r_{\text{out}} \) and width \( d \). We assume the magnesium coil as a continuum and model the layer structure by using a non-isotropic formulation in the equation. In contrast to most former works, we consider the heat transfer problem in Cartesian coordinates instead of the more usual cylindrical coordinates. The reason for solving this problem in Cartesian coordinates is the further simulation of the whole magnesium hot strip rolling process. The rolling process of strip will be considered by modelling in connection with un-coiling/coiling process in Cartesian coordinates. Thereby, the transformation from cylindrical to Cartesian coordinates should not be executed. The developed coil model have been modelled with the dimensions 1.0 m (outer diameter) x 0.3 m (inner diameter) x 0.6 m (width of strip) and 0.005 m (thickness of strip). In figure 2 is shown the generated mesh for the modelled coil geometry. The mesh was so generated that the mesh on the outer and inner side of the coil is fine while at intermediate location is spaced out evenly. The shown coil geometry in figure 2 was meshed through 59300 prism elements with 36210 nodes.

![Fig. 2. The modelled and meshed geometry of coil model.](image)

### 3.3. Reheating of the coil in furnace

The reheating of coil occurs in the special for magnesium engineered air circulation furnace. The principal of heating process in the furnace is based on the circulation of the heated air within a chamber by means of furnace’s fans. Thereby, the temperature distribution in the furnace chamber is set during heating process homogeneous. However, a separated furnace model has not been developed for calculation of furnace temperature. It is assumed therefore the uniform changing of the furnace temperature in the whole chamber during heating.

In order to calculate temperature state of furnace it was measured the temperature within the furnace chamber by means of thermocouple during coil reheating. Therefore, based on the obtained experimental data were correlated the essential thermal coefficient such as heat transfer and thermal conductivity inside in a coil due to inverse method (Beck et al., 1996). Hence the calculation of temperature state of the furnace is carried out based on an intended heat transfer coefficient

$$h = 1.2 \cdot e^{3.285 + 1.57 \cdot 10^{-3} T}$$  \hspace{1cm} (8)

It is also assumed that the variation of coil temperature during coil heating/cooling occurs by
means of a calculated thermal conductivity inside into coil without consideration of exogenous effects/factors:

\[
  k_{rad} = 19.72 \left( 1 - e^{-\left( \frac{T}{131.75} \right)^{0.99}} \right)
\]

These values have been already introduced in (Nam et al., 2014).

3.4. Cooling of the coil

The coil is being transported from the furnace to the coiler as soon as the coil reheating is completed. During the coil transport occurs small losses of temperature due to free convection. The room temperature is assumed as 25°C. The coil is subsequently is being put on the mandrel. It leads to the next heat losses due to contact with the mandrel. However, the mandrel consists of a outer insulation layer and steel. Thereby the heat losses are minimal. Furthermore, the thermal conductivity of mandrel is assumed as 10 times smaller than thermal conductivity of steel.

3.5. Unrolling of the coil

As soon as the coil has been put on the mandrel, the coil unrolling is occurring. This unroll process is modeled in the way that geometrical modelling of coil unrolling and the calculation of heat transfer take place sequentially, but will be performed for every time step. By decreasing the time increment, this process will converge to the continuous simultaneous cooling when unrolling process. Since we consider the coil as a continuum, we define the unrolled coil as union of a hollow cylinder and a thin sheet, where the diameter of the cylinder decreases if the length of the sheet increases. Here, the main difficulty is the transport of the temperature distribution to the next grid. We do this by layer-wise rotation, shrinking of the grid in radial direction and linear interpolation of the data, cf. figure 3.

4. IMPLEMENTATION

The implementation of the model has been carried out in MATLAB® by using the object-oriented FEM Toolbox OOPDE (Prüfert, 2014). The geometry of the coil is a generalized cylinder. By this fact, we can model it by three dimensional prism elements, where we also can use the symmetry of the coil with respect to the x–y-plane. The software is modularized, i.e. for every step in the technological chain there is a separated class. These classes are derived from an abstract super-class that implements all common properties and methods, cf. figure 4.

The classes coilHeating and coilCoolingOnDorn differ in the boundary conditions and the source term, while coilCoolingOnDorn and coilUnrolling differ in the geometry, i.e. in the grid property.

The unrolling method of coilUnrolling class manages the transformation of the 3D grid as well as the transport of the solution from grid to grid. In every time step of the transient heat transfer calculation one unrolling step will be performed.

Note that the transformation of the grid makes it necessary to re-assemble the linear system. From that reason we can also update the nonlinear temper-
ature diffusion coefficient without any additional effort in every iteration step. This is in contrast to the adaptive update strategy for the coil heating procedure suggested in (Nam et al., 2014), where the adaptive update of the coefficients speeds up the computations. A further problem is caused by the fact, that for every step within the rolling process we have a solution on a changing grid, i.e. we should store the solution and the associated grid. This results in an enormous growing of the data within the memory while running the simulation. However, since a call of the unrolling method is rather cheap we do not store the grids and only the solutions at given checkpoints and the information about the length of the sheet. The associated grid will be recomputed from this information if needed. The main advantage of this approach is that we can add very easily further modules as for example a module for the strip rolling, re-heating and so on. These modules will be derived from coil class and only specialized method for e.g. strip rolling must be added, cf. figure 4.

5. RESULTS

The temperature distribution in the coil was calculated in every technological step. The temperature of previous step was used as an initial temperature for the following step. The heating of coil occurred according to a reheating furnace program, which is shown in figure 5.

Fig. 5. Calculated furnace program for magnesium coil reheating.

The total duration of reheating the coil was approximately 10 hours. The temperature distribution within the coil cross-section is shown in figure 6. Based on the results it could be noted that the temperature difference is lesser than 5°C. Hence, the temperature of coil is therefore distributed homogeneously. At the same time the temperature distribution reveals correct behavior from a physical perspective. The quantitative verification of results of coil reheating has been reported in (Nam et al., 2014).

The following figures show the temperature distribution within the coil cross-section at reheating and transport steps. The transport step usually occurs very fast due to loss temperature and takes about 1-2 minutes. The coil is being cooled down due to effects of free convection, cf. figure 7. For lack of specific experimental data pertaining to the present application, the value of 13 W/m²K for the free convection was used (Meyer, 2005).

Fig. 6. Temperature distribution within the coil crosssection after reheating.

Fig. 7. Temperature distribution within the coil crosssection after transport to the coiler.

The development of the temperature during unrolling of the coil is shown in the following figures 8-11. We have simulated 10 seconds of the unrolling process, where a magnesium strip of approximately 1.6 meters was unrolled. The temperature lost at the edges of the strip is visible. Within this time the unrolled strip reaches the guide pulley of reversing four-high mill. It may be also noted that convective heat loss of strip during coil unrolling before the guide pulley does not exceed on an average 10-20°C.
In the near future the experimental test will be carried out the measurements of temperature distribution in the unrolled strip during unrolling process in order to verify the simulation results.

6. SUMMARY

A modelling concept for reversing hot strip rolling of magnesium alloy AZ31 was introduced in this paper. Moreover, it was presented the state of work. A worked out mathematical 3D model is developed to access the temperature state during technological steps as reheating, transport and unrolling of coil. This model considers the heat transfer effects due to a heat conductivity and convection. Besides it was proposed the new modelling approach to modelling of unrolling process. It was solved a geometrical modelling of unrolling process in connection with solving of heat transfer problem. By the solving of heat transfer problem was considered the unrolled strip as well as changed thereby volume of coil in interaction with each other. Finally, it was shown the numerical results for 5 mm thick, 600 mm width strip of magnesium alloy AZ31 after reheating, transport from the furnace to the coiler and during the uncoiling process.

In the near future the experimental tests will be performed to verify the predictions of simulation results.

REFERENCES


KONCEPCJA MODELOWANIA ZJAWISK CIEPLNYCH ZACHODZĄCYCH PODCZAS WALCOWANIA NA GORĄCO BLACH Z MAGNEZU

Streszczenie

W pracy przedstawiono koncepcję technologii nawrotnego walcowania na gorąco blach z magnezu, składającą się z podgrzewanych dwóch walców do odlewania ciągłego materiału wsadowego, transportu zwoju z pieca do zwijarki, odłożenia materiału na trzpień walcowniczy, rozwijania, walcowania na gorąco i zwijania blach. Na wstępie opracowane zostały modele termiczne. Modele te odzwierciedlają zjawiska zachodzące w kolejnych etapach procesu, jak również uwzględniają model szczelin walcowniczej oraz model rozwoju mikrostruktury i własności materialowych blachy. W pracy zaprezentowano dotychczasowy stan prac związanych z opracowywanymi modełami termicznymi. Jak dotąd, zostały zaimplementowane następujące etapy technologiczne: podgrzewanie zwoju, transport z pieca do zwijarki, ułożenie na trzpień walcowniczym i rozwijanie zwoju. Dla wymienionych zagadnień opracowano rozwiązania w trzech wymiarach i przeprowadzono obliczenia numeryczne z użyciem metody elementów skończonych.

Received: October 29, 2014
Received in a revised form: December 17, 2014
Accepted: December 20, 2014
INFLUENCE OF TOOL GEOMETRY ON SURFACE CONDITION OF V-BENT ALUMINUM SHEET

ŁUKASZ MORAWIŃSKI*, ANDRZEJ KOCANIA

Warsaw University of Technology, 3 Narbutta St., 02-524 Warsaw, Poland
*Corresponding author: l.morawinski@wip.pw.edu.pl

Abstract

Bending is one of the processes, which are most commonly performed in sheet metal forming. It is quite difficult to determine the moment of formation of defects in the form of cracks on the surface of the sheet metal. This is very often the criterion for the rejection of the product. For this reason, it is important to evaluate the state of the surface of the sheet metal along with the line of bending, because there are the largest deformations, leading to the formation of defects such as localized necking or cracks.

Application of 3D digital microscope is presented to evaluate the surface condition of the V-bent specimens. The parameter Wmax obtained from the analysis of the geometry of the sheet metal surface with the usage of the 3D visualization is introduced. In contrast to the surface roughness parameters, it does not average the highest and the lowest points in the given area. Its value depends directly on the deepest localized necking currently occurring in the study area, which may be followed by a crack initiation. Due to this, it makes possible to define better the moment of cracking. Images of surfaces of V-bent specimens made of aluminum sheets EN AW-2017A (PA6) commonly used in the automotive and aerospace industries were subjected to detailed analysis. Using computer modeling of the bending process, the values of strain have been determined at the surface of the sheet in the bending line. On this basis, the analysis of dependence of the parameter Wmax on thickness of sheet metal as well as geometry of the tools has been conducted.

The presented possibility of determining the moment of crack initiation allows to safely perform the bending process by avoiding formation of defects on the surface of the aluminum sheet.

Key words: V-bending, tool geometry, 3D digital microscope, computer simulation (FEM), surface condition, grooves, cracks

1. INTRODUCTION

Bending is one of the most commonly performed processes in metal forming. The variety of materials and tooling geometry in the process of bending makes it difficult to determine the moment of the formation of defects on the surface of the sheet metal. Such defects are very often the criteria for the rejection of the products. For this reason, it is important to evaluate the condition of the sheet metal surface along the bending line, because the greatest plastic strain occurs just there, leading to the formation of defects in the form of grooves or cracks as shown by Davidkov et al. (2012), Mattei et al. (2013), Sarkar et al. (2001), Sarkar et al. (2004). For this purpose, observations of the surface condition of the bent aluminum sheet have been carried out by means of the microscope. There have been many attempts made to use the computer simulation for studying the effects of process parameters (e.g. material thickness, punch radius) on spring-back or spring-go (Bakhshi-Jooybari et al., 2009; Thipprakmas & Phanitwong, 2011; Thipprakmas & Rojananan, 2008). In this paper, the results of microscopic examinations of surface condition of V-bent specimens were related to plastic strains obtained from the computer simulations of the V-bending process. It has been helpful to understand the formation of deep grooves and cracks along the bending line. On the basis of the analysis, some recom-
Recommendations have been presented for understanding any state of the aluminum sheet surface subjected to a V-bending process.

2. PREPARATION OF SPECIMENS

Aluminum sheet EN AW-2017A (PA6) has been chosen for this research work. It is a common material applied in the aerospace and automotive industries. Specimens for bending were cut out from the 2 mm thick sheet with a length of 60 mm and a width of 40 mm. V-bending process was carried out on a press brake AMADA HFE M2. The press brake was equipped with a punch with radius $r = 0.5$ mm and a die opening width $w = 12$ mm, see figure 1a. All specimens were bent along the rolling direction on the selected angles ranging from 0° up to the angle where the crack was visible with an unaided eye.

3. EXAMINATION OF THE SURFACE CONDITION

Increasingly used macroscopic vision systems can detect the existing cracks and enable possible rejection of the defective product. However, these systems cannot evaluate the current condition of the surface of the sheet metal and predict the moment of cracking. For this reason a detailed knowledge of the actual 3D geometry of the surface is needed, which can be obtained by using 3D digital microscope. The experimental set-up with a microscope gives a possibility to create a 3D visualization of the surface (figure 2a) from the analysis of the image sharpness and a traditional 2D image (Morawiński et al., 2013). 3D visualization allows the evaluation of the surface condition of the sheet metal in a qualitative way. To clearly evaluate its condition in a quantitative way, the parameter $W_{\text{max}}$ was introduced. It is derived from the analysis of the geometry outline of the sheet metal surface taken from the 3D visualization. These outlines have been approximated by using a second-degree polynomial curve. Then, the distances $h$ between geometry outlines and polynomial curve were calculated for all of measurement points. Distribution of $h$ values is shown in figure 1b. The biggest differences (heights) between surface elevations and the deepest grooves are defined as $W_{\text{max}}$ values. They are different for various surfaces to be analysed. It was assumed, that the parameter $W_{\text{max}}$ is calculated between the distribution points equal to 0.1%. These values take into account the possibility of the measurement errors caused, for example, by a dirty surface.

In contrast to roughness parameters, it does not average the heights of the highest and the lowest points in the given area. Its value depends directly on the deepest grooves currently occurring in the image study area, which may be followed by a crack initiation. Thus, it is possible to define the moment of cracking more precisely. Figure 2b shows the results of measurements of the bent specimens. Each of the dots on the graph represents a result of the analysis of 3D visualization (figure 2a) from which the value $W_{\text{max}}$ was obtained. On the basis of the 3D visualizations and the values of the $W_{\text{max}}$ parameter assigned to them, we can identify significant changes taking place on the outer surface of the bent sample. Deep grooves characteristic for the moment before the formation of cracks have been found at the bending angle 36°. The first cracks found by means of microscope appeared at a bending angle equal to 40°. On the other hand, the macroscopic examinations enabled the detection of cracks only at much higher bending angle 44°. The obtained information was compiled with computer simulation results in order to assign plastic strain values for the given condition of the surface of the bent aluminum sheet.

![Fig. 1. Geometry of the press brake equipment (a) and distribution of $h$ values (b).](image-url)
4. COMPUTER SIMULATION OF THE V-BENDING PROCESS

The change of the sample surface along the bending line is connected with the occurring plastic strains. In order to know the values of these plastic strains, computer simulations of the V-bending process were performed by means of FEM. The material properties used in the simulations were obtained from the uniaxial tensile tests of the aluminum sheet EN AW-2017A. The variety of geometrical parameters of the bending tools used in simulations is given in the table 1. Each simulation was performed at different punch displacement in order to obtain selected bending angles after spring-back of the specimen. The choice of the tool parameters has been carried out in accordance with the recommendations of the press brake manufacturer. According to these recommendations for sheet thickness of 2 mm, the die opening width \( w \) should be from 10 to 25 mm, and the radius of the punch nose from 1.5 mm to 4 mm. In the computer modelling of V-bending process only the plastic strains crucial for the formation of the defects on the sample surface along the bending line were analyzed. Other disadvantages resulting from the applied geometry of tools such as indentation of the inner surface of the aluminum sheet by the rounded nose of the punch, damage of the outer surface of the aluminum sheet by the die opening surfaces or the s-shaped arms of bent aluminum sheet have been omitted.

Figure 3 shows a relationship between the bending angle and the plastic strains which occur along the bending line and are derived from a computer modeling. Figure 3a shows the results of simulations no. 1-3, using the selected die opening widths to the extent provided by the press manufacturer. The radius of 0.5 mm of the punch nose was used in these simulations. In all of the three simulations, there are only two phases of the V-bending process. The first is just curving of the sheet. During this phase the punch meets only with a limited area of the specimen surface. Then, the process proceeds to the wrapping phase, in which the sheet begins to map the shape of the punch. Wrapping phase lasts until the initiation of cracks in the aluminum sheet. The diagram shows that the increase in plastic strain is related to the die opening width. With the small size of the width the material wraps on the punch nose already at small values of the bending angle. Then the plastic strain increases considerably leading to the initiation of cracks. With increasing of the die opening width, the aluminum sheet wraps on the punch at higher values of the bending angle with a consequent reduction in strain increase. The presence of only two phases of the bending process is due to the small die opening width recommended by the press brake manufacturer for the given sheet thickness. The use of bigger die opening width would cause the transition of the bending process into the final setting phase, during which there is

Table 1. Geometry of tools in computer simulations of the V-bending process.

<table>
<thead>
<tr>
<th>Simulation number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>( w, ) mm</td>
<td>26</td>
<td>18</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>( r, ) mm</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>
a significant increase of the strain values at only a slight increase in the bending angle.

Figure 3b shows an influence of the punch radius on the process of the V-bending. There are the results of the simulations 3-8, in which the radii of the punch varied at a constant die opening width of 12 mm. For all radii the bending process is identical in the curving phase. The differences begin in the phase of wrapping. Wrapping the material on the punch causes the increase of strains dependent on the radius of the punch. The smaller the punch radius is, the more it can move in the direction of the die causing an increase in plastic strains. Wrapping the sheet metal around the front of the punch results in defining the plastic strain on the constant level, which is visible in the form of separating lines. The lines for the punches with \( r = 0.5 \) and 2 mm are identical until the end of the process. This is due to the fact that the sheet metal in both cases does not have a possibility to wrap the punches with such a small radius before finishing the bending process.

The results for V-bending process with different die opening widths \( w = 12 \) and 26 mm and the same punch radius \( r = 0.5 \) mm are shown in figure 4. The figures show the moment of the V-bending process in which plastic strain along the bending line for aluminum sheet EN AW-2017A reached values causing the creation of cracks detected by macroscopic vision systems. By comparing the figures, it can be noticed that bending angle between the specimen arms is much smaller in figure 4a than in figure 4b.

Increasing the punch nose radius from 0.5 mm to 6 mm with keeping the same die opening width 26 mm, bottom position of punches and bending angle, resulted in decreasing plastic strain in the outer layer of bent specimen, figure 5. These two specimens have quite different surface condition. As for the specimen shown in figure 5a, macroscopic cracks occur. However, the sample shown in figure 5b hasn’t had even deep grooves.
5. DISCUSSION OF RESULTS

Combining the results of numerical simulations and microscopic examinations has opened the possibility to prepare guidelines in the form of diagram shown in figure 6. There are presented changes of the surface condition of the aluminum sheet for a selected range of bending angles $\alpha$, depending on the die opening width $w$. With the increase of the die opening width, the grooves and cracks are observed at bigger bending angles $\alpha$. 

Fig. 5. The bottom positions of the punches enabling to obtain the same bending angles for punch nose radius $r = 0.5$ mm, $w = 26$ mm (a) and $r = 6$ mm, $w = 26$ mm (b).

Fig. 6. Changes in the surface condition of the aluminum sheet EN AW-2017A for a selected range of die opening width $w$ (punch nose radius $r = 0.5$ mm).

Fig. 7. Changes in the surface condition of the aluminum sheet EN AW-2017A for a selected range of the punch radii $r$ (die opening width $w=12$ mm).
Each marked area shown in the figure 6 corresponds to a different surface condition of the aluminum sheet. Starting from the bottom of the diagram, first there are the V-bending process parameters, for which deep grooves and cracks do not appear. Next area marks the bending angles for which deep grooves are detected on specimen surface. It proceeds with an area where the cracks are initiated in deep grooves. These cracks have a very small size. They are detected only by using the microscopic examinations. The last top area indicates the growth of cracks to the size allowing their detection by using macroscopic vision systems. The further progress of bending leads to an increase of the size of the cracks until they will be visible with the unaided eye.

Figure 7 shows an influence of punch nose radius on the surface condition of V-bent aluminum specimens. Deep grooves appear when the punch nose radius is below 4.5 mm, and the first cracks appear by using the punch nose radius 3.5 mm. The presented possibility of determining the moment of crack initiation allows to safely perform the bending process by avoiding the formation of defects on the surface of aluminum sheet, which would be unacceptable for aesthetic and strength reasons.

6. CONCLUSIONS

The selection of the tools used in the V-bending process of the aluminum sheet is determined not only by the geometrical defects of the product, but also the surface conditions along the bending line.

Information on changes in the surface condition of the bent specimens allows you to predict the moment of creating deep grooves or initiation of cracks and stopping the V-bending process before failure of the material.

Diagrams showing the surface condition of the aluminum sheet would allow the press brake operator to choose the bending angle corresponding with the geometry of tools depending on the surface condition of the aluminum sheet.

REFERENCES


WPŁYW GEOMETRII NARZĘDZI NA STAN POWIERZCHNI BLACHY ALUMINIOWEJ W PROCESIE V-GIĘCIA

Streszczenie

Gięcie jest jednym z najczęściej przeprowadzanych procesów obróbki plastycznej. Różnorodność materiałów i geometrii narzędzi w procesie gięcia powoduje trudności w określaniu momentu powstawania defektów w postaci pęknięć na powierzchniach blachy, a to one bardzo często stanowią kryterium decydujące o odrzuceniu produktu. Z tego powodu bardzo ważna jest ocena stanu powierzchni blach wzdłuż linii gięcia, ponieważ tam występują największe odkłuczenia, prowadzące do powstawania defektów w postaci bruzd czy pęknięć.

Bruzdy tworzą charakterystycznie zafalowaną powierzchnię. Wraz z powiększaniem kąta gięcia następuje pogłębianie się bruzd, aż w największych z nich następuje inicjacja pęknięcia. Od tego momentu rozpoznaje się rozrost i łączenie się pęknięć, które mają zasadniczy wpływ na wytrzymałość wyrobów oraz prowadzą do jego zniszczenia. Stąd określenie momentu powstawania nieciągłości materiału w postaci bruzdy ma kluczowe znaczenie w ocenie stanu powierzchni giętej blachy.

W opracowaniu przedstawiono wykorzystanie mikroskopu cyfrowego 3D do oceny stanu powierzchni próbek z blachy aluminumowej EN AW-2017A (PA6) poddanych V-gięciu. W celu uzyskania wizualizacji 3D powierzchni wykorzystano mapy głębiości uzyskane z analizy ostrości obrazu oraz tradycyjnym obrazem 2D. Wizualizacja 3D pozwala na ocenę stanu powierzchni blachy w sposób jakościowy. Aby jednoznacznie ocenić jej stan w sposób ilościowy, wprowadzono parametr Wmax. Uzyskiwany jest on z analizy zarysu geometrii powierzchni blachy pobranej z wizualizacji 3D. Jego wartość zależna jest bezpośrednio od największych bruzd aktualnie występujących na badanym obszarze, z których może następować inicjacja pęknięcia. Dzięki temu możliwe jest precyzyjniejsze określenie momentu powstawania pęknięcia. Dodatkowo przeprowadzono modelowanie komputerowe procesu V-gięcia w celu uzyskania informacji o odkłuczeniach plastycznych w najsilniej obciążonych warstwach materiału gię-
tych próbek dla szerszego niż w przypadku doświadczeń zakresu zmienności geometrii narzędzi do gięcia.

Przedstawiona możliwość określenia momentu inicjacji pęknięć pozwala na bezpieczne przeprowadzenie procesu gięcia poprzez uniknięcie powstawania defektów na powierzchni blachy, które są niedopuszczalne ze względów estetycznych oraz wytrzymałościowych.

Received: October 31, 2014
Received in a revised form: November 20, 2014
Accepted: November 23, 2014
CALCULATION OF FORMING LIMIT CURVE FOR GRADE 2 TITANIUM USING MODIFIED SAMPLE GEOMETRY

JULITA WINOWIECKA*, PIOTR LACKI

Czestochowa University of Technology, Dąbrowskiego 69, 42-202 Czestochowa
*Corresponding author: winowiecka@itm.pcz.pl

Abstract

Commercially pure titanium Grade 2 has good drawability and tensile strength of about 340 MPa. It is the most widespread grade of titanium in the industry. Grade 2 is characterized by good strength properties, low density, corrosion and external factors resistance. It is widely used in chemical, automotive and aerospace industries. In the aerospace industry titanium Grade 2 is used for production of fuselages, stringers, ventilating ducts, and many other parts. The subject of the work is the forming limit curve (FLC) widely used in industry to determine the possibility of the occurrence of draw-parts defects. FLC is determined based on the relationship between minor and major strains. It is a representation of the limit strains in the plane of the sheets, which in order to avoid cracks, cannot be exceeded during sheet - metal forming.

In the study forming limit curve was determined experimentally and the results were compared with the principal strain calculated in numerical simulations. Numerical simulations of the sheet - metal forming process were prepared in the PamStamp 2G v2012 program, using the finite element method. Forming simulation was carried out for specially designed samples with different lateral cut. The results of experimental studies at a depth corresponding to crack onset were compared with numerical calculations. Distributions of principal strain were determined for all specimens. The minor and major principal strains occurring in the forming samples before rupture onset were analyzed. Based on the results of numerical investigations forming limit curve for the titanium Grade 2 was determined. In experimental studies, in order to determine the plastic deformation ARAMIS system was used that enables non-contact measurements of three-dimensional deformations. The PamStamp 2G program and ARAMIS data acquisition process allowed for analysis of deformation and determination of the values of minor and major principal strains immediately before rupture. The numerical simulations considered technically dry friction and lubrication. This approach allowed for determining the effect of lubrication on strain distributions.

Key words: finite element method, titanium grade 2, forming limit curve

1. INTRODUCTION

Technically pure Grade 2 titanium is the most widespread grade of titanium in the industry. It is characterized by high mechanical strength, low specific gravity, corrosion and external factors resistance. However, titanium materials demonstrate poor suitability to plastic forming using rigid tools at room temperature (Adamus, 2009b). The stamping of Grade 2 titanium elements is connected with the occurrence of return spring-back, causing shape-dimensional inaccuracies (Winowiecka, 2013; Adamus et al., 2011).

The forming limit curve was developed by Keeler (1965) and Goodwin (1968) in the 1960s, while in the seventies a simplified technique of determining FLC was elaborated (Hecker, 1975). Since that time, FLCs are being continually improved and widely used in the optimization of metal sheet forming processes (Ávila & Vieira, 2003; Butuc et al., 2003; Narayanasamy & Narayananan, 2006, 2008). In 1967, Marciniak and Kuczyński (1967) developed a forming limit curve prediction model based on imperfections in sheet metal. This so-called M-K model has been the most commonly used for analytical prediction of FLC and has been the most crucial for further develop-
ment of formability assessment in sheet metal forming processes (Banabic, 2010). Marciniak et al. (2002) recapitulated the research on the M-K model and prepared the theory that the forming limit curve allows one to describe local processes of necking and tearing, but it does not depend on the boundary conditions. According to this theory FLC is a material property curve dependent on the strain state. These curves represent the limit strain, where exceedance is associated with drawpiece defects. FLC could be determined in experiment (Dong & Zhang, 2014). Experimentally determined forming limit curves are used in numerical simulations of metal sheet forming (Adamus, 2009a). In the numerical simulations of stamping, FLCs with the Keeler-Goodwin course are also utilized, which are defined based on experimental material properties and sheet thickness (Winowiecka et al., 2013; Winowiecka et al., 2014; Lacki et al., 2013; Adamus & Lacki, 2014).

The ARAMIS system enables two-dimensional and three-dimensional non-contact deformation measurements. Facets with a unique surface structure are applied on the research objects. The cameras make a series of pictures during the process allowing recording of the deformation. Software based on the images analyzes at each step changes in the surface structure. The photos of the two cameras are compared using pixel coordinates in the images. Then the system calculates the displacement of points on the individual images. On this basis it is possible to determine many distributions: the principal strain, thinning or displacement of the three axes. ARAMIS is used for example to measure three-dimensional deformation during stamping and stretching. In addition, the software enables calculation of the forming limit curve for a set of samples based on the minor and major strains (ARAMIS, 2011).

The use of experimental-numerical methodology is effective to predict the course of forming limit curves. Programs based on the finite element method, e.g. ABAQUS (Djavanroodi & Derogar, 2010) or PamStamp (Oh et al., 2011) are applied to determine the risk of cracks during stamping processes. They allow one to determine the distributions of plastic deformation, principal strain and thinning of the forming drawpiece. Hogström et al. (2009) and Situ et al. (2011) propose the use of the ARAMIS system to determine the experimental strain distribution and ABAQUS program to carry out numerical simulations. Badr et al. (2014) propose the use of optical strain measurement system “Autogrid Vario” to registration of deformation for specimens of various shapes.

## 2. AIM AND SCOPE OF WORK

The aim of the study was to determine the limit strain for Grade 2 titanium using samples with modified geometry. For the experimental studies, the ARAMIS system allowing non-contact dimensional measurement of deformation was used. Numerical simulations of the forming process of the sample set of square specimens with cuts were conducted using PamStamp 2G v2012, using the finite element method. Sample geometry which allowed appropriate crack location was developed. Square samples were designed with cutouts that fit the dimensions of the stamping tool. In this way cracks at the fillet radii of the die or under the blank holder were avoided. The experimental studies have included applying facets and stamping samples together with the registration process in the ARAMIS system. Numerical simulations were prepared by representing the real tool. The results of the calculations allowed us to determine the distribution of limit strain on the drawpiece surfaces for dry friction conditions and lubrication. By comparing the experimental and numerical results, the principal strain distributions in the drawpiece were analysed.

## 3. EXPERIMENTAL STUDIES

To determine the forming limit curve, forcing a hemispherical punch in specimens with different geometry was used. Previous studies led to the development of a set of specimens allowing a wider range of deformation. Initially, stamping sheet metal strips with different widths were used. Such samples, however, often cracked in the close vicinity of the blank holder, preventing proper deformation measurement. Using cut-out metal discs, it was possible to obtain easy crack measurement. The cracks frequently initiated near the center of the drawpiece, making easy and accurate measurements possible. In the case of some sample, sliding of the specimen out from under the blank holder caused by too little surface under the blank holder created a problem, therefore, square samples with cutouts were designed. For such a set, the area under the blank holder is the same for all the specimens, and in addition the material does not slide when using low
blank-holder forces. The samples crack near the axis of the sample, according to the assumption.

It was important to design samples adjusted to the dimensions of the real tool, thus reducing the cracks outside the measurement area. A press tool was used with the following dimensions: die hole 34 mm and filleting radius 5 mm, punch with a diameter of 28 mm and pressure ring hole 34 mm. A set of samples with square geometry shown in figure 1 was prepared for forming. The used sample geometry allows deformation differentiation. The resulting principal strain leads to determining FLC the measurement points. Thanks to the various geometries of the sample set, it is possible to vary the stress state using the same tool. The resulting sizes of the main strains are so diverse that it is possible to determine the measurement points on both sides of the $\varepsilon_2 - \varepsilon_2$ axis. The strain paths for different shapes of specimens is shown in figure 2. Used specimens enable realization of stress range from biaxial tension to pure shear.

**Table 1.** Results of experimental studies developed in ARAMIS system.

<table>
<thead>
<tr>
<th>Before crack</th>
<th>$R = 0$ mm</th>
<th>$R = 8$ mm</th>
<th>$R = 9$ mm</th>
<th>$R = 10$ mm</th>
<th>$R = 11$ mm</th>
<th>$R = 12$ mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height of drawpiece [mm]</td>
<td>14.0</td>
<td>11.2</td>
<td>10.4</td>
<td>9.4</td>
<td>8.7</td>
<td>7.4</td>
</tr>
<tr>
<td>Major strain [-]</td>
<td>0.515</td>
<td>0.745</td>
<td>0.684</td>
<td>0.610</td>
<td>0.562</td>
<td>0.553</td>
</tr>
<tr>
<td>Minor strain [-]</td>
<td>0.070</td>
<td>-0.482</td>
<td>-0.430</td>
<td>-0.354</td>
<td>-0.330</td>
<td>-0.322</td>
</tr>
</tbody>
</table>

Table 1 shows the values of the minor and major principal strains in the drawpiece and height of the drawpiece. The results enable plotting of the experimental forming limit curve.

4. NUMERICAL STUDIES

A numerical model of the tool was elaborated based on the real tool. The rigid shell model of the tool was prepared using Catia System v.5. The parts of the tool were imported into the PamStamp 2G v2012 program as IGS files. Four-node shell elements were automatically generated on the blank and the tool parts. The boundary conditions were as follows: no degrees of freedom for the die, movement in the Z direction for the punch and the blank-holder, all degrees of freedom for the blank. Additionally, the punch and the blank-holder has a defined displacement, and holding-down force was applied to the blank-holder.
The elastic plastic material model was assumed and anisotropic plasticity was used. Hollomon’s equation was used to define the relationship between the flow stress and plastic strain. Equation \( \sigma = K \cdot \varepsilon^n \) defined the hardening curves of the material. The mechanical properties used for defining the material model are given in table 2.

Table 2. Material properties used in definition of material model: \( E \) – Young’s modulus. \( R_y \) – yield point. \( \nu \) – Poisson ratio. \( \rho \) – specific gravity. \( K \) – material constant. \( n \) – strain-hardening exponent.

<table>
<thead>
<tr>
<th>Property</th>
<th>( E ) GPa</th>
<th>( R_y ) GPa</th>
<th>( \nu )</th>
<th>( \rho ) kg/m(^3)</th>
<th>( K ) GPa</th>
<th>( n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 2</td>
<td>105</td>
<td>0.236</td>
<td>0.37</td>
<td>4500</td>
<td>0.465</td>
<td>0.125</td>
</tr>
</tbody>
</table>

In the numerical simulations, the following frictional conditions were considered: lubrication \( \mu = 0.3 \) for contact surfaces between the punch and deformed material, \( \mu = 0.05 \) for the deformed material and blank holder, and deformed material and die, no lubrication - all contact surfaces \( \mu = 0.3 \). The holding-down force was equal to 1.0 kN.

In the experimental studies, the height of the drawn parts at crack moment was measured. The forming process of specimens with lateral cuts was simulated using Finite Element Analysis. The values of major and minor strains were indicated for the same height as the real drawn part at crack.

Figure 3 shows the minor and major principal strain distributions for the analyzed friction conditions for specimens with cutouts with an 8 mm diameter. For specimens for which lubrication was used, the largest major strain obtained is 25% higher in comparison to technical dry conditions. Using the technically dry friction conditions, the largest major strain occurs at the edges of the drawn piece, whereas using lubrication it occurs in the middle. The minor strain in both cases is located at the edges of the cutouts. Using lubrication, the minor strain increased by 40% compared to that obtained without lubrication.

Figure 4 shows the distribution of the major principal strains for the specimen with an \( R = 9 \) mm diameter cutout. The distribution of the major principal strains obtained in the numerical simulations is similar to that obtained in experimental studies. The highest values of the major strain occur at the cutout edge in both cases. Figure 4b presents the major strain distribution in the measuring area in the experimental studies. In this case, the maximum value is reached at one of the cutout edges.

The numerical studies have shown differences between the strain values and distribution depending on the friction conditions. Table 3 shows the numerical studies results of the minor and major principal strains obtained for the samples with cutouts. Using lubrication results in a higher value of major strain of 8÷25%, depending on the geometry of the sam-
ple, as compared with technically dry friction conditions. Lubrication causes the greatest FLC determined experimentally and numerically. For numerical results the values of major strain are higher than for experimental results.

![Fig. 4. Major strain distribution for conditions: dry friction in numerical simulation (a), measurement area in experimental studies(b) (R = 9 mm).](image)

Table 3. Results of numerical studies for dry friction and lubrication condition for stamping depth achieved in experimental studies.

<table>
<thead>
<tr>
<th>R=0 mm</th>
<th>R=8 mm</th>
<th>R=9 mm</th>
<th>R=10 mm</th>
<th>R=11 mm</th>
<th>R=12 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major strain [-]</td>
<td>0.581</td>
<td>0.642</td>
<td>0.644</td>
<td>0.587</td>
<td>0.591</td>
</tr>
<tr>
<td>Minor strain [-]</td>
<td>0.076</td>
<td>-0.210</td>
<td>-0.260</td>
<td>-0.161</td>
<td>-0.253</td>
</tr>
</tbody>
</table>

![Fig. 5. FLC for Grade 2 titanium obtained in numerical and experimental studies.](image)

5. CONCLUSIONS

On the basis of the experimental and numerical studies, the following conclusions can be made:

1. The use of modified square specimen geometry with cutouts allows proper conduct of the forming process. This gives a wide range of principal strains. Cracks that prevent correct measurement — under the blank holder or near filleting radius of die, should be avoided. Crack initiation occurs in the middle of the sample.

2. It is possible to determine the values of the principal strain using experimental-numerical methodology. The numerical process supporting the contactless strain measurement system allows for precise determination of the minor and major strains in the process of stamping specimens with different geometries.

3. The use of dry friction conditions reduces the depth of stamping and principal strain compared to lubrication conditions. Assuming the friction coefficient corresponding to lubrication, at least an 8% increase in the value of major strain was received.

4. Crack initiation with tool lubrication takes place in the middle of the area, compared with crack initiation near the cutout in dry friction conditions.

ACKNOWLEDGEMENT

Financial support of Structural Funds in the Operational Programme – Innovative Economy (IE OP)
financed from the European Regional Development Fund - Project "Modern material technologies in aerospace industry", Nr POIG.01.01.02-00-015/08-00 is gratefully acknowledged.

REFERENCES

ARAMIS, 2009, ARAMIS, ARGUS, SVIEW - FLC Computation v6.1.1 and higher.
ARAMIS, 2011, ARAMIS v6.3 and higher - User Manual.
Adamus, J., 2009b, Titanium and its Alloys as a Material Used for the Drawn-Parts, Inżynieria Materialowa, 30, 310-313.


WYZNACZANIE KRZYWEJ ODKSHAŁCEN GRANICZNYCH DLA TYTANU GRADE 2 PRZY WYKORZYSTANIU ZMODYFIKOWANEJ GEOMETRII PRÓBKI

Streszczenie

Technicznie czysty tytan Grade 2 charakteryzuje się dobrą tlocznioscą i granicą wytrzymałością ok. 340 MPa. Jest najbardziej rozpowszechnionym gatunkiem tytanu w przemyśle. Grade 2 cechuje się dobrymi właściwościami wytrzymałościowymi, małą gęstością, odpornością na korozję i czynniki zewnętrzne. Powszechnie wykorzystywanego jest w przemysłach chemicznym, motoryzacyjnym i lotniczym. W przemyśle lotniczym ze tytanu Grade 2 wykonuje się między innymi elementy kadłubów, wspornik, kanały wentylacyjne. Przedmiotem pracy jest wyznaczenie krzywej odkształceń granicznych (KOG), stosowanej powszechnie w przemyśle do określania możliwości wystąpienia wady wyloczku. KOG wyznacza się w oparciu o zależność pomiędzy odkształceniami głównymi najmniejszymi i największymi. Jest to reprezentacja odkształceń głównych w płaszczyźnie blach, które aby uniknąć pęknięć wyloczki nie mogą być przekroczone w trakcie kształtowania.

W pracy wyznaczono krzywą odkształceń granicznych stosując badania doświadczalne a wyniki odkształceń głównych odniesiono do symulacji numerycznych. Symulacje numeryczne procesu tloczenia przygotowano w programie PamStamp 2G v2012, wykorzystującego metodę elementów skośnych. Symulację tloczenia przeprowadzono dla specjalnie zaprojektowanych próbek o różnych wycięciach. Wyniki tloczenia z badań doświadczalnych przy głębokości tloczenia dla której uzyskano pęknięcie porównano z obliczeniami numerycznymi. Dla wszystkich próbek wyznaczono rozkłady odkształceń głównych. Analizowano odkształcenia główne największe i najmniejsze występujące w tloczonych próbkach przed pęknięciem. W oparciu o rezultaty badań numerycznych wyznaczono krzywą odkształceń granicz-
nych dla tytanu Grade 2. W badaniach doświadczalnych do określenia odkształceń plastycznych zastosowano system ARAMIS, umożliwiający bezkontaktowy trójwymiarowy pomiar odkształceń. Program PamStamp 2G oraz rejestracja procesu w systemie ARAMIS pozwoliły na analizę deformacji i wyznaczenie wielkości odkształceń głównych bezpośrednio przed pęknięciem. W symulacjach numerycznych rozważano tarcie technicznie suche oraz smarowanie. Dzięki temu było również możliwe określenie wpływu smarowania na rozkłady odkształceń.

Received: September 30, 2014
Received in a revised form: December 22, 2014
Accepted: December 2, 2014
NUMERICAL SIMULATION OF THE ROLL LEVELLING OF DP1000 STEEL USING A NONLINEAR COMBINED HARDENING MATERIAL MODEL

ELENA SILVESTRE*, ENEKO SÁENZ DE ARGANDOÑA, LANDER GALDOS, JOSEBA MENDIGUREN

Mechanical and Manufacturing Department, Mondragon University, Loramendi 4, 20500 Mondragon, Gipuzkoa, Spain
*Corresponding author: esilvestre@mondragon.edu

Abstract

The roll levelling is a forming process used to remove the residual stresses and imperfections of metal strips by means of plastic deformations. During the process the metal sheet is subjected to cyclic tension-compression deformations leading to a flat product. The process is especially important to avoid final geometrical errors when coils are cold formed or when thick plates are cut by laser. In the last years, and due to the appearance of high strength materials such as Ultra High Strength Steels, machine design engineers are demanding a reliable tool for the dimensioning of the levelling facilities. In response to this demand, finite element analysis is becoming an important technique able to lead engineers towards facilities optimization through a deeper understanding of the process. Nevertheless, the most commonly used material models, isotropic hardening models, are not able to reproduce the material’s Bauschinger effect and the final numerical results are not accurate enough.

In the present paper, the roll levelling simulation of a DP1000 steel is performed using a combined isotropic-kinematic hardening formulation, firstly introduced by Armstrong and Frederick and subsequently modified by Chaboche. For material parameters’ identification tension-compression tests and shear-tests have been realized and compared. Finally, the influence of the material model in the numerical results is analyzed by comparing a pure isotropic model and a combined Chaboche hardening model.

Key words: roll levelling, kinematic hardening, high strentgh steels, tension-compression test, shear test

1. INTRODUCTION

The development in the last years of new steel grades with high performances has been motivated by new tendencies in the automotive industry. Reducing the weight of a vehicle is a straightforward strategy to improve fuel economy, but it can potentially create safety problems. For that reason, efforts are concentrated in the development of new steel grades with a competitive strength/weight ratio (van der Wiel, 2012), such as DP-Advanced High Strength Steel (AHSS). However, the development of these materials has led to the apparition of undesirable phenomena during forming process which affect the quality of the final product (Banabic, 2010). Although AHSS shows good mechanical properties in terms of durability, strength, stiffness, good crash energy absorption, etc., there are limiting factors for the application of these steel grades: increased springback, poor formability and high level of residual stresses. Furthermore, forming forces, tool wear and crack appearance increase significantly (Mendiguren, 2012).

Prior to sheet metal manufacturing processes, metal sheets are subjected to hot and cold rolling, which determines thickness and mechanical properties. During this process, sheets adopt flatness defects and residual stresses appear inside the material. High level of residual stresses inside the sheet metal
at the end of the process can promote the effect of springback, and this could cause the deformation of the sheet during cutting and affect also other forming processes. Thus, flatness tolerances and materials specifications required by manufacturers cannot be usually met by the rolling process itself, and an additional step before forming operation is necessary. The equipment typically employed for this purpose and which is analysed in this paper is the roll levelling process. Roll levelling is a forming process that aims at correcting flatness defects and minimizing residual stresses. For this purpose, sheets are bent in alternate directions, passing from tension to compression, by a certain number of rolls with adjustable overlapping (figure 1). Flattening of the material is achieved by a selective elongation of the shortest material fibres which ensures strain equalization of all fibres across the width and the thickness, thus removing the initial shape defects (Doege et al., 2002).

Due to the complexity of the process and the search of an optimal roll levelling process, the use of finite element codes which simulate the process has increased significantly in the last years with the aim to reduce cost and optimize the process itself. The finite element method (FEM) is quite successful to simulate metal forming processes, but accuracy depends both on the constitutive laws used and their material parameters identification (Flores et al., 2007). In the last decades, complex hardening models have been developing aiming to properly predict the real material behaviour. For simple applications, isotropic hardening models are usually used assuming proportional expansion of the initial yield surface, according to the Swift and Voce ones. Kinematic hardening laws provide more sophisticated models, where yield surfaces preserve their shape and size but translate through the stress space. A common rule is the Armstrong-Frederick nonlinear hardening law which considers the Bauschinger effect and the transient behavior. Chaboche improved Armstrong-Frederick kinematic hardening model by creating backstresses through superposition of several kinematic models. There are other advanced models, such as Teodosiu and Hu (1995) or Yoshida and Uemori (2002), ones which improve the fitting of the experimental data but bring about difficulties to identify all the material parameters from experimental stress-strain curves.

Material models with mixed nonlinear isotropic and kinematic hardening laws have received increased attention due to their improved ability to predict the Bauschinger effect and cyclic hardening behaviors of the material (Shi, 2008). One of the most popular of such material models is the Chaboche and Lemaitre model (Lemaitre & Chaboche, 1994), which it is the result of the combination of both Voce isotropic hardening law and Armstrong-Frederick nonlinear hardening law. Different authors argue that the predicted results of numerical simulation are significantly influenced by the choice of the hardening model (Saenz de Argandoña et al., 2014).

Each model has its precise requirements in terms of experimental data and testing needed to identify its parameters. For example, isotropic hardening models are identified on the basis of experimental data obtained from monotonic test methods; however these models overestimate the hardening in reversal loading. The characterization of forming operations such as the roll levelling process, in which tensile-compressive deformations take places, cyclic loading experimental tests are usually used in order to consider kinematic hardening, which is able to describe the Bauschinger effect (Bruschi et al., 2014). Different authors have proposed several reverse loading tests, e.g. tension-compression test, pure bending test, three point bending or shear test (Brunet, 2001; Carbonnière et al., 2009). Tension-compression test is the most simple and straightforward test because stress-strain data are obtained directly during the course of the test and an inverse method is not necessary, as occurs in pure and three point bending test. Nevertheless, the test is difficult to perform, due to the tendency of the specimen to buckle in compression (Eggertsen & Mattiasson, 2011). Shear test provides also tension-compression data directly and it has been used by many authors due to the absence of necking and the large range of homogeneous strains (Rauch, 1998).

In this paper, two experimental procedures to characterize the hardening behaviour of DP1000 AHSS are presented. In particular, tension-
compression and simple shear devices have been developed in order to obtain cyclic stress-strain curves. The modeling of hardening behaviour has been carried out by means of the mixed hardening law developed by Chaboche and Lemaitre with the von Mises yield criteria. The material parameters obtained from the fitting have been implemented in a FE numerical simulation of roll levelling process. An experimental roll leveller prototype have been developed in order to validate the results from simulation with both models.

2. CONSTITUTIVE EQUATIONS

The Chaboche and Lemaitre hardening model (Lemaitre & Chaboche, 1994) has been combined with the von Mises yield criteria, since it is recommended to use for cyclic plasticity analyses and it is widely distributed in commercial FE-codes. The von Mises yield criteria can be expressed for the uniaxial loading case:

$$\phi(\sigma, X, \sigma_y) = |\sigma - X| - \sigma_y,$$

(1)

Where $\sigma$ denotes the stress tensor, $X$ is the back-stress tensor and $\sigma_y$ is the initial yield stress. It is a mixed isotropic-kinematic hardening law which describes the movement of the yield surface corresponding to the nonlinear kinematic hardening by means of the evolution of the backstress, and the change in the size of the yield surface, which is introduced by means of the initial value of the yield stress $\sigma_y$ and the isotropic variable $R$. In the proposed model, the evolution of isotropic hardening is defined in function of the accumulated plastic strain $d\varepsilon^p$ by the following law:

$$dR = b \cdot (Q - R) \cdot d\varepsilon^p,$$

(2)

where $Q$ and $b$ are material parameters and the accumulated plastic strain. The kinematic part was proposed by Chaboche and his co-workers. This model is based on a decomposition of the non-linear kinematic hardening rule proposed by Armstrong and Frederik (1966):

$$dX_i = \frac{2}{3} \cdot c_i \cdot d\varepsilon^p - \gamma_i \cdot X_i \cdot d\varepsilon^p,$$

(3)

Chaboche decomposed a stable hysteresis curve in several parts, and it was observed that increasing the material parameters of the hardening rule, a more accurate model was obtained. The use of three components has been recommended by several authors (Bari & Hassan, 2000; Mahmoudi et al., 2011).

3. EXPERIMENTAL PROCEDURE

3.1. Materials

Dual phase steel DP1000 consisting of a ferritic matrix containing a hard martensitic second phase in the form of islands has been used in this study. This material is often used in the automotive industry, in particular for the roof rails. It is supplied in 1.0 mm thickness.

3.2. Tensile test

To determine material properties, a uniaxial tensile test has been performed at 0º to RD. The specimens were cut following ASTM E 8M-04 standard. The experiments have been carried out at 0.05 mm/s on a Universal 5 Tn Zwick/Roell machine. The initial yield stress YS, the ultimate tensile strength UTS, the Young modulus E and the elongation are given in table 1.

<table>
<thead>
<tr>
<th>YS 0.2% (MPa)</th>
<th>UTS (MPa)</th>
<th>E (GPa)</th>
<th>Elongation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>922.19</td>
<td>1055.75</td>
<td>196.80</td>
<td>8.74</td>
</tr>
</tbody>
</table>

3.3. Tension-compression test

The test provides suitable stress-strain curves under small strain ranges, however the maximum plastic strain achieved with this method is limited. Another inconvenience of this test is the buckling for thin sheet in compression loading. For this reason a special tool to avoid the buckling has been developed and its description is presented below. The specimens were cut from sheet for the 0º rolling direction. The specimens are smooth rectangular with the same cross section as the specimen used in the tensile test and a calibrated length of 22.5 mm. The geometry of the specimens has been specially designed for this study with the aim to use them in a tool preventing buckling. Figure 2a shows the experimental test equipment used.

3.4. Simple shear test

A simple shear sample similar to the proposed by Miyauahi (1984) has been developed with a two gauged areas of 1.5x35.2 mm$^2$ (figure 2b). The device developed allows for a reverse simple shear test and can be implemented in any tensile test machine.
Monotonic test has been performed at 0º to RD in the same way that the tension-compression test. Sample surface has been marked with small black dots over a white plane surface. The optical measurement system Aramis® from GOM has been chosen to measure the deformation of the specimen. Surface strains are calculated from the deformation of the pattern relative to a reference point.

4. MATERIAL PARAMETER IDENTIFICATION

Identification of hardening parameters has been carried out by means of the Nelder and Mead (1965) minimization method, which is a nongradient optimization method. The program was implemented as a function of Matlab®, so that the objective function was defined to minimize the difference between the predicted stress values from the model and the experimental data:

\[
f = \text{Min} \frac{1}{n} \sum_{i=1}^{n} \left( \frac{\sigma_i^{\text{exp}} - \sigma_i^{\text{model}}}{\sigma_i^{\text{exp}}} \right) 100 \quad (6)
\]

where \( n \) is the number of experimental data, \( \sigma_i^{\text{exp}} \) is the stress obtained in experimental test and \( \sigma_i^{\text{model}} \) is the stress predicted by the proposed model. The identification method consists of the search of the optimal parameters which minimize the objective function (6). In particular, the model has 4 hardening material parameters, two corresponding to the isotropic equation \((Q, b)\) and other two corresponding to the kinematic equation \((C, \gamma)\).

5. RESULTS AND DISCUSSION

The optimization has been carried out for the experimental data obtained from the tension-compression test and from shear test. In both cases, the parameters obtained from resolution of the model equations by using know states of backstress, \( R \) and increment of plastic strain from experimental
data have been used as initial guesses. The initial parameters obtained are presented in table 2.

**Table 2. Initial parameters for the optimization process.**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q_0$ (MPa)</td>
<td>100.00</td>
</tr>
<tr>
<td>$b_0$</td>
<td>-51.74</td>
</tr>
<tr>
<td>$C_0$ (MPa)</td>
<td>55000.00</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>177.95</td>
</tr>
</tbody>
</table>

5.1. Parameter identification with tension-compression and simple shear test

Fitting of the mixed hardening model to the experimental data from tension-compression test and simple shear test are presented in figure 3a and 3b, respectively. Both figures show an excellent correlation between experimental and simulated data. A good prediction of the Bauschinger effect, cyclic softening and transient behavior are obtained. Material parameters obtained from both fittings are shown in table 3.

**Fig. 4. 2D finite element model geometry and sheet meshing.**

**Table 3. Material parameters of Chaboche and Lemaitre model obtained from tension-compression test.**

<table>
<thead>
<tr>
<th>Test</th>
<th>$Q$ (MPa)</th>
<th>$b$</th>
<th>$C$ (MPa)</th>
<th>$\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tension-compression</td>
<td>-460.02</td>
<td>250.78</td>
<td>80502.77</td>
<td>150.70</td>
</tr>
<tr>
<td>Simple shear</td>
<td>-299.86</td>
<td>113.00</td>
<td>69525.85</td>
<td>172.74</td>
</tr>
</tbody>
</table>

6. NUMERICAL SIMULATION AND VALIDATION OF THE MODEL IN A ROLL LEVELLING PROTOTYPE

6.1. Numerical simulation

2D numerical model has been developed in MSC Marc® software. The model consists of two rows of work introducing a proper deformation to the sheet. The work rolls rotate in order to push the sheet through them as it is shown in figure 4. The sheet has been discretized using a non-uniform mesh of elements with four integration points. The friction coefficient of 0.2 was taken.

The material parameters of Chaboche and Lemaitre model obtained in the previous section from tension-compression test and shear test were evaluated by simulating the roll levelling process using both set of parameters.

6.2. Experimental equipment

A 13-rolls leveller prototype was designed in order to validate numerical simulations. The prototype has several sensors to capture signals of different process variables: loading cells to measure the reaction force supported by the machine, torque sensor

**Fig. 5. Strain path in the surface of the sheet during the process (a), simulation result of force per roll (b).**
in the most critical roll, power consumption device and instrumented sheet with gauge strain at the surface in order to check the deformation of the sheet during the process.

6.3. Validation of numerical simulation of roll levelling with tension-compression and simple shear parameters

The strain path of a point at the surface of the sheet during the levelling process has been determined with both material models. In figure 5a the comparison of the numerical strain path with the experimental data is presented. A negligible difference between both models is perceived. Deviation between the experimental data and numerical results can be appreciated in the last rolls. This difference can be caused by the displacement of rolls during the test due to the reaction forces which the sheet applies in the work rolls. In terms of maximum plastic rate achieved during the process (percentage of thickness which achieves the platic range), only 1% difference was found between the simulation and the experimental data (74% and 73%, respectively).

Simulation using the parameters from tension-compression test has provided slighty lower values of force per roll than shear test simulation (figure 5b). The total force of the machine in the simulation has been calculated as the sum of the reaction forces of the top rolls and has been compared with the total force measured in the prototype. The values are presented in table 4.

### Table 4. Comparison between experimental and simulation values of force and strain.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Experimental</th>
<th>Simulation with tension-compression test parameters</th>
<th>Simulation with simple shear test parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Force (kg)</td>
<td>9000</td>
<td>9291</td>
<td>9973</td>
</tr>
<tr>
<td>Maximum plastic rate (%)</td>
<td>73.2</td>
<td>73.8</td>
<td>74.2</td>
</tr>
</tbody>
</table>

7. CONCLUSIONS

In order to develop an accurate methodology to develop hardening models for Advanced AHSS, such as DP1000, two methodologies of material characterization have been checked.

- Tension-compression and simple shear devices have proven to be a fast and easy test to recover cyclic reverse loading data, and an inverse method is not necessary.
- Chaboche and Lemaitre hardening model has been selected due to its ability to predict the Bauschinger effect and because it is implemented in most of finite element codes.
- The fitting to the experimental data from both characterization test has given different material parameters. Bauschinger effect and transient behaviour has been predicted in both cases.
- Numerical simulation of roll leveling process has been compared with experimental data obtained from a roll levelling prototype. The strain paths along the process obtained from simulation are quite similar to the experimental strain path (1% difference). Total force measured in tension-compression and simple shear simulation were 3.2% and 10.8% respectively higher than the experimental data.

ACKNOWLEDGMENTS

The work presented in this paper has been carried out in cooperation with FAGOR ARRASATE Scoop and the financial support of the INNPACTO National Programme for Public-Private Cooperation (Spanish Science and Innovation Minister).

REFERENCES


---

**SYMULACJA NUMERYCZNA WALCOWANIA POZIOMUJĄCEGO STALI DP100 Z ZASTOSOWANIEM NIELINIOWEGO MODELU UMOCNENIA MATERIALU**

**Streszczenie**

Walcowanie prostujące jest procesem odkształcania stosowanym do usuwania naprężeń szczątkowych i niedoskonałości pasma po walcowaniu, wykorzystującym odkształcenia plastyczne. Podczas procesu blacha jest poddawana cyklicznym odkształceniom rozciągającą-ściskającym, prowadzącym do otrzymywania produktu płaskiego. Proces jest szczególnie istotny, gdy w ostatnim etapie produkcji blacha jest zjwiana lub, jak w przypadku blach grubych, stosowane jest cięcie laserowe, gdyż pozwala uniknąć błędów kształtu. W ciągu ostatnich lat, ze względu na coraz szersze wykorzystywanie materiałów o wysokich własnościach wytrzymałościowych, takich jak stale typu UHSS, przed konstruktorami maszyn pojawiają się zadania projektowania niezawodnych i precyzyjnych narzędzi do walcowania prostującego. Numeryczna analiza procesu, wykorzystująca modelowanie metod elementów skośnych, stała się ważną techniką, pozwalającą inżynierom na optymalizację tego procesu poprzez jego lepsze, głębsze zrozumienie. Niemniej klasyczne modele umocnienia materiału, modele izotropowe, nie oddają efektu Bauschingera obserwowanego w trakcie walcowania prostującego, przez co ilościowe wyniki obliczeń nie są wystarczająco dokładne.

W niniejszej pracy, przeprowadzono symulację procesu walcowania prostującego dla stali DP100 wykorzystując izotropowy, kinematyczny model umocnienia materiału, wprowadzony po raz pierwszy przez Armstronga i Fredericka, a następnie znakomitym modelu Chaboche’a. Identyfikację parametrów modelu przeprowadzono na podstawie wyników testów rozciągania i ściskania oraz prób ściskania. Ponadto w pracy zbadano wpływ modelu materiału w obliczeniach numerycznych porównując wyniki otrzymane dla czystego modelu izotropowego i modelu Chaboche’a umocnienia materiału.

Received: November 4, 2014
Received in a revised form: December 2, 2014
Accepted: December 17, 2014

---

**INFORMATYKA W TECHNOLOGII MATERIAŁÓW**

**ŁÓDŹ**

SYMULACJA NUMERYCZNA WALCOWANIA POZIOMUJĄCEGO STALI DP100 Z ZASTOSOWANIEM NIELINIOWEGO MODELU UMOCNENIA MATERIALU

Streszczenie

Walcowanie prostujące jest procesem odkształcania stosowanym do usuwania naprężeń szczątkowych i niedoskonałości pasma po walcowaniu, wykorzystującym odkształcenia plastyczne. Podczas procesu blacha jest poddawana cyklicznym odkształceniom rozciągającą-ściskającym, prowadzącym do otrzymywania produktu płaskiego. Proces jest szczególnie istotny, gdy w ostatnim etapie produkcji blacha jest zjwiana lub, jak w przypadku blach grubych, stosowane jest cięcie laserowe, gdyż pozwala uniknąć błędów kształtu. W ciągu ostatnich lat, ze względu na coraz szersze wykorzystywanie materiałów o wysokich własnościach wytrzymałościowych, takich jak stale typu UHSS, przed konstruktorami maszyn pojawiają się zadania projektowania niezawodnych i precyzyjnych narzędzi do walcowania prostującego. Numeryczna analiza procesu, wykorzystująca modelowanie metod elementów skośnych, stała się ważną techniką, pozwalającą inżynierom na optymalizację tego procesu poprzez jego lepsze, głębsze zrozumienie. Niemniej klasyczne modele umocnienia materiału, modele izotropowe, nie oddają efektu Bauschingera obserwowanego w trakcie walcowania prostującego, przez co ilościowe wyniki obliczeń nie są wystarczająco dokładne.

W niniejszej pracy, przeprowadzono symulację procesu walcowania prostującego dla stali DP100 wykorzystując izotropowy, kinematyczny model umocnienia materiału, wprowadzony po raz pierwszy przez Armstronga i Fredericka, a następnie znakomitym modelu Chaboche’a. Identyfikację parametrów modelu przeprowadzono na podstawie wyników testów rozciągania i ściskania oraz prób ściskania. Ponadto w pracy zbadano wpływ modelu materiału w obliczeniach numerycznych porównując wyniki otrzymane dla czystego modelu izotropowego i modelu Chaboche’a umocnienia materiału.

---

W-TEMPER FORMING OF AA7075 ALUMINUM ALLOYS AS AN ALTERNATIVE TO THE WARM AND HOT STAMPING

ENEKO SÁENZ DE ARGANDOÑA1*, LANDER GALDOS1, RAFAEL ORTUBAY1, JOSEBA MENDIGUREN1, XABIER AGIRRETXE2

1 Mechanical and Manufacturing Department, Mondragon University, Loramendi 4, 20500 Mondragon, Gipuzkoa, Spain
2 Batz S. Coop., Torrea Auzoa, 2, 48140 Igorre, Bizkaia, Spain
*Corresponding author: esaenzdeargan@mondragon.edu

Abstract

As important light-weight structure material, aluminum alloys have been widely used in automotive and aerospace industries. In the last years, the manufacturing of parts with high strength and good dimensional accuracy has become the main objective in industrial applications. Within the available aluminum alloys, the 7xxx series has attract the interest of the industrial designers due to the high yield strength and ultimate tensile strength they present. However, the formability of these alloys in as-received industrial condition is very poor at room temperature and various studies are being carried out to develop efficient warm and hot forming processes to form them industrially using heated tools. In the present paper, the W-temper forming is studied as an alternative to the warm and hot forming processes. Heat treatment temperatures and critical times are presented and an industrial B-Pillar is formed to validate the new process. In the last chapter, the final mechanical properties of the part are reported, before and after a virtual e-coat process where the W-temper forming is compared with a hot stamping process.

Key words: AA7075, W-temper forming

1. INTRODUCTION AND MOTIVATION

Several examples exist where the principal OEMs have replaced medium strength alloys by ultra-high strength alloys to produce ultra-light body-in-white concepts. A very good example is the early ULSAB project started in 1995 and leaded by the principal steel makers to develop new advanced steels where a 25% of weight reduction was possible keeping the structural behavior. Several successful research projects have been realized since then to introduce new steels in the automotive industry. Components using very high strength Dual Phase, Complex Phase, TRIP and Martensitic steels are a reality and Press Hardened parts are already present in our cars which helps reducing the CO₂ emissions and fuel consumption. If we focus in a future horizon, the high bend stiffness and strength to weight ratio of aluminum alloys combined with their significant corrosion resistance and recyclability mark them as ideal candidates to replace the heavier steel components in the automotive industry.

Despite all these advantages, aluminum alloys still lag behind in their application due to their poor formability as compared to steels and their higher cost. Few OEMs consider the aluminum alloys as a candidate for the production of their body-in-white and full aluminum ones are limited to medium to high class vehicles. However, the multi material body-in-white concepts and the emerging new joining technologies could enable the use of ultra-high aluminum alloys if the lightweighting cost is low. Medium strength 5xxx and 6xxx series aluminum alloys are common and proven alloys for the auto-
motive structural parts and body-in-white in European cars. Good reviews have been realized by Kleiner et al. (2003) and Wang et al. (2012) where warm forming of these alloys is also included to increase their formability and form more complex parts. However, for some high security crash parts like the B-pillar, a higher level of strength to weight ratio is required to satisfy the roof crush and side impact standards and the 7xxx aluminum alloys are the next family to be studied to fulfill with the OEMs requirements. Other authors investigated 7xxx alloys related to forming and crash performance, respectively. They concluded that 7xxx-series alloys have potential to replace steel for structural components like A-pillar, B-pillar and side impact beams. The 7xxx-series alloys are age-hardenable alloys in which strengthening arises due to the formation of fine dispersed meta-stable precipitates during specific heat treatments such as T6, which is the most common industrial available condition.

The formability at room temperature is poor and new forming methods like the warm forming and hot stamping are under research and very few studies using these technologies have been published for the 7xxx alloys. One of the major problems associated with forming at elevated temperatures is degradation of high strength temper of the age-hardenable aluminum alloys, as reported by Lee et al. (2004) for AW-7075-T6. Regarding the forming process, Lee et al. (2004) studied the warm hydroformability of 7075 tubes between room temperature to 300°C. The results showed that sufficient elongation properties of high strength aluminum alloys could be achieved by the selection of pertinent pre-treatment conditions and deformation temperatures. Wang et al. (2012) presented a paper where material characterization and LDR tests were performed using the 7075-T6 material. They found that total elongation at fracture increased between 140°C and 220°C due to the increase in strain rate sensitivity which controls diffuse necking and prevents plastic strain from concentrating in a localized neck.

They further reported that the best drawing and stretching formability can be realized at temperatures between 180°C and 220°C, respectively. No scientific publication has been found where the hot stamping of the 7075 aluminum alloy has been studied. Fan et al. (2013) studied the hot forming of the 6A02 alloy using warm forming dies. The optimal temperature for the forming dies was reported to be near 250°C. On the other hand, Bariani et al. (2013) presented a paper where the hot stamping of the 5083 aluminum alloy was analyzed as an alternative to the super plastic forming process. Cool dies were used in the study and the formability of the material was analyzed at temperatures up to 500°C. The results of industrial trials carried out on an automotive component with a complex geometry confirmed that the forming temperature of 450°C can assure a geometrically sound component, with microstructural and mechanical characteristics comparable of those of the as-delivered blanks.

2. MOTIVATION AND RESEARCH APPROACH

Warm and hot forming often present advantages from the point of view of formability. However, industrial companies frequently avoid the use of these technologies due to its higher cost which is directly linked to the high cycle times, extra investment in heating ovens and fast lank positioning robots and bigger wear and galling problems of the tools as well as the complexity of the temperature control in the tooling. Due to the previous facts, in the present paper the W-temper process is put face to face with the hot forming process and the final properties of the stamped parts are compared. In the W-temper process, the aluminum AW-7075-T6 sheets are first solution heat treated and water quenched and subsequently formed at room temperature.

In order to draw the guidelines for the W-temper process design, tensile tests ranging from room temperature to 400°C are presented to identify the potential application field of the warm and hot forming. Formability is compared to the ones obtained using the W-temper approach which has been optimized using laboratory specimens. Using these results, the final industrial validation case study and the process variables are defined. A B-Pillar is formed using the Hot Stamping and the W-temper process variants and the final mechanical properties are compared after a simulated e-coat process.

3. EXPERIMENTAL TESTS AT LABORATORY LEVEL

3.1. Uniaxial testing at various temperatures

AW-7075-T6 sheets with a thickness of 1.6 mm and chemical composition (in wt.%) of 0.08 Si, 0.28 Fe, 1.57 Cu, 0.022 Mn, 2.35 Mg, 5.64 Zn, 0.19 Cr, 0.027 Ti, have been used in the current work. Tensile samples with a gauge length of 6 mm (see figure
were machined from the as-received sheet in the rolling direction. Tension tests were performed using a MTS 810 servo-hydraulic testing machine equipped with an electrical oven of the same brand. Each tensile sample was heated to the test temperature in 5 minutes before each tensile test. The elongation of the specimen was directly measured from the machine grips movement as very small differences were measured when using an extensometer for the needed forces. A detailed view of the experimental set-up is shown in figure 1b. Tension tests were performed at temperatures between room temperature and 400°C and at strain rates between 0.01 and 0.1 s⁻¹. The tests were repeated at least three times to ensure reproducibility.

Flow curves from the tensile tests of the as-received sheets are shown in figure 2.

The room temperature flow curves at different strain rates are almost identical, so it is observed the strain rate has a negligible influence on the strain hardening rate at this temperature. The flow curves at temperatures between 150 °C and 400°C are found to be sensitive to strain rate. True uniform strain decreases with increasing the temperature and the true fracture strain increases with increasing temperature only above 150°C. The strain hardening decreases a lot above 200°C being a very important parameter to obtain very deep drawn parts. It is known that the formability of the as received AW-7075-T6 alloy is very poor to produce drawn parts.

These tensile test results show the warm drawing temperatures are in the range of 150°C - 300°C since the process designer still can use the strain hardening of the material to pull the material from the blank holder areas to the drawing zones. At higher temperatures, the formability increases but there is a negligible strain hardening. Thus, these temperatures ranging from 300 to 500°C are the most promising ones to use the hot stamping process to pro-
duce semi-opened and small depth parts such as reinforment profiles, pillars and roof members.

3.2. Characterization of the solution heat treatment

Aiming to optimize the future industrial process of the W-temper forming strategy, as received material was longitudinally cut in small pieces of 100x20 mm and solution heat treated using different oven temperatures and times. 470°C was proven to be a good compromise for the reduction of the hardness in a small time. In order to identify the minimum oven time to be used using this oven temperature different solution heat treatment tests were performed varying the oven time followed by a water quenching and direct trough thickness hardness measurement from a transversally cut section in the transversal direction of the sheets. In figure 3 the hardness evolution after solution heat treatment at 470°C at different oven times is shown.

Each hardness data represents an average measurement of three through-thickness measurements from three different parts treated at the same conditions. Two minutes of oven are sufficient to reach the solution state since the sheet thickness is very small. On the other hand, ageing tests were also performed in order to define the maximum allowable time between the solution heat treatment and the forming of the part using same dimension samples and procedure. This is an important parameter for the press-shop since it defines if the material can be solution heat treated and stacked or must directly proceed to the forming operation. The samples were first solution heat treated at 470°C and 5 min of oven time and water quenched. Hardness measurements were later performed just after and at different times after the quenching. In figure 4 the ageing behavior of the alloy is shown. It is observed that the material starts hardening at approximately 10 minutes being the maximum allowable time before forming small.

3.3. Uniaxial testing at different W-temper conditions

After the optimization of the solution heat treatment temperature and time, uniaxial tensile tests were performed from solution heat treated sheets. 100x20 mm samples were first solution heat treated at 470°C and 5 minutes and subsequently water quenched. Just after the quenching, tensile samples were cut by mechanical shearing to guarantee a very fast preparation of the sample using a mechanical press. The specimens were cut following the ASTM E08 standard and using the sub-size specimen having 6 mm of calibrated width. To avoid ageing, the material was tested as soon as possible and always in less than 5 minutes to avoid hardening of the material. The test was repeated at least three times and the experimental work was completed testing the as-received material using the same procedure but air cooling the sample after the solution heat treatment.

This last test was performed to verify the reliability of this approach at an industrial environment and to avoid the water cooling step. The flow curve of the as-received material at T6 state and the solution heat treated and water quenched and air cooled conditions are shown in figure 5. The true fracture strain increases for the both solution heat treated conditions and the yield stress significantly decreases in comparison to the as-received material. The solution heat treated materials present a bigger strain hardening than the original material which is posi-
tive to increase the drawability. Among the solution heat treated materials, the water quenched one presents a bigger hardening than the air cooled one and the formability is bigger too. In comparison to the high temperature flow curves, the solution heat treated and water quenched material presents similar formability values to the one of 300°C. Moreover, the strain hardening is bigger in these new conditions so it seems logical to think the formability of the W-temper approach will be more suitable than the warm forming one for components having deep walls and drawings.

4. FORMING OF B-PILLAR

Finally, a real B-Pillar component was formed using an already available industrial tool to validate the different process variants (see figure 6). Three different approaches were tested: 1) forming of the part using the as received AW-7075-T6 material at room temperature, 2) forming of the part using the W-temper process and c) the hot forming of the as-received material. For the W-temper process, the solution heat treatment was realized at a temperature of 470°C. The sheet was heated during 5 min before water quenching and the forming operation was performed just before the quenching. For the hot stamping process, an oven temperature of 500°C was used since a temperature decrease of 50°C - 75°C was measured during the preheated sheet transport and the time needed from the position to the closing of the press. For both processes a Nabertherm 60/14 oven was used and air recirculation was used to ensure a good homogenization of the temperature along the sheet. The stamped B-Pillar parts are shown in figure 7.

Both hot stamping and W-temper strategies are valid to form a crack free and sound part. On the other hand a completely broken part is obtained using the as-received material at room temperature. Tensile test specimens were cut from the hot formed and W-temper formed B-Pillars to evaluate and compare the final properties of both forming strategies. Cutting area is shown in figure 7 where the specimen is colored in red. Additionally, a solution heat treated and air quenched part was also analyzed following the same approach although small crack were present in the deeper punch radii areas. The hot formed, w-tempered and air cooled specimens were subjected to an artificial heat treatment emulating an industrial e-coat process which was set to 160°C and 20 minutes. Another two w-temper formed specimens were treated using an artificial e-coat process
of 40 minutes and one hour to study the influence of this post forming step in the final component properties. All the flow curves of the formed and e-coated specimens are shown in figure 8. The graph clearly shows that it is impossible to recover the initial as-received material properties. Solution heat treated and air cooled material presents much lower properties than the solution heated + water quenched ones and so it seems not to be a possible process route to be followed in final production. The hot formed and e-coated process and the W-temper and e-coated processes present similar yield strengths of about 340 MPa. The hot formed parts present a slightly higher Ultimate Tensile Strength and hardening and a bigger elongation at rupture (up to 23%) being a priori a better candidate for crashworthiness applications. Regarding the e-coating times, the one hour treatment specimen presents bigger yield strength than the e-coating of 20 minutes being the difference not very significant.

Fig. 8. Final properties of the stamped parts.

5. CONCLUSIONS

a) Tensile test results show that the AW-7075-T6 alloy is temperature and strain rate dependent. The YS and UTS decrease with increasing temperature. Strain hardening is negligible at high temperatures and could limit the drawability of deep parts where the material must be pulled form the blank holder areas to the forming zones.

b) The solution heat treatment and water quenching followed by subsequent direct forming, called as W-temper forming in this paper, seems to be a suitable strategy for forming high complexity industrial parts. A solution heat treatment temperature of 470°C and treating time ranging from 2 to 5 minutes should be used to keep the process time cycle to the minimum.

c) Hot stamping and W-temper forming were applied for the stamping of an industrial B-Pillar with successful results. Final yield strength of the components after forming and e-coating is about 340 MPa. Although it is impossible to recover the initial as-received T6 condition material properties this values are well above the typical alloys used in the automotive industry, namely the 5xxx and 6xxx series alloys. At this stage, industrial manufacturers must evaluate the advantages and disadvantages of both processes and select the best one not only from the cost point of view but environmental impact, weight problems, etc. Further research is needed to compare both strategies using high deep wall components. The authors are currently conducting this work using a 100 mm spherical diameter LDR tooling.

ACKNOWLEDGMENTS

The authors thank the experimental work carried out by the final project student Jon Ander Lopez de Murillas Hurtado. Support from our industrial partner BATZ S.Coop., automotive tool maker, as well as the funding of the Basque Government to perform the Solintbo project are gratefully acknowledged.

REFERENCES


KSZTAŁTOWANIE STOPU ALUMINIUM AA7075
METODĄ W-TEMPER JAKO ALTERNATYWA DLA
TŁOCZENIA NA CIEPŁO IGORĄCO

Streszczenie

Stopy aluminium, jak ważny, lekki materiał konstrukcyjny, są szeroko stosowane w przemyśle motoryzacyjnym i lotniczym. W ostatnich latach, wytwarzanie części o wysokich właściwościach wytrzymałościowych i dużej precyzji stało się jednym z głównych celów zastosowań przemysłowych. Spośród dostępnych stopów aluminium, seria 7xxx jest szczególnie interesująca dla projektantów przemysłowych, ze względu na wysoką granicę plastyczności i wytrzymałość na rozciąganie. Jednakże, plastyczne kształtowanie tego typu stopów w warunkach przemysłowych, w temperaturze pokojowej jest bardzo ograniczone. Stąd prowadzonych jest wiele badań nad opracowaniem efektywnego kształtowania stopów aluminium w procesach na ciepło i gorąco z wykorzystaniem podgrzewanych narzędzi, mogących mieć zastosowania przemysłowe.

W niniejszej pracy analizowano proces kształtowania zwany W-temper jako alternatywę dla odkształcania na ciepło lub gorąco. Walidację nowego procesu wytwarzania przedstawiono dla wybranej, rzeczywistej części produkowanej dla przemysłu motoryzacyjnego, fragmentu słupka B. W ostatniej części pracy omówiono końcowe własności mechaniczne produktu, gdzie kształtowanie typu W-temper zostało porównane z procesem tłoczenia na gorąco.

Received: November 14, 2014
Received in a revised form: November 28, 2014
Accepted: December 1, 2014
DETERMINATION OF HEAT TRANSFER COEFFICIENTS UNDER CLOSED LOOP CONTROLLED CONSTANT CONTACT PRESSURES

LANDER GALDOS1*, ENEKO SÁENZ DE ARGANDOÑA1, JOSEBA MENDIGUREN1, RAFAEL ORTUBAY1, XABIER AGIRRETXE2, JOSÉ MIGUEL MARTÍN2

1Mechanical and Manufacturing Department, Mondragon University, Loramendi 4, 20500 Mondragon, Gipuzkoa, Spain
2Batz S. Coop., Torrea Auzoa, 2, 48140 Igorre, Bizkaia, Spain
*Corresponding author: lgaldos@mondragon.edu

Abstract

During hot forming and tailor tempering of boron steels, heat transfer between work-piece and dies has an important effect on the temperature distribution, microstructure evolution and mechanical properties of the final formed parts. In the present paper the interfacial heat transfer coefficient (HTC) has been determined at different contact pressures. Experimental tests have been realized in a SCHMIDT micro servo-press, which is able to compensate the thermal contraction of the blank and tools to precisely keep constant the contact pressure. Temperature evolution of the tools and the blank has been monitored with nine thermocouples.

For the determination of the heat transfer coefficient (HTC), an analytical-numerical method has been used leading to a fast and reliable calculation method able to determine the evolution of the HTC value during the cooling of the blank. This methodology allows the calculation of different HTC values in function of the contact pressure and the instantaneous tool temperature which will improve the accuracy of the numerical models and the prediction of the final properties of the components.

Key words: hot stamping, heat transfer coefficient

1. INTRODUCTION

The increase of the competitiveness of the automotive market together with the sustainability policies is continuously leading to an increment of the strength of the automotive components. One of the actual manufacturing processes that allows the forming of these very high strength components is the hot stamping process. In the hot stamping process, the boron alloy steel is formed at 950°C achieving very high formabilities at low force values. Once the part has been formed, it is maintained in the closed die forcing it to cool in few seconds generating a quenching process in the material. During this quenching process the austenite transforms into martensite increasing the mechanical properties of the final product, strengths up to 1500 MPa in geometrically complex components are achieved.

Based on the process described before, the automotive industry has also found the possibility of achieving tailored components where different areas of the component offer different mechanical responses: a process named as tailor tempering. These tailored mechanical properties are obtained by imposing different cooling ratios to the component. However, one of the main issues of the hot stamping process and even more important on the tailor tempering is the direct effect that the design of the stamping dies has on the final mechanical properties of the components. These final mechanical proper-
ties directly depend on the microstructure evolution of the material during the quenching process and therefore on the temperature evolution during the quenching process. Among the different variables involved on the thermal evolution the thermal exchange between the die and the component is one of the most important ones. This thermal exchange is governed by the heat transfer coefficient (HTC) which defines the ratio of heat exchange from one interface surface to the other. This way, the greater the HTC is the faster and more efficient the quenching process can be leading to greater mechanical properties in the final component.

In the last decade the HTC determination for boron alloy steel has been an issue under study Abdulhay et al. (2011a, 2011b), Bai et al. (2012), Caron et al. (2013), Chang and Bramley (2002), Hay et al. (2010), Koistinen and Marburger (1959), Lenard and Davies (1992), Malinowski et al. (1994), Merklein and Lechler (2006, 2009), Salomonsson et al. (2009), Tondini et al. (2011), Wang et al. (2012). A review of all the previous works shows that different techniques and methodologies have been developed for this objective, e.g. FEM inverse modelling, analytical-numerical models.

In this work the heat transfer coefficient (HTC) determination between the ORVAR SUPREME tool steel and the USIBOR 1500P boron alloy steel is carried out for different contact pressures. First, the experimental setup used to analyze the heat transfer on the interface of both materials is shown. Then, the numerical methodology followed to calculate the HTC is presented. Next, the determined HTC values are shown and compared to previous work tendencies. Finally, the main conclusions of the work are presented and the advantages and limitations of the presented methodology are defined.

2. EXPERIMENTAL SETUP AND TEST PROCEDURE

In order to emulate the hot stamping process, a laboratory schematic prototype was constructed (figure 1). As shown in figure 1, the workpiece (USIBOR 1500P steel sheet) is placed between the upper and the lower dies (ORVAR SUPREME tool steel). The dies and the workpiece have a cylindrical shape with a diameter of 50 mm and 30 mm respectively and the workpiece is 1.8 mm on thickness.

The experimental procedure is defined as follows. First, the workpiece is heated in an electrical resistance oven to 950°C. In order to allow the complete austenitization of the material, the temperature is maintained then during 5 min. After the austenitization is completed the transfer of the workpiece from the oven to the die is carried out. In figure 2 the positioning of the workpiece on the lower die can be observed.

Once that the workpiece is positioned on the die the forming step starts imposing a specific pressure on the material during the quenching. In order to analyze the influence of the pressure on the heat transfer behavior, different pressures have been studied ranging from 1 MPa to 10 MPa. The pressure is imposed using a high precision micro-press Schmidt 420 with a resolution of 0.0032 MPa. This pressure is maintained until the complete cooling of the workpiece (temperatures between 200°C and 300°C).

A data acquisition system, National Instruments 9215 hardware at 50 Hz, has been used in order to record the temperature evolution during the quenching process. Four TC Direct 12-K-1000-118-1-21-3P2L-1A30, 1 mm diameter thermocouples, have
been used to measure the temperature evolution on the lower die. These thermocouples are located to 2 mm, 4 mm, 6 mm and 8 mm from the interface between the workpiece and the lower die as represented in figure 1. On the workpiece on the other hand, and due to the limited space on the thickness, a single thermocouple on the middle of the thickness, 12-K-1000-118-0.5-2I-3P2L-1A30, 0.5 mm diameter thermocouple has been used. The hole for the thermocouple in the workpiece is critical and EDM has been used for its preparation.

3. NUMERICAL HTC IDENTIFICATION METHODOLOGY

Among the different HTC identification techniques shown on the literature, the analytical-numerical approach has been used in this work. The first hypothesis assumed in this work is that the thermal exchange characterization carried out on the experimental phase (figures 1 and 2) can be simplified as a one-dimensional problem.

The heat transfer coefficient is defined as the ratio between the heat flux and the temperature difference between the surface of the die and the surface of the workpiece

\[ h = \frac{Q}{T_{w1} - T_{d1}} \]  

where \( h \) represents the heat transfer coefficient (HTC) while \( Q \) is the heat flux through the interface. The temperature on the workpiece surface at the exchange interface is defined as \( T_{w1} \) and the temperature on the die surface at the exchange interface as \( T_{d1} \). However, from the experimental tests the temperatures are only known on the thermocouple positions. Therefore, three steps have to be followed for the HTC calculation. First, the temperature on the die surface has to be calculated from the data recorded by the thermocouples at the die. Next, the temperature on the workpiece surface has to be calculated from the thermocouple placed in the workpiece. Then, the heat flux through the interface has to be calculated. Finally, once that the flux and the difference in temperature are known, the HTC can be calculated following the definition shown in (1).

3.1. Temperature definition on the surface of the die

The heat conduction on the die is supposed to be governed by Fourier’s law

\[ \frac{\partial T}{\partial t} = \alpha_d \left( \frac{\partial^2 T}{\partial x^2} \right) \]  

where \( T \) represents the temperature, \( t \) is time, \( x \) is the space variable and the material properties are introduced by \( \alpha_d \) that it is defined as

\[ \alpha_d = \frac{k_d}{\rho_d c_d} \]  

where \( k_d \) is the thermal conductivity, \( \rho_d \) represents the density and \( c_d \) is the specific heat. In order to numerically solve the partial differential equation shown in (2) a backward time centered space method (BTCS) or full implicit method has been used. In this context the partial derivative of the temperature to the time, \( \frac{\partial T}{\partial t} \) is replaced by the first order backward difference and the spatial second partial derivative, \( \frac{\partial^2 T}{\partial x^2} \), is replaced by the BTCS method. Therefore the numerical approximation of the Fourier’s law results on

\[ \frac{T_{d,i}^{n+1} - T_{d,i}^{n}}{\Delta t} = \alpha_d \left( T_{d,i-1}^{n+1} - 2T_{d,i}^{n+1} + T_{d,i+1}^{n+1} \right) \left( \Delta x \right)^2 + o(\Delta t) + o(\Delta x^2) \]  

where \( T_{d,i}^{n+1} \) represents the temperature at the \((n + 1)\) increment on the \((i)\) point. The time increment from increment to increment is represented as \( \Delta t \), while the space increment from point to point is represented as \( \Delta x \). For the development of the numerical model the discretization of equation (4) has been truncated neglecting the error terms \( o(\Delta t) \) and \( o(\Delta x^2) \), these assumptions leads to a first order approximation on time and a second order approximation on space.

The definition of the different points used in this study are shown in figure 3. The first point on the die, \( T_{d1} \), is a point on the surface of the die (not measured in the experiments) and the next four points, \( T_{d2-5} \), correspond to the data measured in the experiments through the thermocouples.

In figure 4 the partial differential equation system to be solved is graphically presented. Supposing Dirichlet boundary conditions (the initial temperature of the surface is the same as the initial temperature of the first thermocouple \( T_{d2} \)) the problem to be solved leads to know the temperature of the first point at the end of the increment.
Previous authors worked on the basics that the thermal properties of the die material are constant as the die temperatures always maintain below 200ºC (Bai et al. 2012). However, the hypothesis of a linear evolution of the properties with the temperature is assumed in this study. From equation (4) and figure 4 the following mathematical relations can be obtained for the temperatures of the \( T_{d2} \)

\[
\frac{\Delta T_{d2}}{\Delta t} = \frac{a_d}{(\Delta x)^2} \left( T_{d1}^{n+1} - 2T_{d2}^{n+1} + T_{d3}^{n+1} \right) \tag{5}
\]

For the \( T_{d3} \) temperature

\[
\frac{\Delta T_{d3}}{\Delta t} = \frac{a_d}{(\Delta x)^2} \left( T_{d2}^{n+1} - 2T_{d3}^{n+1} + T_{d4}^{n+1} \right) \tag{6}
\]

and for the \( T_{d4} \) temperature

\[
\frac{\Delta T_{d4}}{\Delta t} = \frac{a_d}{(\Delta x)^2} \left( T_{d3}^{n+1} - 2T_{d4}^{n+1} + T_{d5}^{n+1} \right) \tag{7}
\]

It is not possible to develop these expressions for the first and fifth points without introducing a new error due to the truncation of part of the discretization of the \( \left( \frac{\partial^2 T}{\partial x^2} \right) \) term. Every temperature on equations (6) and (7) are known (thermocouple readings), leading to different material properties, \( \alpha_d \), to verify the equations. That is why in this study the hypothesis of a linear evolution of these properties is assumed. Therefore the properties of the material, \( \alpha_d \), for both temperatures \( T_{d3}^{n+1} \) and \( T_{d4}^{n+1} \) can be mathematically obtained. Once that the \( \alpha_d \) evolution is known (assuming a linear evolution in this first approach) the material properties at \( T_{d2}^{n+1} \) can be calculated.

Once that the material properties are known for \( T_{d2}^{n+1} \), the temperature at the end of the increment on the die surface \( T_{d1}^{n+1} \) can be calculated from equation (5).

### 3.2. Temperature definition on the surface of the workpiece

Since only a single temperature point on the center of the workpiece, \( T_{w2} \), is known, a different technique has been used to calculate the temperature on the surface of the workpiece, \( T_{w1} \). First, the symmetry property of the problem shown in figure 1 is taken into account. Therefore it is assumed that both workpiece surfaces (lower and upper) follow the same thermal evolution. If the discretization of the Fourier’s law, equation (4), is applied in this case, the temperature on the surface results

\[
T_{w1}^{n+1} = \frac{\Delta T_{w2}}{2\Delta t} \frac{(\Delta x w)^2}{\alpha_w} + T_{w2}^{n+1} \tag{8}
\]

where \( \Delta x w \) represents the distance between \( T_{w1} \) and \( T_{w2} \) points (half of the thickness) and \( \alpha_w \) represents the material properties defined as

\[
\alpha_w = \frac{k_w}{\rho_w c_w} \tag{9}
\]

The evolution of the density \( \rho_w \), specific heat \( c_w \) and thermal conductivity \( k_w \) of the USIBOR 1500 P has been previously studied by Hay et al. (2010). Both thermal conductivity and specific heat are supposed to be dependent on the martensite volume fraction (each phase, martensite and austenite, have different properties) while the density remains constant in both temperature and crystallographic phase. The specific heat of the austenite is

\[
c_{wa} = 426.0 + 0.1538 T \text{ (J/kg K)} \tag{10}
\]

while the one of the martensite

\[
c_{wm} = 311.2 + 0.439 T \text{ (J/kg K)} \tag{11}
\]
Fig. 5. Temperature evolution on both die (continuous line) and workpiece (line) for the different testing pressures: 1 MPa (a), 3 MPa (b), 5 MPa (c) and 10 MPa (d).

Fig. 6. Heat transfer coefficient values on normalized time of process.
where $T$ represents temperature in both cases. The thermal conductivity on the other hand results

$$k_{wa} = 16.27 + 0.010\ T \ (W/m\ K) \quad (12)$$

for the austenite while for the martensite is

$$k_{wm} = 83.73 - 0.245\ T + 5.79 \times 10^{-4}T^2 - 5.18 \times 10^{-7}T^3 \ (W/m\ K) \quad (13)$$

The martensite volume fraction $f_m$ can be calculated following the law Koistinen and Marburger (1959)

$$f_m = 1 - \exp[-0.011(673 - T)] \quad (14)$$

where the martensite volume fraction is set to zero for workpiece temperatures above 673 K.

The thermal conductivity has been calculated following Hay et al. (2010) as

$$k_w = f_m k_{wm} + (1 - f_m) k_{wa} \quad (15)$$

while the specific heat including the heat due to the austenite-to-martensite transformation is defined as

$$c_w = f_m c_{wm} + (1 - f_m) c_{wa} + (1 - f_m)(0.011)(1.35 \times 10^5) \quad (16)$$

Knowing the material properties at different temperatures together with the symmetry assumption and the numerical discretization of the Fourier’s law of equation (4), the temperature on the surface of the workpiece can be calculated.

In figure 5 the evolution of the temperature at different points on both the workpiece and the die for the different testing pressures is shown. On continuous lines the temperatures of the die points are shown, the lower temperature is the $T_{d1}$ while the higher one is the temperature on the surface of the die $T_{d1}$. The dashed lines on the contrary represent the temperature on the workpiece, the higher values corresponds to the center point $T_{w2}$ and the lower ones to the surface of the workpiece $T_{w1}$.

### 3.3. Heat flux on the interface

The last variable needed to calculate the heat transfer coefficient is the heat flux though the interface. On this regard in this work it has been assumed the heat flux as

$$Q = k_w \frac{T_{d1} - T_{d2}}{\Delta x} \quad (17)$$

where the thermal conductivity of the ORVAR SUMPREME depending on the temperature has been obtained by interpolating the suppliers specifications with a second order polynomial law

$$k_w = -9.528 \times 10^{-6}T^2 + 1.453 \times 10^{-2}T + 24.71 \quad (18)$$

where the temperature $T$ has been taken on the surface of the die.

### 4. RESULTS AND DISCUSSION

Following the numerical methodology shown in the previous section, the HTC value has been determined at each time increment of the cooling/quenching process. Figure 6 shows the different values of the HTC on the cooling normalized time. As can be observed in figure 5, each test has a different cooling time. Therefore in order to be able to compare the different evolutions, the normalized cooling time from 0 (before start of the cooling) to 1 (end of the cooling) has been used.

The main result observed in figure 6 is that the HTC values increase with the values of the contact pressure. It can also be concluded from figure 6 that there exists an increment of the HTC value during the process as well. These results agree with previous works where an increase of the HTC was shown in Bai et al. (2012), Caron et al. (2013), Merklein and Lechler (2009), Salomonsson et al. (2009). The increase of the HTC during the process was also previously reported as well by Chang and Bramley (2002).

### 5. CONCLUSIONS

In this work the heat transfer coefficient (HTC) for the interface between the boron alloy steel USIBOR 1500P and the ORVAR SUMPREME tooling steel has been determined under different contact pressures. An analytical-numerical method has been used for the HTC determination. This numerical methodology allows calculating the evolution of the HTC during the process under some hypothesis. The obtained results agree with the trends shown by previous authors showing an increase of the HTC with the contact pressure and a variation during the process. These results will allow the correct simulation of tailored hot forming parts.

### ACKNOWLEDGMENTS

The authors thank the experimental work carried out by Jon Ander Lopez de Murillas Hurtado. Support from our industrial partner BATZ S. Coop.,
automotive tool maker, as well as the funding of the Basque Government to perform the Solintbo project is gratefully acknowledged.

REFERENCES


ARTIFICIAL NEURAL NETWORKS AND RESPONSE SURFACE METHODOLOGY AS A TOOL FOR ANALYSIS THE SPINDLE TORQUE IN FSP PROCESS

MAREK ST. WĘGŁOWSKI

Institute of Welding, Bl. Czesława Str. 16-18, 44-100 Gliwice
*Corresponding author: marek.weglowski@is.gliwice.pl

Abstract

The article presents the effect of rotational and travelling speeds and down force on the spindle torque acting on the tool in friction stir processing (FSP) process. To find a dependence combining the spindle torque acting on the tool with the rotational speed, travelling speed and the down force, the artificial neural networks (ANN) and response surface methodology (RSM) were applied. Good correlation between experimental set and model was achieved. The best results were gained for the multilayer perceptron type 3-9-1. The results obtained in artificial neural network were compared with those through response surface methodology. Based on achieved results ANN, quadratic and linear models can be recommended to predict the value of spindle torque acting on the tool during FSP process carry out on alloy AlSi9Mg.

Key words: friction stir processing, artificial neural networks, response surface methodology, cast aluminium alloy

1. INTRODUCTION

A new production method of surface layers is the friction stir processing (FSP). This technology offers control over shaping the functional properties of materials being processed. FSP consists in heating and plasticising a material (parent metal) as a result of friction with a tool, provided (or not) with a pin, rotating and moving along an element surface subjected to processing. This technology originates from the friction stir welding (FSW) process, yet in comparison with this method, the phenomena taking place in the interface between the stirring area and the parent metal will have a decisive effect on the functional properties of a layer obtained through this process. The application, the course of the process, as well as the applied tools and equipment were discussed in the previous work (Ma, 2008). The research of the FSP of surface layers, so far has been focused mainly on the metallurgical analysis of microstructural changes in modified aluminium alloys (Węglowski, 2011; Charit & Mishra, 2005). However, from a practical point of view it is important to determine the impact of FSP conditions, i.e. a tool rotational speed, travelling speed, down force as well as the shape and type of tool on the moment acting on the tool, temperature in the stirring area, and the amount of heat generated in the stirring area. The heat generated in the area being processed and the level of plastic strain are factors having a decisive effect on microstructural changes, and, consequently, on the mechanical and functional properties of newly formed areas (Węglowski & Dymek, 2012).

Up to now many different methods was used to analyse the FSP/FSW processes. Among other experimental techniques (Darras, 2005; Węglowski et al., 2013; Węglowski & Dymek, 2013), analytical and numerical modelling (Kovacevic et al., 2012; Neto & Neto, 2013) are applied. The new and the...
most interesting ways is to use response surface methodology (Elangovan et al., 2008; Palanivel et al., 2011; Palanivel et al., 2012; Venkateswarlu et al., 2012) and neural network techniques (Asadi et al., 2012; Buffa et al., 2012; Tansel et al., 2010; Okuyucu & Kurt, 2007; Ebnonnasir et al., 2011). The response surface methodology (RSM) mainly was used to analyze the effects of FSW process parameters on quality or strength of FS welded joints. Elangovan et al. (2008) indicated that the RSM is useful technique to select and control the welding process parameter for obtaining maximum strength of FS welded joints also Palanivel et al. (2011) revealed that RSM allows to predict the mechanical properties of FS aluminum welded joints. RSM was also used to maximize the wear resistance (Palanivel et al., 2012) of FS welded dissimilar aluminum alloy. Venkateswarlu et al. (2012) FSP of AZ31B alloy has been modeled using RSM. Authors determined the total elongation values of tensile tested samples for various input parameters namely rotational and travelling speeds, and tool tilt angle. Recently, in the fields of FSP/FSW technology, the neural network technique has been used to predict the grain size in the stir zone (Asadi et al., 2012), mechanical properties of welded joints (Buffa et al., 2012) and optimization of the process (Tansel et al., 2010). An ANN model was also developed for the analysis and simulation of the correlation between the friction stir welding (FSW) parameters of aluminium (Al) plates and mechanical properties. The input parameters of the model consist of weld speed and tool rotation speed (TRS). The outputs of the ANN model include property parameters namely: tensile strength, yield strength, elongation, hardness of weld metal and hardness of heat affected zone (HAZ) (Okuyucu & Kurt, 2007). Ebnonnasir et al., (2011) determined the effect of FSP parameters on hardness of stir zone.

Even though sufficient literature is available on FSW of aluminium alloys, no study has been reported so far to correlate the process parameters and torque acting on the tool in friction stir processing cast aluminium alloy. Hence, in this investigation, the experimental results are compared with results obtained from the response surface methodology and neural networks.

2. EXPERIMENTAL SET UP AND METHODOLOGY

FSP was tested by means of an FSW station located at Institute of Welding in Gliwice. The station was composed of a conventional milling machine FYF32JU2, system for fixing test plates and a measurement head LOWSTIR (LOWSTIR - LOW cost processing unit for Friction Stir Welding). The tests were carried out using a tool with a shoulder of a 20mm diameter with pin. The tool was made of high speed steel grade H6-5-2. The tests were conducted on a 6mm-thick test plate made of aluminium casting alloy AlSi9Mg, using 30 technological parameters. The technological parameters were selected on the basis of previous experiments related to the FSW process and the milling machine operating conditions.

Table 1. Results of measurement during FSP process carry out on AlSi9Mg plate.

<table>
<thead>
<tr>
<th>No</th>
<th>Rotational speed $\omega$ [rpm]</th>
<th>Travelling speed $v$ [mm/min]</th>
<th>Spindle torque $M$ [Nm]</th>
<th>Down force $F_d$ [kN]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>112</td>
<td>112</td>
<td>141.84</td>
<td>23.92</td>
</tr>
<tr>
<td>2</td>
<td>560</td>
<td>112</td>
<td>33.43</td>
<td>12.26</td>
</tr>
<tr>
<td>3</td>
<td>900</td>
<td>112</td>
<td>17.82</td>
<td>8.87</td>
</tr>
<tr>
<td>4</td>
<td>1400</td>
<td>112</td>
<td>10.66</td>
<td>10.55</td>
</tr>
<tr>
<td>5</td>
<td>1800</td>
<td>112</td>
<td>8.20</td>
<td>14.46</td>
</tr>
<tr>
<td>6</td>
<td>112</td>
<td>224</td>
<td>156.82</td>
<td>27.67</td>
</tr>
<tr>
<td>7</td>
<td>560</td>
<td>224</td>
<td>41.43</td>
<td>14.62</td>
</tr>
<tr>
<td>8</td>
<td>900</td>
<td>224</td>
<td>25.59</td>
<td>11.05</td>
</tr>
<tr>
<td>9</td>
<td>1400</td>
<td>224</td>
<td>17.04</td>
<td>14.32</td>
</tr>
<tr>
<td>10</td>
<td>1800</td>
<td>224</td>
<td>12.09</td>
<td>19.36</td>
</tr>
<tr>
<td>11</td>
<td>112</td>
<td>560</td>
<td>192.37</td>
<td>40.90</td>
</tr>
<tr>
<td>12</td>
<td>560</td>
<td>560</td>
<td>59.81</td>
<td>19.64</td>
</tr>
<tr>
<td>13</td>
<td>900</td>
<td>560</td>
<td>45.34</td>
<td>23.43</td>
</tr>
<tr>
<td>14</td>
<td>1400</td>
<td>560</td>
<td>27.69</td>
<td>23.38</td>
</tr>
<tr>
<td>15</td>
<td>1800</td>
<td>560</td>
<td>21.07</td>
<td>24.04</td>
</tr>
<tr>
<td>16</td>
<td>112</td>
<td>710</td>
<td>209.0</td>
<td>46.51</td>
</tr>
<tr>
<td>17</td>
<td>560</td>
<td>710</td>
<td>67.75</td>
<td>22.63</td>
</tr>
<tr>
<td>18</td>
<td>900</td>
<td>710</td>
<td>52.00</td>
<td>25.15</td>
</tr>
<tr>
<td>19</td>
<td>1400</td>
<td>710</td>
<td>32.47</td>
<td>26.39</td>
</tr>
<tr>
<td>20</td>
<td>1800</td>
<td>710</td>
<td>22.81</td>
<td>25.44</td>
</tr>
<tr>
<td>21</td>
<td>112</td>
<td>900</td>
<td>229.0</td>
<td>53.71</td>
</tr>
<tr>
<td>22</td>
<td>560</td>
<td>900</td>
<td>75.10</td>
<td>21.11</td>
</tr>
<tr>
<td>23</td>
<td>900</td>
<td>900</td>
<td>57.43</td>
<td>28.76</td>
</tr>
<tr>
<td>24</td>
<td>1400</td>
<td>900</td>
<td>34.80</td>
<td>29.51</td>
</tr>
<tr>
<td>25</td>
<td>1800</td>
<td>900</td>
<td>21.57</td>
<td>27.19</td>
</tr>
<tr>
<td>26</td>
<td>112</td>
<td>1120</td>
<td>253.0</td>
<td>62.06</td>
</tr>
<tr>
<td>27</td>
<td>560</td>
<td>1120</td>
<td>81.6</td>
<td>28.72</td>
</tr>
<tr>
<td>28</td>
<td>900</td>
<td>1120</td>
<td>59.57</td>
<td>32.21</td>
</tr>
<tr>
<td>29</td>
<td>1400</td>
<td>1120</td>
<td>40.02</td>
<td>31.80</td>
</tr>
<tr>
<td>30</td>
<td>1800</td>
<td>1120</td>
<td>30.14</td>
<td>36.30</td>
</tr>
</tbody>
</table>
range (FYF32JU2). The range of travelling and rotational speeds are given in table 1. The plates were fixed to the machine with special grips and then processed. The roughness and surface quality of test plates were similar to the qualities after milling. The plates were not cleaned.

During experiments the mean values of the spindle torque and down force were calculated from 100 points in the area of the fully stabilized FSP process.

3. RESULTS AND DISCUSSION

It should be noted that the signals recorded during FSP are depended on tool geometry, parameters of the process, parent material, measurement system as well as cooling and clamping system. The influence of the rotational and travelling speeds on the torque acting on the tool is shown in figure 1 and 2, respectively. The data for rotational speed have been least square fitted with semiempirical relation:

$$M = a \exp\left(-\frac{\omega}{b}\right) + c$$  \hspace{1cm} (1)

![Fig. 1. Influence of rotational speed on the spindle torque acting on the tool.](image)

Function $M(\omega)$ is presented in figure 1 and the results of calculations with equation (1) are given in table 2. While the data for travelling speed have been least square fitted with linear empirical relation. The results are presented in figure 2 and calculations in table 3. The fitting is rather good.

![Fig. 2. Influence of travelling speed on the spindle torque acting on the tool.](image)

As can be seen the spindle torque strongly depends on the rotational speed of the FSP tool. This is due to the fact that the rotational speed stimulates the process’s temperature (Węglowski et al. 2013) within the FSP volume (modified material). Temperature linearly increases with the increase of the rotational speed. Hence, the friction coefficient also decreases. Higher temperature causes a decrease in the material resistance for the travelling tool. It would be expected that this decreases the torque. However, the torque is less affected by the change in the travelling speed. Such behaviour can be rationalized when assuming that for a constant rotational speed and decreasing travelling speed, the volume of material being processed per each tool revolution decreases, hence the heat is generated in a smaller volume, and this in turn may lead to rise in the temperature and decrease in the flow stress. Modest influence of the travelling speed on the torque is likely caused by a weaker relation between the travelling speed and temperature compared to the influence of the rotational speed on temperature. During the experiments, the penetration depth was kept constant (control by the machine operator). Hence, the value of the down force depends on rotational and travelling speeds and also machine operator. The influence of the down force on spindle torque is shown in figure 3. It is evident that the increase of down force causes the increase in the torque. The

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Travelling speed, mm/min</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>112</td>
</tr>
<tr>
<td>$a$, Nm</td>
<td>186.42±15.11</td>
</tr>
<tr>
<td>$b$, mm/min</td>
<td>344.19±69.91</td>
</tr>
<tr>
<td>$c$, $10^{17}$ Nm</td>
<td>7.20±6.41</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.98</td>
</tr>
</tbody>
</table>
data for down force have been least square fitted with linear empirical relation. The results of calculations are given in table 4.

Table 3. Fitted values of the linear dependence of spindle torque and travelling speed for the selected rotational speed, the function is presented in figure 2.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Rotational speed, rpm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>112</td>
</tr>
<tr>
<td>$a$, Nm·min/mm</td>
<td>0.11±0.0</td>
</tr>
<tr>
<td>$b$, Nm</td>
<td>131.04±0.83</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.99</td>
</tr>
</tbody>
</table>

The spindle torque is depended on rotational and travelling speeds, down force, type and shape of the tool, and the modified material. In the present work the response surface methodology (RSM) has been applied in order to find a dependence combining the torque acting on the tool with the rotational speed in a wider range (112÷1800 rpm), the travelling speed in the range of 112÷1120 mm/min and down force. RSM has been built in Statistica software. Rotational and travelling speeds and down force were introduced as independent variables. During analysis of the FSP process a first and second-order response surface models were used for mathematical modelling. The interactions between the these variables were assumed in the models.

The calculation results of the regression coefficients for the linear (2) and quadratic models (3) are shown in table 5. The calculation results indicate the significance level $p$ where linear main effects are statistically significant ($p<$0.05), suggesting that a linear model containing the interaction is sufficient while the quadratic parameter $B_9$ at the down force is statistically insignificant ($p>$0.05).

In the third step to analysis the interaction between selected FSP parameters and spindle torque acting on the tool the artificial neural network techniques was used. In this present study, BFGS algorithm was applied. The architecture of ANN used in this study is 3-n-1, with 3 corresponding to the input values, n to the number of hidden layer neurons and 1 to the output. Statistica ver. 10 was used for train-
ing the network model for spindle torque prediction. The training parameters used in this investigation are listed in Table 6. The neural network described in this work, after successful training, was used to predict the spindle torque acting on the tool during friction stir processing of AlSi9Mg aluminium alloy within the trained range.

Table 5. Results of calculation of regression coefficients for linear model.

<table>
<thead>
<tr>
<th>Models</th>
<th>The regression coefficients</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$B_0$</td>
<td>$B_1$</td>
<td>$B_2$</td>
<td>$B_3$</td>
<td>$B_4$</td>
<td>$B_5$</td>
<td>$B_6$</td>
<td>$B_7$</td>
<td>$B_8$</td>
</tr>
<tr>
<td>linear</td>
<td>-13.35</td>
<td>-0.09</td>
<td>-0.04</td>
<td>6.68</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>quadratic</td>
<td>27.01</td>
<td>-0.06</td>
<td>0.0</td>
<td>4.11</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.07</td>
</tr>
<tr>
<td>$p$</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.02</td>
<td>0.0</td>
<td>0.01</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$p$</td>
<td>0.0</td>
<td>0.0</td>
<td>0.03</td>
<td>0.0</td>
<td>0.0</td>
<td>0.02</td>
<td>0.0</td>
<td>0.02</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Remarks: $p$ – probability, for the clarity of the writing the value were rounding.

Table 6. Parameters of ANN used in the investigation.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Step 1</th>
<th>Step 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Networks to train</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>Networks to retain</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Type of networks</td>
<td>multilayer perceptron</td>
<td></td>
</tr>
<tr>
<td>Random sample size [%]</td>
<td>Training</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Validation</td>
<td>15</td>
</tr>
<tr>
<td>Type of multilayer perceptron</td>
<td>3:2-9:1</td>
<td>3:2-8:1</td>
</tr>
</tbody>
</table>

Fig. 4. Comparison of experimental and calculation results.

Comparison of the results for experiments, neural networks and response surface methodology is shown in Figure 4. The results indicate that the best fitting was achieved for ANN net type 3:2-9:1. Comparable results can be observed for a quadratic model. This result is caused by the fact that for complex model the interaction between input data were not omitted, also travelling speed was assumed as significant factor for value of torque. Based on achieved results ANN and quadratic model can be recommended for predicting the value of spindle torque acting on the tool during FSP process being carried out on cast aluminium alloy AlSi9Mg. It should be noted that from practical point of view the linear model should be also enough good solution.

4. CONCLUSIONS

The present study has examined the relationship between selected parameters of FSP process and spindle torque. The conclusions are as follows:

− the increase in the rotational speed decreases the torque acting on the tool,
− the increase in the travelling speed and down force increases the torque acting on the tool,
− the relationship between rotational speed and spindle torque can be fitted by the exponential function, while for travelling speed and down force by the linear,
− the surface response methodology is a useful technique to determine the impact of the parameters of the process on the spindle torque. In this study, a square model with interaction assured the better fitting,
− the results of ANN models indicate that more accurate in estimating the values of torque is multilayer perceptron type 3:2-9:1 than 3:2-8:1.

ACKNOWLEDGEMENTS

This work has been performed with funding from Polish Ministry of Science and Higher Education in the frame of statutory activity of Instytut Spawalnictwa (Institute of Welding).

REFERENCES

Sztuczne sieci neuronowe i metoda powierzchni odpowiedzi jako narzędzia do analizy momentu obrotowego w procesie FSP

Streszczenie

W artykule przedstawiono wpływ prędkości obrotowej i prędkości przesuwanego oraz siły docisku na moment obrotowy działający na narzędzie w trakcie procesu tarciowej modyfikacji warstw wierzchnich FSP. Do wyznaczenia zależności łączącej moment obrotowy działający na narzędzie z prędkością obrotową, prędkością przesuwu i siłą docisku, zastosowano sztuczne sieci neuronowe (SSN) i metodę powierzchni odpowiedzi (RSM). Osiągnięto zgodność pomiędzy wartościami z badań doświadczalnych i modelami. Najlepsze wyniki uzyskano dla perceptronu wielowarstwowego typ 3-9-1. Uzyskane wyniki dla sztucznej sieci neuronowej porównano z tymi, które uzyskano przy zastosowaniu metody powierzchni odpowiedzi. Na podstawie uzyskanych wyników SSN, modele liniowe i kwadratowe mogą być zalecane do przewidywania wartości momentu obrotowego działającego na narzędzie podczas modyfikacji FSP stopu AlSi9Mg.

Received: October 5, 2014
Accepted: November 12, 2014
INTELLIGENT INTERFACE FOR DECISION SUPPORT SYSTEM IN METALLURGICAL DOMAIN

KRZYSZTOF REGULSKI1*, GABRIEL ROJEK1, DOROTA WILK-KOŁODZIEJCZYK1, STANISŁAWA KLUSKA-NAWARECKA2, EDWARD NAWARECKI1, GRZEGORZ DOBROWOLSKI1

1AGH University of Science and Technology, al. A. Mickiewicza 30, Krakow, Poland
2Foundry Research Institute, Zakopianska 73, Krakow, Poland
*Corresponding author: regulski@agh.edu.pl

Abstract

Extensive decision support system composed of several modules providing different functionalities for the user can be a challenge in designing of a usable, plain and consistent user interface. A characteristic of advisory systems in the domain of metallurgy is their complex architecture designed to deal with multiple types of tasks in many areas. The objective of this paper is to present the possibility of creating an intelligent interface providing the user with complete freedom to act without precise knowledge of the multiple functionality of the system. The paper presents the architecture of developed advisory system in the domain of metallurgy enriched with the description of mechanisms, which enable obtaining an intelligent user interface. The intelligent user interface should guide the user at navigation through all system modules in order to meet all needs that are the reason for this system usage. In designing of such interface, the Case-Based Reasoning methodology is proposed to be used. Case-Based Reasoning ensures a pattern for building systems that are able to learn on the basis of human actions associated with human decisions and use such knowledge to resolve the currently appearing problems. Each user query directed to the system, and each of his decisions about the choice of functionalities of the system modules supplies the system with the crucial information on proper use of the whole decision support unit related with the needs of a user. Skillful use of this knowledge can provide convenient and efficient performance of intelligent interface.

Key words: intelligent interface, metallurgy, decision support system, Case-Based Reasoning, information technology

1. INTRODUCTION

For a few years, Foundry Research Institute in Krakow has been leading an international project entitled: "Developing solutions for the conceptualization and sharing of knowledge components about foundry technologies in the context of innovation and improvement of production processes" (No. 820/N-Czech Republic/2010/0), within the framework of which system modules have been developed constituting information and functional equivalent of the components of knowledge about the foundry technology. These modules make up the platform for knowledge sharing, which provides the user with a constant supply of current information and knowledge in the field of casting technology, derived not only from the literature, but also from other sources, such as expert works, databases or algorithmically acquired knowledge from online sources.

The information system can largely automate the process of decision-making supplying to the user knowledge of the issues, which are technological problems after a previous dialogue with the user to establish the facts. It can be anticipated that the problems the user will want to solve may involve a variety of issues and provide very different tasks for the system. What emerges is the goal that one has to put in front of the system, and in particular in front of its interface - design knowledge base in such
a way that it is interactive, thus enabling the user to get through dialogue with the system only the knowledge that is necessary to solve the problem and the system to anticipate on the basis of this dialogue what kind of knowledge the user has in mind.

A key element of the system from the point of view of the user becomes the interface which enables using diverse knowledge stored in databases that make up the local knowledge base. Within the framework of the discussed system, an integrated knowledge base of this type will hold: 1) literature and scientific publications, 2) materials (articles, data) obtained from the Internet, 3) various types of production and experimental data gathered by the users, 4) the accumulated knowledge in the form of rules and ontological models, 5) data on the characteristics of the materials, including dilatograms.

So rich a variety of document types will require diverse approaches and many record formalisms, as well as a preprocessing of user queries. Imagine a user who plans to use the system. He begins by entering a single phrase, a search query. It is expected that the system will return the output documents related to this query. They are supposed to be publications, photographs, data, reports and rules related to the topic. The task of the system in such a situation is not only going through a variety of integrated resources looking for the searched term, but also finding topics related to this term, i.e. "similar" topics. For the user, this process has to be transparent. As an output the user may obtain a list of suggested documents divided into sections comprising publications / photos / data / rules.

Without an intelligent interface, the user is forced to use the system as a set of modules, in which each has a separate query window and an individual way of interaction with the user - sometimes it is enough to give the search query, sometimes the user has to select a different mode of action (give domain list, select the set of descriptors, or click on a properly formulated query to the database).

This what happens inside such a system becomes crucial, especially during the integration of different sources of knowledge and semantic integration, i.e. the search for documents thematically "similar" to the searched query.

An example of the flow of events in a typical case of using the system with intelligent interface could look like this:
1. User opens the system interface.
2. User enters the subject.
3. System (interface) analyzes the subject
   a) searching for similar queries in the history of use,
   b) searching for data on the Internet and in the documents and databases in natural language,
   c) screening databases of materials if the user has mentioned the name of the material,
   d) searching data on treatment if the name of the technology has been mentioned,
   e) retrieving statistical data about the searched resources,
   f) searching alternative sources of knowledge (source materials, rule-based expert knowledge, database of publications, database of innovation, designs and patents, database of dilatograms, etc.).
4. The results are saved in XML files.
5. The results are displayed in a proper form.

The system not equipped with an interface agent would not have had step 3 in exchange for what in step 2 it could have: "the user selects an appropriate module of the system and enters his question in an appropriate form". Our goal is to design an effectively operating system which, based on information provided by the user (query, selection of module, automatically collected history of interaction with the system), will carry out the task of knowledge management, i.e. the location, acquisition and sharing. Such an interface should be fully compatible with the modules directly cooperating and should be as much as possible susceptible to subsequent modification or development. It will constitute in itself a learning system based on user behavior.

2. METALLURGICAL-DOMAIN DECISION SUPPORT SYSTEM

Implemented in the Foundry Research Institute in Cracow, the advisory system has a modular architecture, which consists of the following modules already described in most of other publications:
- CastSemaWiki - a platform for sharing and storing of documents based on the wiki mechanism, enriched with module for semantic annotation and semantic search based on the domain ontology for cast iron (Haratym et al., 2013).
- Module for the integration of material data for cast iron along with the application which allows collecting the experimental data on cast iron treatment and changes in its properties under the influence of this treatment (figure 1) (Kluska-Nawarecka et al., 2013).
− System for information retrieval from the Web (SWATLocalCrawler) - an automatic system for searching the web data resources to extract information about the foundry products. The detectors operate in conjunction with SWATLocalCrawler and allow searching for predefined patterns in the content of Web pages in domains that are selected for searching (Opaliński et al., 2011).
− InnoCast Database - a relational database with a web application on innovation, patents, publications and research projects issued in recent years for problems related to forging, rolling and casting.
− Dedicated ETL module - system for integration of data from metallurgical databases (e.g. SINTE: database of publications in the field of casting, NORCAST: database of standards, and InnoCast database) consistent with the ETL methodology - Extract, Transform and Load (Wilk-Kołodzieczyk et al., 2014).
− CastExpert + - the existing and developed at the Foundry Research Institute expert system in the field of diagnosis of casting defects enhanced with image simulations in the field of iron casting and multiscale modeling (Nawarecki et al., 2012).
− Domain ontology for cast iron supplemented by a set of rules in SWRL regarding changes in cast iron under the influence of treatment (Kluska-Nawarecka et al., 2013).
− Base of dilatograms together with an application that allows the digitization of graphs and searching.
− RoughCast – attribute casting defects classification system implemented according to the theory of rough sets (Dobrowolski et al., 2011).
− OntoGrator - system of knowledge integration based on the domain ontology (Regulski & Kluska-Nawarecka, 2012).
− The semantic integration system for unstructured text documents executed by algorithms of latent semantic indexing (Gurda et al., 2013).

Each of these modules has a different functionality, characteristics, architecture and method of operation. User interaction with these modules has different requirements - simple queries are addressed to relational databases, some modules perform operations instead of returning the results to the screen, etc. The user may use the individual modules in different situations - some of them will be used in the search for documents containing information, some will be used to find specific material or process parameters, others will serve visualization or monitoring of the Internet or rule-based inference. Development of an intelligent interface is meant to improve navigation in the complex structure of the system functionality.

3. INTELLIGENT INTERFACE

The concept of an intelligent interface is not a new concept. IUI - Intelligent User Interface - is a group of issues well-known in computer science for years. Interest in such systems is due to the still developed database systems whose interfaces are inflexible and ill-suited to user different requirements. Many also are the reasons why the computer application interfaces can be called intelligent:
− interfaces that have knowledge about the functionality of the application,
− interfaces that have knowledge of user preferences,
− self-adaptive interfaces, i.e. the interfaces that on the basis of interaction with the user can adapt themselves to his needs,
− interfaces processing natural language and using semantic analysis.

Creating smart interfaces goes hand in hand with the development of artificial intelligence techniques and uses its achievements. Computer systems are
becoming increasingly better adapted to the needs of users, often having more intelligent interface than an application that lies beneath it. At a pace relatively most slow is developing the first of the above mentioned groups of intelligent interfaces, which might be due to the fact that it emerged as the first one and therefore could be treated as a problem already solved. Everyone has come across the infamous MS Office Assistant, which further testifies to the fact that in this field there is still much to be done.

3.1. CBR methodology

The proposed solution for the described interface is to use case-based reasoning to assist the use of the system. User after each use of the system leaves his mark in the form of a given query, selection of individual modules of the system or indications what type of knowledge is needed. Any such interaction with the system is saved as one of the cases. Using the accumulated knowledge about how to interact with the system (cases base), the interface may next time suggest to the user the keywords, potentially the best modules for his problem, or the type of data that are worth to search on a basis of revised former cases that fits the best to the user's query.

The main paradigm of CBR methodology is inference regarding the current problem by using knowledge of the similar problems solved in the past (Aamodt & Plaza, 1994). The main component of the CBR system is a case-base, which contains a collection of problems and solutions (cases) that occurred in the past. When a new problem emerges, CBR system searches the database for similar cases, and then the solution of the case searched is adapted to the current situation. CBR methodology also allows for learning - at any time instant, when the solution of the new case takes place, the knowledge related to this experience is attached to the case-base. Thanks to this, relevant knowledge will be available in subsequent uses of the CBR system to solve the following problems. The CBR methodology can be compared to the process of “scooping” and using experience.

Despite many differences in the construction of various systems using CBR methodology, a common element is the algorithm called CBR cycle. CBR cycle consists of four phases (Rojek et al., 2014):

1. searches for the most similar case or cases,
2. the re-use of information and knowledge contained in the case most similar to solution of the current problem,
3. revision of the proposed solution,
4. preservation of information and knowledge regarding current case together with its solution for use in the future.

CBR cycle starts when there is a new problem to solve - a new user or a new query. Then the four phases shown above are executed sequentially. In the first phase, there is search for the case or group of cases which are most similar to the current problem. Each case stored in the database corresponds to one interaction with the system. The case consists of the keywords specified by the user, possibly indicating modules, which he wants to use and the type (types) of the information sought. Search among the collected cases will be done using a similarity measure that best suits the attributes compared. For example, for the vector of keywords, a cosine distance relative to the vectors of keywords from the previous interactions can be used. In many known applications of CBR methodology, a similarity measure uses a Euclidean or Hamming distance (Kusiak & Rojek, 2012).

In the phase of re-use, the case searched in the previous phase is used to provide the user with hints on the use of the system. This phase can be very simple when the solution related to the searched case is returned as the solution of the current case. For example, it is corresponding to the query result that was obtained in the past results. In many cases, however, an adaptation of the solution is necessary, which may occur by transforming the previous query or using other modules associated with the new keywords.

During revision, the proposed solution is evaluated, and in the case of a negative evaluation, introduction of corrections to the solution is also possible. Evaluation takes place by matching result to the user's needs and is the subjective opinion of the user. The use of an expert is also anticipated. He could serve as a "model user" and provide information on the assessment of individual results, or the necessary amendments. This first phase of the implementation of the system anticipates creating the preliminary case-base consisted of cases created by this "model user". The phase of platform testing provides dozens of examples of the use cases of the system. That is why the case-base is never empty, even at "first use".

The phase of storing information and knowledge is the phase in which learning of the system occurs. This process typically involves a simple addition of the current event (after the revision phase) to the
case-base. With this, the knowledge of the problem, which has just been solved, can be used in the future functioning of the system (figure 2).

Information flow and saving "past cases" to a base forming the above illustrative embodiment shown in figure 3. This diagram clarifies the general concept of the intelligent interface shown in figure 2.

For example, user can ask a question expressed in natural language: "How is ductile iron obtained?". Agent of the interface by analyzing the sentence chooses keywords: "obtaining" + "ductile iron". This query devoid of other attributes, i.e. selected modules and the type of knowledge, is directed to the advisory system. Individual modules return a variety of results, e.g. CastSemaWiki will return an article on ductile iron, in which the user will find an answer to the question posed; the database of material characteristics will indicate the parameters of ductile iron, but in order to get information on how to obtain this material, the user will have to find in a domain ontology a suitable SWRL rule indicating that the grey cast iron subjected to modifications and appropriate heat treatment changes into ductile iron. One can also search the base of publications with the specified keywords, or check whether in the field of ductile iron in recent years some research works have been performed.

The interface stores the actions performed by the user thereby complementing the case of interaction with the type of information and system modules.
used. When re-call of this query occurs, the system may propose further proven steps of search.

4. SUMMARY

This paper describes a proposal to apply the techniques of inference drawn from the studies of artificial intelligence to support user interface design in large systems of decision support in the field of metallurgical technology and metal processing. Complex systems can be a valuable tool for the user who is an employee of industry, but the level of complexity in their use can be a psychological barrier to their widespread application. Designing of intelligent interfaces that are guides during operation of the system may decide about the success of the operation of information science tools. The proposed solution uses the CBR methodology in the process of building a database of cases of user interaction with the system, in order to give further support and advice in the course of working with the advisory system. User can easily take advantage of the proposed functionalities when his demand for information is not fully specified, or has problems with proper expression of the query. In the literature on information science, the concept of intelligent interface is well-known. Innovative is the use of methodologies in the field of episodic inference to control contact with the advisory system in the field of metal processing. The advantage of the use of a CBR methodology comparing to the machine learning methods (classification methods, Bayesian Networks, etc.) commonly used in IUI ist that the case-base need not necessarily be very complex, especially at the beginning - each "use-case" is represented by one case in base, not like in others methods - by group of cases.

ACKNOWLEDGEMENTS

The work was financed within the framework of the international project No. 820/N-Czechy/2010/0 of 30 November (part: modules of knowledge and databases). Financial support of the National Centre for Research and Development (LIDER/028/593/L-4/12/NCBR/2013) is also gratefully acknowledged (in the part of intelligent interface engine model).

REFERENCES


INTELLIGENTNY INTERFEJS DLA SYSTEMU WSPOMAGANIA DECYZJI W OBSZARZE METALURGII

Streszczenie

Rozległy system wspomagania decyzji złożony z kilku modułów zapewniających różne funkcjonalności użytkownikowi stanowi wyzwanie w projektowaniu użytkowniczego, prostego i spójnego interfejsu. Specyfikę systemów doradczych w zakresie metalurgii jest stopień skomplikowania ich architektury zaprojektowa-
nej celem wspierania różnorodnych typów zadań w wielu obszarach. Celem tego artykułu jest prezentacja możliwości stworzenia inteligentnego interfejsu zapewniającego użytkownikowi pełną swobodę działania nawet w sytuacji, gdy nie zna on dokładnie wszystkich funkcjonalności systemu. Artykuł przedstawia architekturę tworzonego systemu wspomagania decyzji w zakresie metalurgii wzbogaconą o opis mechanizmu pozwalającego skonstruować inteligentny interfejs. Interfejs taki powinien wspomagać użytkownika w nawiązaniu po modułach systemu tak, aby użycie systemu było najefektywniejsze. Do tworzenia takiego mechanizmu zaproponowano metodykę wnioskowania epizodycznego (Case-Based Reasoning, CBR). CBR stanowi wzorzec tworzenia systemów uczących się na podstawie przeszłych zachowań użytkownika związanych z wykorzystaniem systemu w rozwiązywaniu problemów. Wiedza o poprzednich zachowaniach i wyborach użytkownika służy do wspomagania bieżących problemów. Każdze zapytanie użytkownika skierowane do systemu oraz każdy jego wybór użycia poszczególnych funkcjonalności modułów systemu wzbogacza system o cenną wiedzę: właściwe wykorzystanie całego systemu doradczego związane z potrzebami użytkownika. Umiejętnie wykorzystanie tej wiedzy zapewni wygodną i efektywną pracę inteligentnego interfejsu.

Received: October 1, 2014
Received in a revised form: November 21, 2014
Accepted: November 24, 2014
CONTROL OF LEAD REFINING PROCESS WITH THE USE OF CASE-BASED REASONING APPROACH

GABRIEL ROJEK*, KRZYSZTOF REGULSKI, PIOTR JAROSZ, JACEK GABRYEL, JAN KUSIAK

AGH University of Science and Technology, al. A. Mickiewicza 30, 30-059 Krakow, Poland
*Corresponding author: rojek@agh.edu.pl

Abstract

Presented in this work research concerns pyrometallurgical method of lead refining process. The main objective of the considered lead refining process is to remove impurities of the input lead ore in order to obtain output material of PB985 type. The content of lead in the output material should be at least 99,985%. The considered pyrometallurgical lead refining process consists of 11 subprocesses, which are performed sequentially. The output material of each of the first 10 subprocesses is the input material for the consecutive subprocess. Each of subprocesses is controlled autonomously, however it can influence the control of following subprocesses due to obtained results. The main goal of the work is to design a computer system, which controls the whole lead refining pyrometallurgical process. The essential idea of the control system is to elaborate control advices at every step of the refining process. The proposed computer system should consist of subsystems giving advices according to individual subprocesses, so every subprocess should be controlled by individual control subsystem. Case-Base Reasoning is chosen as the methodology for creation of each control subsystem.

Key words: lead refining process, process control, Case-Based Reasoning

1. INTRODUCTION

The input material for lead refining process is lead containing impurities, which have to be removed in order to obtain output material of PB985 type – the content of lead in output material should be at least 99,985%. Process of lead refining can be realized with the use of two different methods: either pyrometallurgical or electrolytic refining is used (Habashi, 1997). Presented in this work research concerns the pyrometallurgical method, because this method is used in known industrial plant and real industrial data concerning this method is available to be used in presented next research according the design of control system. The whole pyrometallurgical process of lead refining consists of 11 subprocesses, which are performed sequentially in an industrial plant:

1. Arranging input materials into batches.

2. Drossing (Liquation) – the first stage of decoppering; cooper and other impurities precipitate, when the lead cools.

3. Sulfur decoppering – the second stage of decoppering; selective sulfidizing of cooper by the Colcord process.

4. Initial refining – removal of arsenic, tin and antimony by selective oxidation and the Harris process.

5. Desilvering – removal of noble metals, which involves intermetallic precipitation with zinc by the Parkes process.

6. Dezincing – removal of zinc, which remains after desilvering; this subprocess is achieved by vacuum distillation.

7. Debismuthizing – the removal of bismuth can be achieved by the Kroll-Betterton process, which uses calcium and magnesium.
8. Final refining I – the first stage of removal of alkali metals and alkaline earth metals remaining after debismuthizing.

9. Removing of thallium (with the use of zinc chloride).

10. Final refining II – the second stage of removal of alkali and alkaline earth metals.

11. Casting – casting temperature is around 400 °C.

The output material of each of first 10 subprocesses is the input material for next subprocess. Each of subprocesses is controlled autonomous, however it influences the control of next subprocesses due to obtained results. Such run of the whole refining process results in the main idea of the computer system that should aid at control of the whole process. The designed computer system should consists of many subsystems giving advices according individual subprocesses, so the purpose of the whole system is to give advices at every one subprocess performed in the industrial plant. Such advices can be a help for the operator of the process in order to find out proper control of the industrial process. This feature can be very important for the operator, that has a little experience according control of the analyzed process. In such case the system will give advices related to experiences of other operators working in the enterprise and will help the inexperienced operator to undertake decisions.

The scientific domain of this paper is the design of an advisory computer system, which should give help for the operator of an industrial process. Advices giving by the designed system are based on experiences gained by the control of the industrial process done in the past, because it is not possible to model the analyzed process in order to simulate physicochemical processes. All experiences being the basis for system functioning concern only results, which in the goal of the whole industrial process was obtained – the output material was of PB985 type, but the specific content of lead in the output material is not know. Such characteristics of the data being the basis for knowledge of the system prevents formulating any objective according optimization of the industrial process performed by the system. The only one functionality providing by the designed system can be giving advices, which

---

**Fig. 1.** Consecutive subprocesses in the whole process of lead refining.
should help the operator to obtain desirable results, but the optimization of the industrial process is not possible due to inadequate data that are held by the authors.

2. CONTROL OF LEAD REFINING PROCESS

As presented in the previous section, the whole lead refining process consists of 11 subprocesses, which are controlled in an autonomous way. The consecutive subprocesses are referenced as $S_1, S_2 \ldots S_{11}$. The schematic view of the whole process control is presented in figure 1.

As presented in figure 1, the input material for a subprocess $S_i (1 \leq i \leq 11)$ is denoted as $M_i (1 \leq i \leq 11)$. The control of a subprocess $S_i (1 \leq i \leq 11)$ is denoted as $C_i (1 \leq i \leq 11)$. The output of a subprocess $S_i (1 \leq i \leq 11)$ is divided into $R_i (1 \leq i \leq 11)$ and $M_{i+1} (1 \leq i \leq 11)$, where $R_i$ means rejectamenta after performing of $S_i$ (materials, which are not used in next subprocesses of the whole analyzed process) and $M_{i+1}$ means materials, which are used in the next subprocess $S_{i+1}$ (in other words, these materials are input materials for a subprocess $S_{i+1}$). $M_{12}$ denotes output materials of the last subprocess (casting) and is output of the whole lead refining process realized with considered pyrometallurgical method.

The control $C_i (1 \leq i \leq 11)$ of a subprocess $S_i$ relates to temperatures (temperature at the process start and temperature at the process end), duration time and amount of technological additives, which are added (e.g. S, PbS, NaOH, NaNO₃, ZnCl₂, sawdust). As it is earlier mentioned, the control $C_i (1 \leq i \leq 11)$ is autonomously performed in the subprocess $S_i$. The control $C_i$ should be adjusted to the input materials $M_i$ of the subprocess $S_i$ in order to obtain the output materials $M_{i+1}$, which are desired materials for the next subprocess $S_{i+1}$. The rejectamenta $R_i$ relates to amount of rejected materials after performing of the subprocess $S_i$ (e.g. amount of dross after the $P_2$ subprocess).

3. DESIGN OF THE CONTROL SYSTEM

The control of the whole lead refining process is realized by autonomous controls of every subprocess $S_1, S_2 \ldots S_{11}$, what is the basis for the main idea of the control system design as many control systems giving advices at each subprocess. The whole control system should consists of 11 units related to 11 subprocesses, the control unit $CU_i$ is related to the subprocess $S_i (1 \leq i \leq 11)$, what is presented in figure 2.

The main functionality of the control unit $CU_i (1 \leq i \leq 11)$ is giving advices at control of the subprocess $S_i (1 \leq i \leq 11)$. The control unit $CU_i$ should indicate proper control $C_i$ of the subprocess $S_i$ according to the current specification of used input material $M_i$. Because any rules according proper control realized by control units are not possible to formulate, every control unit should use available industrial data according past made actions. The available industrial data concerns several runs of the whole lead refining process and can be source of knowledge according proper control advices giving by all control units related to consecutive subprocesses.

Analyzing the possibilities of design of the control unit $CU_i (1 \leq i \leq 11)$ related to the subprocess $S_i (1 \leq i \leq 11)$ the basic remark is, that analytical model of any subprocess $S_i (1 \leq i \leq 11)$ is not possible to formulate due to complexity and not completely known nature of the physico-chemical processes. This remark is the reason for searching of solutions concerning design of the control units in the domain of Artificial Intelligence (AI). One of wide used approaches related to AI in the domain of control of industrial processes consists of building the model of the industrial process with the use of Artificial Neural Network (ANN) and using this model by an optimization procedure. In Sztangret et al. (2009) ANN model is based on the architecture of Multi-Layer Perceptron (MLP) and particle swarm optimization (PSO) is used as the optimization procedure. This approach requires enough amount of data in order to

![Diagram of the control system](image-url)
train the ANN, which should predict outcomes of a hypothetical control. Presented further remarks on system functioning indicate, that the available amount of data (for presented here research) seems to be not sufficient in order to use the approach, which uses ANN as the model of the analyzed process. Another approach related to AI, which is applied to control of industrial process, relies on Case-Based Reasoning (CBR). In Kusiak et al. (2013) the CBR system is presented, which is applied to the oxidizing roasting process of sulphide zinc concentrates. In Sun (2008) the CBR methodology is used to control of combustion control of blast furnace stoves. Another application of CBR in the domain of industrial production is presented in Michael and Khemani (2002). The CBR methodology enables to construct an advisory system even the amount of the available data is very scant, so this methodology is chosen for creation of each control unit \( CU_i \) (\( 1 \leq i \leq 11 \)).

### 3.1. General view of Case-Based Reasoning

The CBR methodology involves reuse of previously collected data by solving of a current problem. After the current problem is solved, the data related to this problem and its solution is saved in order to be used at next system functioning, what enables learning on the basis of performed by system actions (Bergmann et al., 2009). In Rojek and Kusiak (2012) application of Case-Based Reasoning to a generic industrial process is presented, however it is assumed, that a generic process does not consist of subprocesses. The main notion of CBR is a case, which represents a past event (referenced also as an episode or an experience item). A case is represented usually as an ordered pair (\( \text{problem}, \text{solution} \)), where \text{problem} describes a specific situation and \text{solution} describes actions, parameters and any other data related to the performed solution to this specific problem. One case represents one event, that took place in the past. Because a CBR system uses data according many past events, notion of a case base is used. The case base is collection of all cases representing all registered past events.

A system developed with Case-Based Reasoning performs the CBR cycle, that consists of four phases. The CBR cycle start, when a new problem (called the current problem) has to be solved Aamodt and Plaza (1994):

1. **Retrieve phase** – select in the case base a case, which is similar to the current problem; the selected case is referenced next as the retrieved case.
2. **Reuse phase** – use the information and knowledge in the retrieved case in order to solve the current problem; the solution in the retrieved case is returned as the solution to the current problem (with possible adaptation).
3. **Revise phase** – the proposed in the previous step solution is revised by e.g. being applied to the real environment of the CBR system.
4. **Retain phase** – the information concerning the current problem and its revised solution is saved in order to be useful for future problem solving; this step usually happens by adding a new case to the case base.

### 3.2. Definition of a case of a control unit

Because a control unit \( CU_i \) should give advices at control of the subprocess \( S_i \) (\( 1 \leq i \leq 11 \)), a case is denoted as \( \text{case}_i \) equal to a pair (\( \text{problem}_i, \text{solution}_i \)) concerning the subprocess \( S_i \). It means, the control unit \( CU_i \) has its own case base, which consists of cases related to control of the subprocess \( S_i \). The problem in the domain of subprocess control relates to particular input materials. The solution relates to done control and its results, so finally it is proposed to define a case as a pair \(( M_i, (C_i, M_{i+1}, R_i))\), where:

- \( M_i \) denotes problem in the form of the specific materials, used as the input materials for the subprocess \( S_i \),
- \( C_i \) denotes control of the subprocess \( S_i \),
- \( M_{i+1} \) denotes the output materials of the subprocess \( S_i \) after performing of the control \( C_i \) and using input materials \( M_i \),
- \( R_i \) denotes the rejectamenta of the subprocess \( S_i \) after performing of the control \( C_i \) and using input materials \( M_i \) (the rejectamenta are materials not used in next subprocess of the whole industrial process).

### 3.3. The CBR cycle of a control unit

As it was mentioned before, a control unit \( CU_i \) gives advices at control of the subprocess \( S_i \). The control unit \( CU_i \) has its own case base, which contains cases according past runs of the subprocess \( S_i \). When a need is to give an advice of control of the subprocess \( S_i \), the control unit \( CU_i \) creates the representation of the current problem (by contents of individual elements: Ag, Cu, Zn etc.) and performs the CBR cycle:
1. Retrieve phase: selects in the case base a case \((M_i^R, (C_i^R, M_{i+1}^R, R_i^R))\), which is most similar to the current problem related to the specified input materials \(M_i^C\).

2. Reuse phase: presents the control \(C_i^R\), the specification of the output material \(M_{i+1}^R\) and the rejectamenta \(R_i^R\) as the solution to the current problem.

3. Revise phase: enables the user to change the data presented in the previous step according to the real control and its results made in the industrial environment – the control \(C_i\), the specification of the output material \(M_i^C\) and the rejectamenta \(R_i\) can be inputted by the user.

4. Retain phase: creates a new past case \((M_i^C, (C_i^C, M_{i+1}^C, R_i^C))\), which relates to the current problem \(M_i^C\) with its solution \((C_i^C, M_{i+1}^C, R_i^C)\) and adds this case to the case base.

The goal of the retrieve phase is to select a case \(case^i = (M_i^R, (C_i^R, M_{i+1}^R, R_i^R))\), which is most similar to the current problem related to the specified input materials \(M_i^C\). This involves use of a similarity measure, which is usually formalized as a function \(sim: P \times P \rightarrow [0, 1]\), which compares descriptions of two problems from \(P\) and produces a similarity assessment as a real value from \([0, 1]\). Taking into consideration the analyzed industrial process and the subprocess \(S_n\), the similarity function is proposed to be equal to the inverse Euclidean distance:

\[
sim(case^i, M^C) = \frac{1}{1 + \sqrt{\sum (m_i^x - m_i^C)^2}}
\]

where: \(case^i = (M_i^R, (C_i^R, M_{i+1}^R, R_i^R))\), \(M_i^R = [m_i^{x,1}, m_i^{x,2}, \ldots, m_i^{x,N}]\) and \(M_i^C = [m_i^{C,1}, m_i^{C,2}, \ldots, m_i^{C,N}]\).

The reuse phase is very simple in presented here application of CBR methodology – the system presents the control \(C_i^R\), the specification of the output material \(M_{i+1}^R\) and the rejectamenta \(R_i^R\), which are directly copied from the retrieved case \(case^i\). Adaptation is not made in the reuse phase due to the lack of knowledge related to physicochemical nature of the analyzed process. The available industrial data is also to scant for the use of any statistical methods at adaptation of the retrieved case \(case^i\) to the current problem related to the specified input materials \(M_i^C\).

The revise phase requires implementation of the system to the real environment of its functioning. The system enables only to input the real control and its results made in the industrial environment – the control \(C_i^R\), the specification of the output material \(M_{i+1}^C\) and the rejectamenta \(R_i^C\) have to be inputted by the user. The insertion of this data is the basis for the last – the retain phase, which enables to enlarge the case base with currently made experience. In the retain phase a case \((M_i^C, (C_i^C, M_{i+1}^C, R_i^C))\) is created and added to the case base. During the next system functioning this case is treated equally with other past cases – is processed in the retrieved phase and can be chosen as the retrieve case being the basis for future problem solving.

### 4. REMARKS ON SYSTEM IMPLEMENTATION AND FUNCTIONING

The control system is implemented according to previously presented remarks. By implementation jCOLIBRI framework is used as the code library for classes related to operations performed during the CBR cycle. Functioning of the system is presented on the base of the \(P_7\) subprocess (debismuthizing). The available industrial data enables to create 9 cases according the debismuthizing done in the past. An example of a case can be given as:

1. \((\text{case } 5)\)
2. \((\text{Ag:4, Cu:5, Zn:2, Bi:553, As:3, Sn:3, Sb:33, Cd:2, Ni:1, Ti:360, Fe:12, S:7, Al:1, Te:18})\)
3. \((\text{Ag:4, Cu:5, Zn:2, Bi:553, As:3, Sn:3, Sb:33, Cd:2, Ni:1, Ti:360, Fe:12, S:7, Al:1, Te:18})\)
4. \((\text{Bi_foam: 4000})\)
5. \((\text{Bi_foam: 4000})\)
6. \((\text{Bi_foam: 4000})\)
7. \((\text{Bi_foam: 4000})\)

The second line in the above given example specifies the problem \(M^5\), the fourth line specifies the control \(C^5\), the fifth line specifies the output material \(M^8\) and the sixth line specifies the rejectamenta \(R^5\). The main influence on the system functioning has the similarity measure, which is presented by equation (1) and which operates on descriptions of problems. The results of system functioning depend on the similarity measure and input data, which characteristic according description of problem is presented in table 1.

At the start of the functioning of the subsystem \(CU7\) giving advices for the debismuthizing subprocess the user inputs description of the current problem e.g.: \((\text{Ag:4, Cu:4, Zn:1, Bi:545, As:3, Sn:3, Sb:34, Cd:1, Ni:1, Ti:360, Fe:10, S:8, Al:1, Te:18})\). The subsystem moves to the retrieve phase and computes the similarity measure for all cases: 0.016,
0.017, 0.012, 0.014, 0.105, 0.008, 0.019, 0.011, 0.025. The most similar is case 5, so case 5 (presented in the example at the beginning of this section) is selected as the retrieved case and the subsystem moves to the reuse phase – presents the control, the specification of the output material and the reject, which are directly copied from the retrieved case 5. Next subsystem moves to the revise phase, which in the user can change presented data according to performed control in the real environment. Data revised by the user are saved in the case base as a new past case in the last retain phase.

5. CONCLUSION

The problem of building the computer advisory system, which helps at control of the industrial lead refining process can be resolved using Case-Based Reasoning (CBR). Presented in this work design of such computer system is adjusted to the specific of the analyzed process that consists of 11 subprocesses, which are performed sequentially in the industrial plant. The efficiency of presented solution due to specific of CBR cannot be evaluated without full deployment of the system to the real environment, which is the industrial plant in the case of the analyzed process. Influence on the system efficiency which is the industrial plant in the case of the analyzed process has amount of data related to operations done in the past, because this data collected in the case base of CBR system are source for advices giving by the system. Future evaluation of the system should be done in long period of time – half a year or even longer in order to make possible use of data that is add to the case base by functioning of the system.

ACKNOWLEDGEMENTS

Financial support of the Ministry of Science and Higher Education (AGH UST, project no. 11.11.110.300) is acknowledged.

REFERENCES


Bergmann, R., Althoff, K. D., Minor, M., Reichle, M., Bach, K., 2009, Case-Based Reasoning, Introduction and Recent Developments, *Künstliche Intelligenz: Special Issue on Case-Based Reasoning*, 23, 5-11.


Table 1. Characteristics of data related to description of problem.

<table>
<thead>
<tr>
<th>parameter</th>
<th>min</th>
<th>max</th>
<th>average</th>
<th>median</th>
<th>variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ag</td>
<td>1</td>
<td>4</td>
<td>2.22</td>
<td>2</td>
<td>1.44</td>
</tr>
<tr>
<td>Cu</td>
<td>1</td>
<td>7</td>
<td>3.33</td>
<td>2</td>
<td>4.50</td>
</tr>
<tr>
<td>Zn</td>
<td>1</td>
<td>9</td>
<td>2.67</td>
<td>1</td>
<td>9.50</td>
</tr>
<tr>
<td>Bi</td>
<td>480</td>
<td>668</td>
<td>584.78</td>
<td>591</td>
<td>3150.69</td>
</tr>
<tr>
<td>As</td>
<td>3</td>
<td>8</td>
<td>3.78</td>
<td>3</td>
<td>2.69</td>
</tr>
<tr>
<td>Sn</td>
<td>3</td>
<td>5</td>
<td>3.33</td>
<td>3</td>
<td>0.50</td>
</tr>
<tr>
<td>Sb</td>
<td>1</td>
<td>33</td>
<td>13.56</td>
<td>13</td>
<td>91.78</td>
</tr>
<tr>
<td>Cd</td>
<td>1</td>
<td>2</td>
<td>1.44</td>
<td>1</td>
<td>0.28</td>
</tr>
<tr>
<td>Ni</td>
<td>1</td>
<td>2</td>
<td>1.22</td>
<td>1</td>
<td>0.19</td>
</tr>
<tr>
<td>Tl</td>
<td>340</td>
<td>380</td>
<td>360.00</td>
<td>360</td>
<td>100.00</td>
</tr>
<tr>
<td>Fe</td>
<td>12</td>
<td>27</td>
<td>18.44</td>
<td>18</td>
<td>29.78</td>
</tr>
<tr>
<td>S</td>
<td>6</td>
<td>15</td>
<td>8.22</td>
<td>7</td>
<td>10.19</td>
</tr>
<tr>
<td>Al</td>
<td>1</td>
<td>2</td>
<td>1.22</td>
<td>1</td>
<td>0.19</td>
</tr>
<tr>
<td>Te</td>
<td>6</td>
<td>36</td>
<td>20.56</td>
<td>21</td>
<td>72.28</td>
</tr>
</tbody>
</table>

STEROWANIE PROCESEM RAFINACJI OLOWIU Z WYKORZYSTANIEM PODEJŚCIA OPARTEGO O WNIOSKOWANIE EPIZODYCZNE

Streszczenie

Przedstawione w pracy badania dotyczą pirometalurgicznej metody rafinacji ołowiu. Głównym celem procesu rafinacji ołowiu jest usunięcie zanieczyszczeń z materiału wejściowego w celu otrzymania materiału wyjściowego typu PB985, w którym zawartość ołowiu wynosi co najmniej 99,985%. Rozważano pirometalurgiczny proces rafinacji ołowiu złożony z 11 podprocesów, które są wykonywane w sposób sekwencyjny. Materiał wyjściowy każdego z pierwszych dziesięciu podprocesów stanowi materiał wejściowy do następującego po nim podprocesu. Każdy podproces jest sterowany autonomicznie, jednakże może wpływać na następujący po nim podproces poprzez osiągnięte rezultaty. Głównym celem niniejszej pracy jest projekt systemu komputerowego, który pozwala na sterowanie całym procesem rafinacji. System ten powinien prezentować wskazówki odnośnie sterowania na każdym z etapów rafinacji. Proponowane rozwiązanie jest w postaci systemu komputerowego, który składa się wielu pod-
systemów sterowania przyporządkowanych do odpowiednich podprocesów rafinacji, więc każdy podproces jest sterowany poprzez indywidualny podsystem. Przy projektowaniu każdego z podsystemów sterowania użyto podejścia opartego o wnioskowanie epizodyczne (ang. Case-Based Reasoning).

---

Received: October 14, 2014
Received in a revised form: November 26, 2014
Accepted: December 5, 2014
EXTRACTING KNOWLEDGE FROM INTEGRATED EXPERIMENTAL DATA ON THE ADI MANUFACTURE

BARBARA MRZYGLÓD*, IZABELA OLEJARCZYK-WOŻEŃSKA, MIROSŁAW GŁOWACKI, ANDRZEJ OPALIŃSKI

AGH University of Science and Technology, Al. Mickiewicza 30, 30-059 Krakow, Poland

*Corresponding author: mrzyglod@agh.edu.pl

Abstract

The problem discussed in this article relates to the integration of knowledge about the design process of products from ADI. Specialised domain knowledge, often resulting from costly experiments, published in the world in a large number of magazines, is an invaluable source of information for its users and researchers. The authors draw attention to the need to develop a system that will facilitate selective access to specific passages of this knowledge, allowing for its automatic processing. The cost of developing such a system is certainly lower than the cost of multiple repetition of the experiments. Proposed under the system, aggregation and centralisation of the results of studies already carried out may be the basis for planning and execution of subsequent experiments, covering areas hitherto unexplored. Another outcome may relate to the creation of new knowledge through the discovery of relationships and dependencies that are not visible in individual, single, experiments, but emerge when the results of a large number of different tests are compared.

This paper proposes the use of artificial neural networks to explore the relations between the properties of ADI and selected heat treatment parameters based on a set of integrated experimental data from various publications. In the future, this form of knowledge representation may be used in intelligent computer systems in the knowledge acquisition module on the manufacturing process of ADI.

Key words: data extraction, manufacture of ADI, table of attributes, decision support systems, artificial neural networks

1. INTRODUCTION

Specialised domain knowledge, often resulting from costly experiments, published in the world in a large number of magazines, is an invaluable source of information for both technologists and researchers. However, getting familiar with all the articles in a particular subject (the production of ADI) and in a particular field of study (metallurgy-founding) is often a great problem arising, among others, from the restricted access to certain databases, language barriers, time constraints, etc.

The creation of a system facilitating selective access to specific passages of domain knowledge and allowing for an automatic processing of this knowledge seems to be the goal deliberate and advantageous. The cost invested in the development of such a system will certainly be lower than the cost of multiple repeating of experiments that have already been carried out. Moreover, this combination of results (centralisation) can be a source of inspiration for the planning and execution of research covering the areas unexplored so far.

With integrated expertise in a particular field, one can attempt to design inference algorithms and systems that will enable automatic processing of the knowledge (Kochański et al., 2010).

2. CHARACTERISTICS OF THE SOURCE MATERIAL

The sphere of the authors’ interest includes published knowledge on the manufacturing process of ADI (Austempered Ductile Iron). An overview of
journals and an analysis of articles relating to this area aimed at the selection of the ADI production parameters ensuring the achievement of certain properties in the final material. Most experimental studies were designed to investigate the effect of selected agent on the properties of the material. For example, on samples of a given chemical composition, several variants of the heat treatment were performed, or for a predetermined variant of the heat treatment, the chemical composition was changed, introducing, e.g., a new element and examining its effect on the structure of the material and hence on the properties (Zahiri et al., 2003; Dymski, 2001; Putatunda, 2001a, 2001b; Lin et al., 1996; Janowak & Morton, 1984; Biel-Golaska & Kowalski, 1996; Kowalski et al., 1990).

Logical diagram, observed in most of the reviewed articles, used for the presentation of the results of experiments can be expressed in the form of an implication where the premises include the chemical composition of the base material, the dimensions of the heat treated product, and the heat treatment parameters, while the resulting variables are the values of selected properties of the examined product. The logical diagram is presented in figure 1. From the examined articles, using the adopted scheme, the experimental data were extracted. An attribute table was developed containing the specific values of the variables, which are used as the premises. These are the following parameters:

- **chemical composition**: C, Si, Mn, Mg, Cu, Ni, Mo, S, P, B, V, Cr, Ti, Sn, Nb, Al;
- **dimensions** (usually these are dimensions of samples according to standards adopted in the experiment);
- **heat treatment parameters**: temperature of austenitising ($T_a$), time of austenitising ($t_a$), temperature of isothermal transformation ($T_i$), time of the transformation ($t_i$), and the cooling medium.
The variables which are conclusions include:
- tensile strength \((R_m)\);
- elongation \((A_5)\);
- reduction of area \((Z)\);
- hardness \((HRC)\);
- toughness \((K_C)\);
- yield strength \((R_e)\).

A model of the table was created taking into account all the variables that have occurred in the described experiments. Since only some of the experiments allow for all the variables, the table also includes areas for which no data are available. Table 1 presents a fragment of the developed table.

The table summarises the results of 260 variants of experiments carried out in the production of ADI. ADI is formed by the properly conducted heat treatment (HT) of ductile iron, which is a starting material for the production of ADI. The heat treatment consists of two stages, i.e. austenitising and austempering. The heat treatment parameters, i.e. the parameters of both austenitising \((T_a - \text{temperature of austenitising})\) and austempering \((T_i - \text{temperature of isothermal holding})\) and \(t_a - \text{time of austenitising}\) and \(t_i - \text{time of isothermal holding}\) are selected depending on the properties that are to be obtained in ADI, e.g. high ductility combined with moderate strength and hardness, or vice versa a low ductility at the expense of high hardness and strength. The combination of properties that are required must be optimal from the point of view of later conditions of the casting operation.

Table 2 presents the classification of ADI by American ASTM A897 Standard. For ADI the classification is based on properties and not on the chemical composition. Below are the properties obtained by heat treatment that decide about the grade of the material.

### Table 2. ADI classification according to ASTM A897.

<table>
<thead>
<tr>
<th>Grade by ASTM A897</th>
<th>Tensile strength (R_m) [N/mm²]</th>
<th>Yield strength (R_{p,0.2}) [N/mm²]</th>
<th>Elongation (A_5) [%]</th>
<th>Hardness (H_B)</th>
<th>Fracture toughness (K_C) (unnotched) [J]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASTM 897 Grade 1</td>
<td>850</td>
<td>550</td>
<td>10</td>
<td>269-321</td>
<td>100</td>
</tr>
<tr>
<td>ASTM 897 Grade 2</td>
<td>1050</td>
<td>700</td>
<td>7</td>
<td>302</td>
<td>363</td>
</tr>
<tr>
<td>ASTM 897 Grade 3</td>
<td>1200</td>
<td>850</td>
<td>4</td>
<td>341-444</td>
<td>60</td>
</tr>
<tr>
<td>ASTM 897 Grade 4</td>
<td>1400</td>
<td>1100</td>
<td>1</td>
<td>388-487</td>
<td>35</td>
</tr>
<tr>
<td>ASTM 897 Grade 5</td>
<td>1600</td>
<td>1300</td>
<td>-</td>
<td>444-555</td>
<td>-</td>
</tr>
</tbody>
</table>
4. THE USE OF ANN FOR MODELLING THE MANUFACTURING PROCESS OF ADI

Artificial neural networks are mathematical structures, the operation of which is a simplified representation of the human brain. Artificial neural network consists of a group of interconnected cells (neurons) processing in parallel the received information. Appropriately designed neural network is capable of self-formulating the interdependences between the parameters of a phenomenon during the learning process based on empirical cases. The learning process is an iterative process, repeated many times, step by step, with the primary objective to optimise the network parameters, i.e. the weighting factors. Each of the variables entering the network input gets initially a randomly assigned weight, which is the strength of its impact on the value of the output variable. The weighting factors are determined in a learning process, which is the source of knowledge and the intelligence of a neuron. The higher is the weight value, the more significant is the variable (Lula et al., 2007). The main way to use a neural network is by the creation of models, which are a formalised structure mapping a process or phenomenon. One of the most important issues is an appropriate set of empirical data describing the examined phenomenon.

The collected experimental data on ADI production parameters may give rise to the development of a model of this process. A very important step in the development of intelligent systems (e.g. supporting the decision making process) is to clarify the purpose of the analysis and define its basic assumptions. The following describes various stages of the work carried out during the construction of a formal model of the phenomenon of the ADI production using artificial neural networks, based on the collected experimental data.

Table 3. The adopted filtering criteria.

<table>
<thead>
<tr>
<th>C</th>
<th>Si</th>
<th>Mn</th>
<th>Mg</th>
<th>Cu</th>
<th>Ni</th>
<th>Mo</th>
<th>S</th>
<th>P</th>
<th>B</th>
<th>V</th>
<th>Cr</th>
<th>temp. A, ºC</th>
<th>time A, s</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.64</td>
<td>2.53</td>
<td>0.31</td>
<td>0.028</td>
<td>0.031</td>
<td>1.53</td>
<td>0.32</td>
<td>0.008</td>
<td>0.024</td>
<td>0</td>
<td>0</td>
<td>0.049</td>
<td>900</td>
<td>7200</td>
</tr>
</tbody>
</table>

4.1. Stage I - filtering of the global data set according to the adopted criteria

The stated aim of the analysis was to examine the influence of selected parameters of heat treatment ($T_i$ - temperature of isothermal transformation, $t_i$ - time of isothermal transformation) on selected properties of ADI (tensile strength - $R_m$, elongation - $A_5$, hardness - $HRC$). Based on the adopted aim of the analysis, appropriate criteria for filtering the global data set were chosen. It was found that among all the analysed cases, only variants with a similar chemical composition (there should be no major differences in the content of individual elements in the base ductile iron) would take part in the analysis, and the values of the austenitisation process parameters should be similar. The criteria adopted in the filtering operation are given in table 3. It was also very important that the set selected by these criteria comprised measured values of all the analysed properties, i.e. the tensile strength - $R_m$, elongation - $A_5$, hardness - $HRC$.

Of the 260 cases included in the global table, the established criteria were met by 70 cases, which in further course of the study were adopted as a base set for analysis using the ANN.

4.2. Stage II - design of a neural network and choice of the environment for its implementation

With explicit nature of the problem and established set of data for analysis, the design of the network has begun. The process of model building consisted of the following stages: defining the explanatory variables and the explained variables; type selection and determination of the neural network structure; learning of the neural network; evaluation of the network model. The input (explanatory) variables were:
- $T_i$ – the temperature of isothermal transformation,
- $t_i$ – the time of isothermal transformation.

The dependent (explained) variables were:
- $R_m$ – tensile strength,
- $A_5$ – elongation,
- $HRC$ - hardness.

To determine the detailed, yet optimal, network architecture, the STATISTICA software and its Automatic Neural Networks module were used. Several architectures with different number of hidden neurons (arbitrarily it was assumed that the number of neurons should not be less than 4 and more than 10)
and different activation functions (linear, sigmoidal, tangesoidal and exponential) were tested. The STATISTICA Generator allows implementing only one hidden layer, so this parameter was not subjected to analysis. At the time when the learning process was initiated, the neurons between layers were interconnected with each other. Of all the networks generated by the programme, three MLP type networks, which gave the best answers, were finally selected. Networks had 2 input neurons and 3 output neurons, and differed in the number of hidden neurons. Summary of the learning process of selected neural networks and their specific characteristics are given in table 4.

### Table 4. Summary of the network learning process.

<table>
<thead>
<tr>
<th>Network name</th>
<th>MLP 2-7-3</th>
<th>MLP 2-5-3</th>
<th>MLP 2-6-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality (learning)</td>
<td>0.970961</td>
<td>0.972785</td>
<td>0.973469</td>
</tr>
<tr>
<td>Quality (testing)</td>
<td>0.963774</td>
<td>0.957228</td>
<td>0.956306</td>
</tr>
<tr>
<td>Quality (validation)</td>
<td>0.904937</td>
<td>0.922499</td>
<td>0.932627</td>
</tr>
<tr>
<td>Error (learning)</td>
<td>424.604</td>
<td>379.930</td>
<td>298.828</td>
</tr>
<tr>
<td>Error (testing)</td>
<td>1202.007</td>
<td>1420.773</td>
<td>1472.218</td>
</tr>
<tr>
<td>Error (validation)</td>
<td>1828.513</td>
<td>1525.045</td>
<td>1887.503</td>
</tr>
<tr>
<td>Learning algorithm</td>
<td>BFGS 84</td>
<td>BFGS 67</td>
<td>BFGS 129</td>
</tr>
<tr>
<td>Error function</td>
<td>SOS</td>
<td>SOS</td>
<td>SOS</td>
</tr>
<tr>
<td>Activation (hidden)</td>
<td>Logistic</td>
<td>Tanh</td>
<td>Logistic</td>
</tr>
<tr>
<td>Activation (output)</td>
<td>Tanh</td>
<td>Tanh</td>
<td>Linear</td>
</tr>
</tbody>
</table>

4.3. Stage III - analysis of the network parameters

There are many algorithms for learning the MLP networks widely described in the literature (Tadeusiewicz, 1993; Rutkowska et al., 1997; Lula at al., 2007). One of the groups are methods based on Newton’s algorithm. These methods include the BFGS (Broyden-Fletcher-Goldfarb-Shanno) method used in the designed networks. The numbers, shown in the description of the applied learning algorithm, in table 4, indicate in which epoch the network learning has been completed. In the case of MLP2-6-3 networks, termination of the learning process took place in epoch 129.

Data set describing the phenomenon modelled was divided into learning set (70%), validation set (15%) and test set (15%). At the stage of the network learning, the learning set is used and then the validation set, which allows control of the learning process by examining the level of learning of neurons. In fact, learning consists of two phases: selection of weights for the learning set and testing of weights on samples from a validation set. Modification of weight values continues until it reaches a minimum approximation error, or if an error in the validation set begins to grow. The network improves its action solely on the basis of data from the learning set, so if in the course of learning it is observed that the decline of validation error has stopped or the error begins to grow, it indicates that the network has started to be too fit to the learning data and loses its ability to generalise the learning outcomes. Figure 2 shows the change in errors (both learning and validation) in successive epochs for the MLP2-6-3 network.

The final form of the model (trained by the learning set and validated by a validation set) is further tested by a test set. The error is the sum of squared deviations between the setpoint and the output of the network (indicated in table 4 as SOS). After completion of the learning process, the testing process starts using a test set which contains the values not used previously in the network learning process.

**Fig. 2. Changes in errors (learning and validation) depending on the epoch number for the MLP2-6-3 network.**
Table 4 shows the summary done for the three selected networks giving the best results. The given errors (learning, validation and test) refer to the total network error calculated as a mean squared residue for all output variables. The large values of these errors are the result of the order of magnitude of the source data. The magnitude of the validation and testing errors is much larger than that of the learning error - this is due to the low power of the available data set. In the future, the prospect of completing and augmentation of data set will allow reducing this difference.

A detailed analysis of errors for each output variable \((\text{Rm}, \text{HRC}, A_5)\) was performed on the MLP2-6-3 network and is presented in tables 5-7. The network consists of 6 neurons in the hidden layer and the adopted activation functions are: in the hidden layer – a logistic function, in the output layer - a linear function.

The most reliable results include the last three columns, i.e. the relative mean squared error, the relative average deviation and the correlation coefficient, which are independent of the unit of measurement of individual variables, and thus can be mutually compared. In all cases, the relative mean squared error and the relative average deviation are lower for testing than for validation.

### 4.4. Stage IV - Analysis of the results obtained during the modeling

Table 8 presents several examples of the resulting values of properties estimated by the designed model of an artificial MLP 2-6-3 neural network, compared with actual results.

Figure 3a presents the graph of \(\text{Rm}\) (tensile strength) dependence on the time and temperature of austenitering plotted with the use of a neural

![Graph](image-url)
network model. Figure 3b shows the course of this relationship when developed from actual data. Figures 4a and 4b show the dependence between hardness and the examined parameters when derived from a neural network model and real data, respectively.

The construction of more complex structures, and thus the introduction of new input variables to the neural network, is further objective of the authors’ activities. Work is underway to complete the global database. Its development will enable the achievement of this goal.

However, despite a significant reduction in the amount of input parameters adopted in this study, this project can still bring utilitarian benefits. Modelling the effect of isothermal transformation parameters \( T_a, t_i \) on the properties of cast iron of defined chemical composition and constant austenitising conditions (filter installed) may be an excellent tool supporting the work of technologist or researcher in determining the optimum parameters of this treatment (even for a single composition) and helping in

5. SUMMARY AND CONCLUSIONS

The manufacturing process of ADI is a complex process and the resulting mechanical parameters depend on a lot of factors. Building a neural network model, the authors have established a goal of the analysis, which is to study the effect of two selected parameters \( T_a, t_i \) on selected mechanical properties \( (R_m, A_s, HRC) \). The choice of only two variables to build the network was due to the fact that in the whole global set, only between the values of these two parameters, significant differences were observed.

Fig. 3. Plotted relationship between the tensile strength \( R_m \) and the time and temperature of austempering, based on selected neural model (a), based on real data (b).

Fig. 4. Plotted relationship between hardness and the time and temperature of austempering: based on selected neural model (a), based on real data (b).
future elimination of expensive material-involving experiments.

The article presents the developed and tested basic version of a concept, which is based on the integration of experimental results relating to the production of ADI, and the ability to use this concept in the design of inference algorithms and systems enabling automatic processing of the knowledge acquired.

The development and provision of such a system will in the future prevent the multiple repetition of experiments that have already been carried out and can be a source of inspiration to:
- search for new areas of studies;
- search for these fragments of knowledge that can be used to solve a specific problem;
- create new knowledge (discover relationships and dependencies that are not visible in the individual experiments but can occur during the compilation of results obtained in a large number of studies).

The possibility of using the collected data to develop a formal model of the phenomenon of the ADI production based on specific assumptions and with the use of artificial neural networks was presented.

The obtained results should be regarded as preliminary, indicating the possibility of using Artificial Neural Networks to build a model of the ADI manufacturing process. Further work aimed at improving this model will be connected with the process of network optimisation and the introduction of more learning data obtained from the subsequent published experimental data. There is also the possibility of building a neural network for another configuration of the dependent and independent variables based on the sets of learning data filtered from a global set by criteria other than those specified in this study.

ACKNOWLEDGEMENTS

The study was financed under Project AGH No. 11.11.110.300.

REFERENCES


Streszczenie

Podejmowany w artykule problem dotyczy integracji wiedzy o procesie projektowania wyrobów z żeliwa ADI. Specjalistyczna wiedza dziedzinowa, często będąca wynikiem kosztownych eksperymentów, opublikowana na świecie w dużej liczbie czasopism, stanowi bezcenne źródło wiedzy dla jej użytkowników i dla badaczy. Autorzy zwracają uwagę na potrzebę opracowania systemu informatycznego ułatwiającego selektywny dostęp do poszczególnych fragmentów tej wiedzy, pozwalającego na jej automatyczne przetwarzanie. Koszt opracowania takiego systemu jest z pewnością niższy niż koszty wielokrotnego powtarzania eksperymentów. Proponowana, w ramach systemu agregacja i centralizacja wyników przeprowadzonych już badań, może być podstawą do planowania i wykonania kolejnych eksperymentów, pokrywających obszary dotąd niezbadane. Innym rezultatem może być wytworzenie nowej wiedzy, poprzez odkrycie relacji i zależności, które nie są widoczne w poszczególnych, pojedynczych eksperymentach, ale można je ujawnić podczas zestawienia wyników dużej liczby zróżnicowanych badań.

W artykule zaproponowano wykorzystanie sztucznych sieci neuronowych do odkrywania zależności występujących pomiędzy własnościami żeliwa ADI a wybieranymi parametrami obróbki cieplnej w oparciu o zestaw zintegrowanych danych eksperymentalnych, pochodzących z różnych publikacji. Taka forma reprezentacji wiedzy może w przyszłości być wykorzystana w inteligentnych systemach komputerowych w module pozyskiwania wiedzy o procesie wytwarzania żeliwa ADI.

Received: October 27, 2014
Received in a revised form: November 19, 2014
Accepted: December 20, 2014
OPTIMIZATION OF GEOMETRICAL STRUCTURES OF MODERN MATERIALS BY USING A HYBRID EVOLUTION STRATEGY

JACEK JACKIEWICZ

Institute of Technology, Faculty of Mathematics, Physics & Technical Sciences
Kazimierz Wielki University
ul. Chodkiewicza 30, 85-064 Bydgoszcz, Poland
Corresponding author: jaceksnd@aim.com

Abstract

In this paper, a new hybrid evolutionary-gradient algorithm is proposed to solve global numerical problems of topology optimization. The proposed algorithm is developed as a compromise between the direct and gradient-based optimization approaches. A short comparative study of computational results demonstrates the validity of the proposed method, as well as the necessity of its use.

Key words: topology optimization, gradient optimization algorithms, hybrid evolution strategy

1. INTRODUCTION

One of the most fundamental areas of materials science is development of mechanical material properties. In recent years, artificial materials engineered to have properties, which may not be found in nature. Generally, they have extraordinary effective properties, such as negative dynamic modulus and/or density, superior thermoelectric properties, phononic bandgaps and high specific energy absorption. These metamaterials are usually arranged in repeating patterns and are assemblies of multiple individual elements fashioned from conventional microscopic materials such as metals or plastics. Their extraordinary properties are created not by the choice of their chemical composition, but by their exactly-designed geometric structures constituted by precise shape, size, orientation and arrangement.

Geometrical structures of modern materials can now be much more controlled in order to design the structural anisotropy, which determines the deformation pattern and directions of lines of action of the internal forces. Thus by fixing the structural anisotropy, geometrical structures of modern materials could be designed to create complex stress-strain paths to protect a certain internal volume. Natural and produced artificially co-continuous structures, which include hard and soft materials, can provide superior combinations of properties such as stiffness, strength, impact resistance, toughness, and energy dissipation. Optimizing topology and geometric arrangement of components of such a structure provides the possibility of obtaining the combination of desired properties in the macroscale. The topology optimization techniques can be applied, among others, to designing materials at micro and nano-levels. Numerical methods needed to carry out an effective topology optimization of continuum structures have been investigated extensively. Most of these methods are established on the basis of finite and/or boundary element analysis, where the design domain is discretized into a fine mesh of elements.

For a given set of loads and boundary conditions, the primary goal of the topology optimization
procedure is to find the topology of a structure by determining for every point in the design domain whether the specific point of space become a part of the material matrix or the void space. Using the optimization of topology, namely, the optimization of such properties of geometrical shapes that remain the same even when the shapes change, engineers can find the best concept design, which meets the design requirements for performance and manufacturability. Rather than limiting the changes to the sizes of structural components, topology optimization provides much more freedom and allows the designer to create totally novel and highly efficient conceptual designs for continuum structures. Since topology optimization of continuum structures is the most difficult technically as well as the most advantageous economically in comparison with other types of structural optimization, further publication sections will be devoted to this issue.

2. FORMULATION OF A STRUCTURAL OPTIMIZATION PROBLEM (SOP)

In stating any structural constrained optimization problem the objective function \( f \), which value indicates the goodness of the design is always introduced together with the following variables (Christensen & Klarbring, 2009):

- Design variables (collected in a vector \( x \)) that describe the design (i.e. geometrical shape or choice of material), and which can be changed during optimization.
- State variables (collected in a vector \( y \)) that represent the responses of the structure for a given design \( x \) (note that displacements, stresses, strains, forces or vibrations can be assigned to “structural responses”).

Hence, a general structural optimization problem (SOP) can be stated as follows:

\[
\text{SOP} \left\{ \begin{array}{l}
\text{minimize or maximize } f(x,y) \text{ with respect to } x \text{ and } y \\
\text{subject to } g_j(x,y) \leq \bar{g}_j, \\
\text{behavioral constrains on } y, \\
\text{design constrains on } x, \\
\text{equilibrium constraints.}
\end{array} \right.
\]

In structural optimization all kinds of requirements due to the function \( f(x,y) \) that act as constraints can be expressed by

- Inequalities:
  \[ \forall j = 1, ..., r \ (\text{for } r \text{ passive constrains}) \quad g_j(x,y) \leq \bar{g}_j, \]
- Equalities:
  \[ \forall j = r + 1, ..., s \ (\text{for } s-r \text{ active constrains}) \quad g_j(x,y) = \bar{g}_j. \]

In addition there can be constraints on the variables \( x \) and \( y \) themselves, and they are defined as

\[ x \leq \bar{x} \quad \text{and} \quad y \leq \bar{y} \] (2)

or defined by means of subsets of \( x \) and \( y \) with assigned (fixed) values. The process of solving a system with continuously distributed parameters requires a discretization, which produces a discrete parameter system modeled with a finite element method (FEM) or a boundary element method (BEM). Note that a FEM mesh (or a BEM mesh) is only fixed during a given iteration step of the sequential approximate optimization. For static optimization problems discretized by means of the FEM, equilibrium constrains take the form:

\[ K(x) \cdot U = F(x) \] (3)

where \( K(x) \) is the stiffness matrix of the structure, which is a function of the design variables, \( U \) is the displacement vector and \( F(x) \) is the static force vector, which is also depend on the design variables. Note that displacements (collected in a vector \( U \)) take the role of general state variables.

A multi-objective optimization problem (MOOP) deals with several objective functions (collected in an optimization vector \( f \)), such that

\[ \text{minimize}_{x,y} \left\{ f_1(x,y), f_2(x,y), ..., f_l(x,y) \right\} \] (4)

where \( l \) (>1) is the number of objective functions. It is worth to stress that the requirements due to the objective functions (implemented in the forms of constraints) are the same as for the SOP. Note that tasks of minimization and maximization of objective functions are trivially related to each other, since one objective function \( f_k(x,y) \) \((\forall k=1,2,\ldots,l)\) could just as well be converted into another function with minus sign: \( -f_k(x,y) \).

To handle with problems of multi-objective optimization, the Pareto optimal solutions are sought. A structural design is Pareto optimal, if there is no any other design that satisfies all of the objectives better. Thus, the meaning of Pareto optimal solution is as follows: \((x^*,y^*)\) satisfying taken constrains is Pareto optimal, if there is no other feasible solution \((x,y)\) satisfying the constraints, which may combine an value
decrease of at least one objective function without value increases of the remaining objective functions in comparison to the solution \((x^*, y^*)\), and what can be mathematically written as

\[ \exists k \in \{1, 2, ..., 1\}: f_k(x^*, y^*) < f_k(x, y) \quad \forall k = 1, 2, ..., l : f_k(x^*, y^*) \leq f_k(x, y). \quad (5) \]

In multi-objective optimization problems, it is characteristic that no unique solution exists but a set of mathematically equivalently good solutions, which can be described as nondominated, efficient and noninferior.

The method of Lagrange multipliers can be applied to identify candidate optimal points of equality-constrained optimization problems. If functions: \(f\) (objective or performance) and \(g\) (constraints) are entirely functions of \(m\) design variables \((x_1, x_2, ..., x_m)\), then the constrained optimization problem corresponding to only active constraints can be reformulated with a Lagrangian function \(L\) as

\[ L(x, \lambda) = f(x) + \sum_{j=1}^{p} \lambda_j \left[ g_j(x) - \bar{g}_j \right], \quad (6) \]

where the \(\lambda_j\) are the Lagrange multipliers corresponding to the active constraints. Let us assume that the stationary point of Lagrangian function \(L\) corresponds to the local minimum. Partial differentiation of \(L\) with respect to the design variables gives

\[ \frac{\partial L(x, \lambda)}{\partial x_i} = \frac{\partial f(x)}{\partial x_i} + \sum_{j=1}^{p} \lambda_j \frac{\partial g_j(x)}{\partial x_i}. \quad (7) \]

Necessary conditions for \(x^*\) to be a local minimum of \(L\) on the intervals \(x_i^\text{min} \leq x_i \leq x_i^\text{max}\) \((\forall i = 1, 2, ..., m)\) are that

\[ \frac{\partial L(x, \lambda)}{\partial x_i} \bigg|_{x_i=x_i^*} = 0. \quad (8) \]

Conditions (8) are necessary, what means that if they are not satisfied, then \(x^*\) is not a local minimum. Assuming that the Lagrangian function \(L\) is strictly convex, a unique solution to the minimization problem may be derived on the basis of the Karush–Kuhn–Tucker stationary conditions (Christensen & Klarbring, 2009).

3. NUMERICAL METHODS FOR TOPOLOGY OPTIMIZATION

It seems, that there are two main directions of development of topology optimization, which are referred to the following methods: the evolutionary structural optimization (ESO) method and the solid isotropic microstructure with penalization (SIMP) method (Rozvany, 2009). The ESO method is a design method based on the simple concept of gradually removing the inefficient, low-stressed material from a structure by deleting elements with the lowest strain energy density from the finite element model. The idea of the ESO method has developed from the simple one-directional optimization schemes to the bi-directional evolutionary structural optimization (BESO) method, which allows to remove solid elements with the lowest strain energy density from the structure and to change the void elements with the highest strain energy density into solid elements, simultaneously. The initial research on BESO was conducted by Young et al. (1999) for stiffness optimization. According to the SIMP approach the variable density is used to do the topology optimization. In the optimization process each finite element \(e\) is assigned to the density \(\rho_e\), which determines its Young’s modulus \(E_e\) as stated in the following formula:

\[ E_e(\rho_e) = E_{\text{min}} + \rho_e^p (E_0 - E_{\text{min}}) = [y + (1 - \gamma)\rho_e^p] E_0, \quad (9) \]

where \(E_0\) is the real Young’s modulus of the material, \(E_{\text{min}}\) is a very small value of the Young’s modulus, which is assigned to void regions in order to prevent the stiffness matrix from becoming singular, \(p\) is a penalization factor (a given value usually between 2 and 4 to penalize the small values of \(\rho_e\), typically \(p = 3\) acc. Andreassen et al. (2011)) and \(\gamma \approx E_{\text{min}}/E_0\). To obtain a solid-void design with maximum stiffness, the mean compliance is minimized for a given volume of material. The topology optimization problem can be stated as

\[ \min_{\rho} f(\rho) = U \cdot K \cdot U = \sum_{e=1}^{n} [y + (1 - \gamma)\rho_e^p] u_e \cdot k_e \cdot u_e, \quad (10a) \]

subject to

\[ g_0(\rho) - \bar{g}_0 = \frac{v(\rho)}{v_0} - f_V = \frac{1}{v_0} \sum_{e=1}^{n} v_e \rho_e - f_V = 0 \quad (10b) \]

where \(u_e\) is the element displacement vector, \(U\) and \(F\) are the global displacement and force vectors, respectively, \(K\) represents the global stiffness matrix, while \(k_e\) is the element stiffness matrix for the \(e\)-th element with material, which has the Young’s modulus \(E_0\). There are \(n\) finite elements used to dis-
cretize the design domain, with the corresponding vector of elemental densities \( \rho = \{ \rho_1, \rho_2, \ldots, \rho_n \} \) (i.e., the vector of the design variables). Furthermore, \( V(\mathbf{p}) \) represents the material volume, \( v_e \) the volume of the \( e \)-th element, \( V_0 \) the total volume of the design domain, \( f_v \) a prescribed volume fraction and the mark “\("\) means the inner product.

In order to solve the topology optimization problem (10) Bendsoe (1995) has proposed the following heuristic updating scheme for the design variables:

\[
\rho_e^{\text{new}} = \begin{cases} 
\rho_e B_e^{1/\alpha} & \text{if } \rho_e B_e^{1/\alpha} < \rho_e, \\
\rho_e & \text{if } \rho_e B_e^{1/\alpha} \leq \rho_e, \\
\rho_e B_e^{1/\alpha} & \text{if } \rho_e B_e^{1/\alpha} \geq \rho_e,
\end{cases}
\]

\[\rho_e^{\text{new}} = \begin{cases} 
\rho_e B_e^{1/\alpha} & \text{if } \rho_e B_e^{1/\alpha} < \rho_e, \\
\rho_e & \text{if } \rho_e B_e^{1/\alpha} \leq \rho_e, \\
\rho_e B_e^{1/\alpha} & \text{if } \rho_e B_e^{1/\alpha} \geq \rho_e,
\end{cases}
\]

where \( \delta > 0 \) is a prescribed move limit. Then, the algorithm (11) is iteratively solved until the following stopping condition is satisfied:

\[
\| \mathbf{p}^{\text{new}} - \mathbf{p} \| < \epsilon_b \tag{11a}
\]

(11) with \( \epsilon_b > 0 \) and prescribed. In the updating scheme (11), the \( B_e \) are found from the stationary conditions (8) as:

\[
B_e = - \left( \frac{\partial g_0(\mathbf{p})}{\partial \rho_e} \right)^{-1} \frac{\partial f(\mathbf{p})}{\partial \rho_e} \tag{13}
\]

During applying the algorithm (11), the Lagrange multiplier may be found using a simple bisectioning strategy (although more advanced strategies are also possible). The first derivatives of the compliance defined as \( C(\mathbf{p}) = \mathbf{F} \mathbf{U} (\equiv f(\mathbf{p})) \) with respect to the design variables \( \rho_e \) (\( \forall e = 1, 2, \ldots, n \)) can be obtained by the following derivations:

\[
\frac{\partial C(\mathbf{p})}{\partial \rho_e} = \mathbf{F} \cdot \frac{\partial \mathbf{U}}{\partial \rho_e} \cdot (\mathbf{a}) \mathbf{F} \text{ is the static force vector, and}
\]


(b) \( \mathbf{F} \) is independent of \( \rho_e \),

and moreover, \( \partial \mathbf{U}/\partial \rho_e \) is to be

\[
\frac{\partial \mathbf{U}}{\partial \rho_e} = \frac{\partial (K^{-1} F)}{\partial \rho_e} = \frac{\partial K^{-1}}{\partial \rho_e} \cdot \mathbf{K} \cdot \mathbf{U} = -K^{-1} \cdot \frac{\partial \mathbf{K}}{\partial \rho_e} \cdot \mathbf{U}. \tag{15}
\]

Substituting (15) into (14) yields

\[
\frac{\partial C(\mathbf{p})}{\partial \rho_e} = -\mathbf{F} \cdot K^{-1} \cdot \frac{\partial \mathbf{K}}{\partial \rho_e} \cdot \mathbf{U} = -\mathbf{u}_e \cdot \frac{\partial \mathbf{K}}{\partial \rho_e} \cdot \mathbf{u}_e \cdot K^{-1} \text{ is symmetric}. \tag{16}
\]

To reduce the overall number of iterations needed to attain the optimum the objective and constraint functions from the formulation (10) are replaced with their approximations at the current design point (Fadel et al., 1990). In topology optimization the linear, reciprocal, and exponential approximations of the objective and constraint functions are all based on a truncated first-order Taylor series expansion (Groenwold & Etman, 2008).

Substituting the following exponential intermediate variables:

\[
y_e = \left( \frac{z_e - \bar{\rho}_e}{\bar{\rho}_e - \bar{\rho}_e} \right)^a \tag{11b}
\]

(\( \forall e = 1, 2, \ldots, n \)) into the exponential approximation of the objective function \( f(\mathbf{p}) \) described in the formulation (10) about the current design point \( \mathbf{z}^0 \) gives

\[
f_{\text{app}}(\mathbf{z}) = f(\mathbf{z}^0) + \sum_{e=1}^n \frac{\partial f}{\partial z_e}\bigg|_{z_e=\bar{z}_e} \left(\frac{z_e - \bar{\rho}_e}{\bar{\rho}_e - \bar{\rho}_e}\right)^a - 1 \tag{18}
\]

The constraint function \( g_0(\mathbf{p}) \) introduced in the formulation (10) about the current design point \( \mathbf{z}^0 \) is approximated linearly

\[
g_{\text{app}}(\mathbf{z}) = g(\mathbf{z}^0) + \sum_{e=1}^n \frac{\partial g}{\partial z_e}\bigg|_{z_e=\bar{z}_e} (z_e - \bar{\rho}_e) \tag{19}
\]

Necessary conditions (8) for \( \mathbf{z} = \mathbf{z}^* \) to be optimal provide the relationships between the Lagrange multiplier \( \lambda \) and design variables:

\[
\forall e = 1, 2, \ldots, n : \frac{\partial f_{\text{app}}}{\partial z_e} + \lambda \frac{\partial g_{\text{app}}}{\partial z_e} = 0. \tag{20}
\]

By substitution equations (18) and (19) into equation (20) yields:

\[
\left(\frac{z_e - \bar{\rho}_e}{\bar{\rho}_e - \bar{\rho}_e}\right)^{1-a} = B_e \left( \frac{1}{\lambda} \frac{\partial f_{\text{app}}}{\partial z_e}\bigg|_{z_e=\bar{z}_e} \right). \tag{21}
\]

Thus, the explicit expression, which assesses the optimal design variables \( z_e^* \) for the next iteration takes the form:

\[
z_e^* = \bar{\rho}_e + B_e \left( \frac{1}{\lambda} \frac{\partial f_{\text{app}}}{\partial z_e}\bigg|_{z_e=\bar{z}_e} \right). \tag{22}
\]
It is noteworthy that the SIMP and BESO methods are general and can handle any type of structural models and objective functions. However, a straightforward implementation of these methods to perform topology optimization in the finite element framework is, in some cases, hampered owing to drawbacks such as: mesh dependency and numerical instabilities. An adequate remedy to these drawbacks seems to be the application of homogenization methods (Bendsøe & Kikuchi, 1988; Allaire, 2002) but they are mainly restricted to linear elasticity and particular objective functions. Therefore, another type of method known as the level-set method (LSM), which was first proposed by Sethian and Wiegmann (1999) in structure optimization, is becoming more popular because of its computational efficiency. In structural optimization the LSM provides a general framework for the propagation of a material interface between material and void domains by means of a level-set function \( \phi \), which is a continuous Lipschitz function. For the given function \( \phi \), the material domain \( \Omega^+ \), the void domain \( \Omega^- \) and the material interface \( \Gamma \) inside the design domain \( \Omega \) are defined as,

\[
\begin{align*}
\forall x \in \Omega^+ \setminus \Gamma & \quad (\text{material}) : \quad \phi(t, x) > 0 \\
\forall x \in \Gamma (\equiv \partial \Omega^+) & \quad (\text{interface}) : \quad \phi(t, x) = 0 \\
\forall x \in \Omega^- (\equiv \Omega \setminus \Omega^+) & \quad (\text{void}) : \quad \phi(t, x) < 0.
\end{align*}
\]

(23)

So, the level-set function \( \phi \) is used both to represent and evolve the interface \( \Gamma \). In the LSM, apart from the implicit function \( \phi \), the most commonly used functions are the Heaviside function \( H(\phi) \): \[
H(\phi) = \begin{cases} 
0 & \text{if } \phi < 0, \\
1 & \text{if } \phi \geq 0,
\end{cases}
\]

(24)

and the one-dimensional Dirac delta function \( \delta(\phi) \): \[
\delta(\phi) = \frac{dH(\phi)}{d\phi}.
\]

(25)

A small but arbitrary non-zero change in the ordinate of \( H(\phi) \) for a particular value of \( \phi \) within the design domain \( \Omega \) can be determined as follows:

\[
\delta H(\phi) = \frac{dH(\phi)}{d\phi} \delta \phi = \delta(\phi) \delta \phi|_{\phi=0} \Rightarrow \delta(\phi) = 0 \text{ for } \phi \neq 0.
\]

(26)

Following the optimization process, the shape \( \Omega(t) \) is going to evolve according to a fictitious time \( t \). To be precise, assume that the shape \( \Omega(t) \) evolves in time \( t \in \mathbb{R}^+ \) with a velocity \( V(t) (\equiv \dot{x}(t)) \) of each point \( x(t) \) on the boundary \( \Gamma(t) (\equiv \partial \Omega(t)) \). Then

\[
\forall x(t) \in \Gamma(t) : \quad \phi(t, x(t)) = 0.
\]

(27)

Differentiating the above equation in \( t \) yields the Hamilton–Jacobi equation:

\[
\frac{\partial \phi}{\partial t} + \nabla \phi \cdot \dot{x} = \frac{\partial \phi(x,t)}{\partial t} + V_n(x,t) \|\nabla \phi(x,t)\| = 0 \quad \text{(28)}
\]

\[\begin{pmatrix}
\text{tangential components of } V (\equiv \chi) \\
\text{vanish: } \nabla \phi \cdot \dot{x} = \nabla \phi \cdot (V_n n) \equiv V_n \nabla \phi \cdot \|\nabla \phi\|
\end{pmatrix},
\]

which governs the evolution of the level-set function \( \phi \) in time. In equation (28) \( V_n(x,t) \equiv \nabla \cdot n \) is the normal velocity function.

By assuming that points of the material interface \( \Gamma \), which is represented by the zero level set: \( \phi(t, x) = 0 \), move only in the normal direction \( n (\equiv \nabla \phi / \| \nabla \phi \|) \) to \( \Gamma \), a small change in a variable quantity \( \delta \phi |_{\phi=0} \) due to an arbitrarily small increment (i.e., a virtual change) \( \delta l \) along the normal direction \( n \) can be expressed as

\[
\delta \phi |_{\phi=0} = \phi(x + \delta l n) - \phi(x) \equiv \nabla \phi \cdot (\delta l n) = \| \nabla \phi \| \delta l.
\]

(29)

The Lagrangian method can be utilized to convert the original constrained optimization problem (10) into an unconstrained problem as

\[
L_{LS}(t, \Omega, \phi, u, v, \lambda) = \int_\Omega [E(\phi) D : \varepsilon(u)] + \lambda \int_\Omega [H(\phi) d\Omega - f_v V_0],
\]

(30)

where \( \varepsilon \) is the strain field, \( u \) is the actual displacement field, \( v \) denotes the arbitrary virtual displacement field, which displacements are kinematically admissible, \( \lambda \) means the Lagrange multiplier of the volume constraint, \( D \) is the elasticity matrix with unit Young’s modulus and \( E(\phi) \) indicates the design variable, which is defined by means of the following equation:

\[
E(\phi) = E_0 H(\phi) + [1 - H(\phi)] E_{\min}.
\]

(31)

In equation (31) \( E_0 \) and \( E_{\min} \) represent the material elasticity module and the minimum relative elasticity module, respectively. According to the principle of minimum potential energy, the displacement field \( u \) corresponds to the absolute minimum of the potential energy functional.

It is known from the paper by Allaire et al. (2004) that the standard version of the level-set method is in general not capable of creating new holes in the structure during the optimization process. However, the level-set method can change the
topology by closing holes or merging several holes together. To further improve computational efficiency and overcome the difficulty that the standard version of the level-set method cannot generate new holes during the optimization process, it is assumed that the level-set function $\phi(x)$, considering also as phase-field function (Gain & Paulino, 2012), evolves in the direction, which minimizes the objective functional.

It is widely believed (Sigmund, 2011) that for the majority of topology optimization problems the use of non-gradient, nature inspired methods like genetic algorithms is not reasonably practicable. However, the most successful methods nowadays for multiobjective optimization are based on the non-dominated evolutionary or genetic algorithms (Deb, 2001; Madeira et al., 2006). While many of the evolutionary algorithms such as genetic algorithm, cuckoo search algorithm, particle swarm optimization algorithm, artificial bee colony algorithm, harmony search algorithm and artificial immune algorithm are capable of generating Pareto-optimal topologies, a common drawback is that they require numerous FEA-runs, and therefore, from a practical point of view, their computational cost discourages their use to problems of large-scale topology optimization.

4. PROPOSED METHOD OF A HYBRID EVOLUTION STRATEGY AND NUMERICAL EXPERIMENTS

Optimization methods can be arranged into two groups: direct methods (represented, for instance, by optimization techniques that mimic the process of natural selection) and gradient-based methods. In direct search methods, only the objective function and the constraint values are used to conduct the search method, whereas gradient-based methods (e.g., the method of steepest descent) consist in computing the steepest-descent direction associated with the gradient, and then the optimal step-size along this direction. Compared to direct methods gradient-based methods faster converge near an optimal solution, but are not efficient in non-differentiable or discontinuous problems. One of the essential drawbacks of the steepest descent algorithm for problems of topology optimization is that it requires an initial guess. Based on local information arisen from the initial guess, the algorithm suggests a search direction. A unidirectional search is then performed along the search direction to find better solution. This better solution becomes the new solution and the above procedure is continued for a number of times. Figure 1 illustrates this procedure.

In addition, local search technique of gradient-based methods can easily be trapped in local minima. Hence, there is no guarantee that a steepest descent algorithm finds an optimal solution. In general, one can only say that it terminates in a critical point.

Fig. 1. Path of optimum design.

In this research, to overcome the above major problems in the gradient-based optimization methods a hybrid evolutionary-gradient algorithm (HEGA) is proposed to solve global numerical problems of topology optimization with continuous variables. The HEGA combines the multi-objective optimization technique of an evolution strategy based on ideas of adaptation and evolution with the gradient-based algorithm for solving the discrete problem of topology optimization (10). The hybrid evolution strategy has a powerful global exploration capability and uses natural problem-dependent representations, and primarily mutation and selection, as search operators to find the optimum-design domain $\Omega$ for the effective gradient-based algorithm, which searches one of local optimal topologies. In
common with evolutionary algorithms, the operators of the HEGA are applied in a loop. An iteration of the loop is called a generation. The sequence of generations is continued until a termination criterion is met.

Assuming that initial population has N different individuals, which consist of finite sets of n points generated randomly, implicit descriptions of domain geometries can be provided by calculating the convex hull areas. In the plane the convex hull of a finite set \( P \) of points is the unique convex polygon, whose vertices are points from \( P \) and that contains all points of \( P \). According to an alternative definition of the convex hull, \( n \) points, which form the convex hull are compared to nails sticking out of the plane. Thus, the convex hull of \( P \) is an area enclosed by an elastic rubber band (tightened around all nails), which snaps minimizing its length. As far as real-valued search spaces are concerned, mutation is normally performed by adding a normally distributed random value to each vector component of point coordinates. The step size or mutation strength (i.e. the standard deviation of the normal distribution) is often governed by self-adaptation (see evolution window). Only fixed points, which are the support points do not change their coordinates in the numerical exploration of the evolution strategy.

Fig. 3. Impact of changes of the design domain and number of finite elements on the optimal topologies and corresponding compliances for \( f_V = 0.25 \) (the \( f_V \) is calculated for the design domain from figure 2).

Fig. 4. Impact of changes of the design domain on the optimal topologies and corresponding compliances for \( f_V = 0.4238 \) (the \( f_V \) is calculated for the design domain from figure 2).

5. CONCLUSIONS

Topology optimization has been implemented through the use of the hybrid evolutionary-gradient algorithm. The proposed method is a compromise between the computational speed of the effective gradient-based optimization methods and overall accuracy of evolution strategy calculations to determine the global optimum. The calculation results shown in figures 3 and 4 indicate that the choice of the design domain in the task of topology optimization has a significant impact on designs of compliant structures. However, if changes of the design domain are very limited as indicated in figure 5, they will have no significant effect on the compliance
outcome. It should be stressed that the HEGA is an universal method for topology optimization, which is susceptible to modifications due to the specific design tasks.

REFERENCES


Young, V., Querin, O.M., Steven, G.P., Xie, Y.M., 1999, 3D and Multiple Load Case Bi-Directional Evolutionary Structural Optimization (BESO), *Struct Optimization*, 18(2–3), 183–192.
OPTYMALIZACJA PRZESTRZENNYCH STRUKTUR NOWOCZESNYCH MATERIAŁÓW ZA POMOCĄ HYBRYDOWEJ STRATEGII EWOLUCYJNEJ

Streszczenie

W niniejszej pracy zaproponowano nowy, hybrydowy algorytm gradientowo-ewolucyjny do numerycznego rozwiązywania zagadnień z zakresu globalnej optymalizacji topologii. Proponowany algorytm stanowi kompromis między bezpośrednimi i gradientowymi metodami optymalizacji. Króta analiza porównawcza wyników obliczeń wykazuje walidację proponowanej metody, a także konieczność jej stosowania.

Received: November 2, 2014
Received in a revised form: November 18, 2014
Accepted: November 23, 2014
HYBRID PARALLEL EVOLUTIONARY ALGORITHM IN OPTIMIZATION OF 2D GRAPEHENE-LIKE MATERIALS

ADAM MROZEK1*, WACLAW KUŚ2, TADEUSZ BURCZYŃSKI3

1AGH University of Science and Technology, al. Mickiewicza 30, 30-059 Krakow, Poland
2Silesian University of Technology, Akademicka 2A, 44-100 Gliwice, Poland
3Institute of Fundamental Technological Research Polish Academy of Sciences, Pawinskiego 5B, 02-106 Warszawa, Poland
*Corresponding author: amrozek@agh.edu.pl

Abstract

Development and application of the hybrid parallel evolutionary-conjugated gradient algorithm for searching for new, stable atomic arrangements of the two-dimensional graphene-like carbon lattices was described in this paper. The main goal of the optimization is to find stable arrangements of carbon atoms under imposed conditions (e.g. density, shape and size of the unit cell). Such configurations correspond to the minimal values of the total potential energy of the atomic system. Thus, the fitness function is formulated as the total potential energy of the atoms. Interactions between carbon atoms are modeled using Adaptive Intermolecular Reactive Bond Order potential. The parallel approach used in computations allows significant reduction of computation time. Validation of the achieved results and example of the model of new 2D material obtained using presented method were presented in this paper. The numerical scalability tests of the algorithm were performed on the IBM BlueGene/Q supercomputer.

Key words: nano-scale modeling, carbon materials, AIREBO potential, evolutionary algorithm, conjugate gradient minimization, parallel computing

1. INTRODUCTION

Carbon has many allotropes: the ones occurring naturally such as diamond, graphite and amorphous phase, as well as numerous synthetic structures like graphene, nanotubes and their derivatives. This phenomenon is caused by existence of carbon atoms in various hybridization states i.e. atoms of carbon with different electronic configurations, which determine types of bondings, angles between them and spatial arrangement of neighboring atoms.

In the recent years, graphene and similar two-dimensional materials are the subjects of particular interest of researchers (Cranford & Buehler, 2011; Enyashin & Ivanovskii, 2011; Narita et al., 2000; Peng et al., 2012; Scarpa et al., 2009) because of unique electronic, thermal and mechanical properties of such structures. Additionally, two-dimensional graphene-like materials can be used to create another, more complex class of nanostructures, like nanotubes.

Two-dimensional graphene-like materials can be considered as periodic, flat atomic networks, made of stable configurations of carbon atoms in certain hybridization states. It can be observed (figure 1), that each flat carbon network consists of certain parts like benzene rings (or other polygonal elements) and acetylenic linkages of different length connected and replicated periodically in each crystallographic direction. Thus, depend on arrangement of the considered structure, rectangular or triclinic unit cell of given size and atomic density can be identified in the each type of the flat periodic network. An overview of such structures (like graphyne...
and supergraphene), along with detailed description and investigation of their structural and electronic properties using thight-binding method can be found in the work by Enyashin and Ivanovskii (2011). The stress-strain relations and mechanical properties were obtained by Mrozek and Burczynski (2013).

Many examples of theoretically-predicted, two-dimensional graphene-like materials, their periodic, modular structure suggest that potentially another, similar stable constructions can be built. In the combination with unique properties of the two-dimensional graphene-like materials, these facts became an inspiration to develop the method for finding new types of stable flat carbon networks. Such algorithm was proposed by authors and presented in this paper, along with validation of the achieved results and numerical examples.

Since the stable configurations of atoms correspond to the global (or local – in the case of isomers) minima on the Potential Energy Surface (PES), such a task can be considered as an optimization problem. However, searching for the global minimum on the PES is a non-trivial, NP-hard problem, because the number of local minima increases almost exponentially with the number of atoms in the considered structure.

Full quantum ab-initio methods, due to high costs of computations, are usually applied to the relatively small atomic systems. On the other hand, classical optimization methods (e.g. gradient-based ones) have problems with more complex, multimodal functions and reveal tendencies to stuck in local minima. That is why many various heuristic and artificial intelligence-based approaches have been recently applied to deal with similar problem of minimization of the potential energy of the atomic clusters (small, isolated atomic systems). Historically, the first group of the methods is based on searching of the PES, combined with simulation of the certain physical processes, e.g. Random Searches and Monte Carlo (MC) Simulated Annealing (Lloyd & Johnson, 1998) and Basin Hopping Monte Carlo (Wales & Doye, 1997). The second group of computational intelligence methods is inspired by biological mechanisms, present in natural environment and live organisms. The bio-inspired optimization methods of atomic structures, such as Genetic Algorithm (Roberts et al., 2000), Artificial Immune System (AIS) (Shao et al., 2004), and Particle Swarm Optimization (PSO) (Zhou et al., 2008) have become very popular in last years. Both of them give high probability of finding global minimum on the PES, however effectiveness is ransomed by long time of computations.

Authors of this paper successfully implemented set of bio-inspired algorithms: Distributed Evolutionary Algorithm (EA) (Mrozek et al., 2005), AIS and PSO (Mrozek et al., 2010) for investigation of small aluminium clusters with pair-wise Morse and Murrell-Motttram potentials.

Searching for new two-dimensional, graphene-like structures can be performed in the same manner, however needs more sophisticated interatomic interaction model, so called bond-order potential, should be applied. The bond-order potential is able to handle various hybridization states of carbon atoms, allowing creation of bondings with proper, neighborhood-dependent geometry. Additionally, in opposite to the isolated for environment atomic clusters, new algorithm should impose periodicity of the created structure.

Proposed and presented in this work hybrid algorithm combines Parallel Evolutionary Algorithm, prepared by the authors, and conjugated-gradient optimization, built-in LAMMPS software package (Lammps, 2014) which assists forming of the new atomic configuration. Behavior and potential energy of carbon atoms is determined using Adaptive Intermolecular Reactive Bond Order (AIREBO) potential, as developed by Stuart et al. (2000). Presented...
algorithm has modular construction, thus each component can be replaced with functional equivalent (e.g. EA with AIS, gradient optimization with molecular dynamics, etc.) or adapted to use on new computer architectures. It should be noted that routines built-in LAMMPS program are effectively parallelized and can be used on multiprocessor computers (using MPI), as well as on GPUs (via CUDA). Proposed method can be extended to optimization of the three-dimensional molecular structures and may be considered as an alternative approach to existing ones such as the \textit{ab-initio}/PSO algorithm called CALYPSO (Wang et al., 2010).

Presented work, is a continuation of author’s investigations and modeling of atomic systems (Mrozek & Burczynski, 2013) and developed version of approach applied to the minimization of energy of atomic clusters (Mrozek et al., 2010). In the following chapters of this paper, proposed algorithm was described in detail, along with necessary validation of the results and numerical example.

The parallel version of algorithms for IBM BlueGene/Q supercomputer were implemented. The results of scalability tests for up to 512 cores are presented in the paper.

2. HYBRID EVOLUTIONARY-GRADIENT OPTIMIZATION METHOD

Proposed method incorporates two optimization techniques: the heuristic Evolutionary Algorithm (EA) and classical Conjugated Gradient (CG) routine. The EA (Michalewicz, 1996) mimics the mechanisms well-known from biological evolution of species, i.e. EA process the population of individuals. In this application, like in most practical cases, the individuals contain only one chromosome with vectors of genes (so-called design variables). The floating point representation of genes is used in optimization algorithm (Kuś & Burczyński, 2008). The genes (g1, …, gm) are the real-valued Cartesian coordinates of each atom. The range of coordinates is constrained by dimensions of the unit cell of newly created atomic lattice. Thus, the total number of design variables (genes) equals 2m, where 2 is the dimension of the problem and m denotes the number of atoms. Such construction is schematically presented (for three dimensional case) in figure 2.

In the initial population, each individual represents certain, constant number of atoms with randomly generated coordinates placed in the area of the unit cell with periodic boundaries. Dimensions and type (rectangular or triclinic) of the unit cell, as well as number of atoms, are the part of set of the parameters of the simulation. Such approach allows regulation of atomic density of created structure. It should be noted, that periodicity of the atomic structure significantly reduces number of design variables.

![Fig. 2. Construction of the chromosome.](image)

The fitness function is formulated as the total potential energy of the considered atomic system, i.e. sum over all potential energies of particular atomic interactions given by equation (1). Choosing the proper interaction model is crucial for this kind of task. In the presented case, potential energy, as well as neighborhood-dependent behavior of the carbon atoms is determined using Adaptive Intermolecular Reactive Bond Order (AIREBO) potential (Stuart et al., 2000), an extended REBO ($E^{REBO}$) model proposed by Brenner et al. (2002) with additional torsion ($E^{TORSION}_{ij}$) and long range terms ($E^{LJ}_{ij}$):

$$FF = \sum_{i} \sum_{j \neq i} \left( E^{REBO}_{ij} + E^{LJ}_{ij} + \sum_{k \neq i,j} \sum_{l \neq i,j} E^{TORSION}_{ij} \right)$$

Application of AIREBO potential has a drawback, because this kind of model treats the long range interactions in the simplified way, using Lennard-Jones-like function ($E^{LJ}$ term in equation 1) for computation of interactions of this type. Other, more accurate, approach of modeling carbon-carbon interactions can be applied: e.g. ReaxFF model, based on the first-principles calculations, developed by Chenoweth et al. (2008) or second generation Long-range Carbon Bond Order potential (LCBOPII) (Los et al., 2005). However, the AIREBO potential is fitted to handle different spatial configurations and hybridizations types of carbon atoms and is computationally more effective than the ReaxFF approach, which requires additional equilibration of the atomic charge every certain number of iterations (Rappe & Goddard, 1991; Nakano, 1997). Additionally, application of the AIREBO potential to the examination of mechanical
properties of various two dimensional graphene-like materials has been already performed by Mrozek and Burczynski, (2013) yields good agreement with results obtained by other researchers.

It should be noted, that the way of generation initial positions of the atoms may result in very small distances between them, or even overlapped atoms. Such configurations have high potential energies and are very unstable. Similar phenomenon may occur after mutation and crossover operations, performed by EA. That is why initial and offspring populations have to be equilibrated, i.e. potential energy has to be minimized by correction of the positions of the atoms, before the first run of the EA. The coordinates of atoms are transferred to the LAMMPS program and the equilibration process is performed using potential energy minimization routine based on the Polak-Ribiere CG algorithm (Press et al., 2007). The minimized functional was defined in the same way as the fitness function (1), used in EA. Such approach, combined with bond order potential, ensures that structure is not only properly equilibrated, but additionally spatial configurations of atoms conform the rules characteristic for the molecular geometry of flat carbon networks. The periodicity of newly-created structure is also guaranteed by this routine, as well as proper boundary conditions, imposed on the unit cell.

After the conjugated-gradient minimization, the total potential energy (i.e. fitness function) is computed for each individual in the population. This part of algorithm is also handled by the LAMMPS.

The CG optimization is the most time-consuming part of the algorithm. To overcome this problem, authors decided to parallelize proposed algorithm and make it suitable for running on the multiprocessor computers (Kuś & Burczynski, 2008). Thus, the population is scattered into certain number of parts using MPI library. In the next step, each part is further processed in parallel way using dedicated instance of LAMMPS running on separate node of the computer.

Such minimized atomic structures, along with estimated values of fitness function are gathered together and imported to EA. The EA performs selection and invokes evolutionary operators like mutation and crossover. The selection chooses chromosomes for a new parent subpopulation taking into account values of the fitness function. Evolutionary operators change chromosome’s genes and create new chromosomes for the offspring population. The uniform and Gaussian mutations, the simple crossover and ranking selection were implemented in presented algorithm (Kuś & Burczynski, 2008).

The main hybrid evolutionary-gradient algorithm works in iterative way: the offspring population is scattered into parts again, exported to the conjugated-gradient minimizer built-in LAMMPS. In the subsequent step all the data are gathered and exported to the EA, which perform selection of individuals and other evolution operations. These steps are repeated until the stop condition is not satisfied. The stop condition can be formulated as given maximum number of iterations or lack of improvement of the fitness functions during certain time interval. The flowchart of the Hybrid Parallel Evolutionary Algorithm is presented in figure 3.

Algorithm in the form presented in figure 3 is dedicated for optimization rather small atomic systems (up to tens of atoms). However, since the computational routines built-in LAMMPS software package, including CG
energy minimization, are parallelized itself, it is possible to adapt described method to work effectively even with large, three dimensional structures. It can be done by the additional domain decomposition of each atomic system (individual) and running CG in parallel on the cores of each computational node. Parallelization in LAMMPS can be performed in classical manner, using MPI or OpenMP libraries, as well on GPUs, using CUDA. Such massive-parallel allows application of presented method to run effectively on supercomputers or new computer architectures.

3. VALIDATION AND RESULTS

In the first step authors tried to find arrangements of carbon atoms which are already known from literature. Such approach was applied to verify the results, obtained with use of proposed algorithm. Two examples of known 2D graphene-like materials: supergraphene and network made of distorted hexagons, were selected for testing purposes (Enyashin & Ivanovskii, 2011). It should be noted, that presented configuration of the unit cells (shown in figure 1b and 1c, respectively) is non-unique and represents only one of the possible variants. It is an important fact, since proposed algorithm works on unit cells with given shape and atomic density. The results of validation, as well as examples of the obtained new carbon-based structures were presented and described in the following chapter. In the all of presented cases the population consists of 100 individuals. The probabilities of mutation and crossover operations were equal to 10%. Supergraphene (figure 1b) has a honeycomb structure, similar to graphene, where regular carbon-carbon bonds were replaced with acetylenic linkages, thus contains \( sp^2 \) and \( sp^3 \) hybridized atoms. Progress of energy minimization in function of iterations of evolutionary algorithm was presented in figure 4. The final structure of the supergraphene (figure 5a) was obtained in 34th generation. Triclinic unit cell (10 Å × 6 Å), along with bond lengths, was shown in figure 5b. For comparison, the bond lengths computed using tight binding method (Enyashin & Ivanovskii, 2011) equals to 1.248 Å between \( sp^1 \)-\( sp^1 \) atoms and 1.402 Å in the case of \( sp^1 \)-\( sp^2 \) bonds. Analogical distances, time-averaged during run of classical Molecular Dynamics (MD) with the same AIREBO potential, were equal to 1.33 Å and 1.39 Å, respectively (Mrozek & Burczynski, 2013).

The second structure, used for testing purposes, is a combination of supergraphene and standard graphene lattices (figure 1c). This structure is built of distorted hexagons i.e. vertical acetylenic linkages of supergraphene were replaced with double carbon-carbon bonds. This structure also contains \( sp^2 \) and \( sp^1 \) hybridized atoms. The orthogonal unit cell was found with use of proposed algorithm in 5th iteration (figure 6). Resultant lattice and corresponding unit cell were shown in the figure 7a and 7b, respectively. Obtained unit cell has dimensions of 5 Å × 5 Å and contains only 6 atoms instead of the one presented in paper by Enyashin and Ivanovskii, (2011): 7.03 Å × 6.91 Å and 36 atoms, (see figure 1c).
Again, similar difference between bond lengths, obtained using various computational methods and models occurred. The interatomic distances, revealed during optimization were presented in figure 7b. The tight binding computations, performed by Enyashin and Ivanovskii (2011) predicted distances of 1.237 Å between $sp^1$-$sp^1$ atoms. Distances between $sp^1$-$sp^2$ and $sp^2$-$sp^2$ atoms equals to 1.413 Å and 1.422 Å, respectively.

Since test cases yields promising results, hybrid optimization algorithm was applied to search for the new stable configurations of given number of carbon atoms in unit cell of given size and periodic boundaries. In the following example 14 carbon atoms placed in the 7 Å×6 Å rectangular unit cell have undergone the process of minimization of total potential energy. The rest of the algorithm’s parameters remained unchanged.

The progress of optimization was shown in figure 8. The final arrangement of the atoms was obtained in 127th generation. The result – a flat polycyclic network made of dodecagons of $sp^2$ hybridized carbon atoms was presented in figure 9a. More detailed view on unit cell, along with atom’s potential energies and bond lengths, was shown in figure 9b.

The scalability of parallel algorithm was tested on IBM BlueGene/Q supercomputer located at ICM Warsaw. The supercomputer is build from 16 core one chip IBM PowerPC A2 1.6 GHz nodes connected by using 5D torus 40Gb low latency network. Each of the nodes has 16 GB RAM, if all cores are used the 1 GB per core is available. The number of atoms in unit cell is low and the memory consumption is low, and the 1 GB per core is sufficient in discussed problem. The tests were performed for 1, 8, 16, 32, 64, 128, 256 and 512 cores. The lowest possible numbers of nodes were chosen (1 node for 1, 8, 16, 2 nodes for 32, 4 nodes for 64, 8 nodes for...
The number of cores was 128, 16 nodes for 256 and 32 nodes for 512 cores). The number of chromosomes in evolutionary algorithm was 512, the gradient minimization and objective function was computed using single core. The work was divided evenly between cores, ex. for 128 cores test case, each of the 128 nodes performs gradient minimization for 4 chromosomes. The MPI was used in parallelization of evolutionary algorithm. The speedup of the computations were computed and presented in figure 10. The maximum speedup was about 167 for 512 cores. The wall time was measured from start of the program till the end of all computations (including creation of objects etc.). The performance of the algorithm is not close to the linear and optimization of the code is planned in the future.

4. CONCLUSIONS

The hybrid parallel evolutionary algorithm, applied to searching for the new two-dimensional graphene-like materials was described and obtained results were presented in this paper. Proposed method was able to find already-known structures like supergraphene, as well as new, stable one. The final form and properties of optimized structures depends on the assumed size, type and atomic density of the unit cell. Thus, the considered optimization problem can be reformulated, even using multi-criteria approach, and applied to searching for molecular structure with predefined material properties (e.g. stiffness tensor). Every component of the presented algorithm can be replaced with functional equivalent (e.g. optimization method, atomic potential), additionally proposed approach is ready to use in the optimization of the three-dimensional molecular structures. The parallel approach used in the most time consuming parts of algorithm such as conjugated gradient minimization and evaluation of the fitness function significantly reduces the time of optimization. The scalability of the algorithm for up to 512 cores on IBM BlueGene/Q was tested and presented in the paper. Analysis of the times of computations yields great scalability of implemented algorithm. Obtained results and possible applications of described method make it a potential tool for molecular optimization, not only for the flat, carbon systems.

ACKNOWLEDGEMENT

Calculations were performed at the Interdisciplinary Centre for Modelling Mathematical and Computational Modelling at the University of Warsaw under the grant G56-1.
REFERENCES


POSZUKIWANIE NOWYCH, PLASKICH MATERIAŁÓW GRAFENOPODOBNYCH PRZY UŻYCII HYBRYDOWEGO ALGORYTMU OPTYMALIZACJI

Streszczenie


Received: November 13, 2014
Received in a revised form: December 12, 2014
Accepted: December 17, 2014
MULTI-OBJECTIVE OPTIMIZATION OF PHTHALIC ANHYDRIDE CATALYTIC REACTOR USING GENETIC ALGORITHM WITH SIMULATED BINARY JUMPING GENES OPERATOR

VIBHU TRIVEDI, MANOJKUMAR RAMTEKE*

Department of Chemical Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110 016, India
*Corresponding author: mcramteke@chemical.iitd.ac.in

Abstract

Multi-objective optimization problems of chemical industry have been efficiently solved by evolutionary algorithms (EAs). However, due to high computational costs, different concepts are introduced in evolutionary framework for the improvement of convergence speed. One such concept is ‘jumping genes’ which has been adapted in binary-coded genetic algorithm and found to be improving the performance of the algorithm significantly. However, its adaptation in real-coded form lacked the similar success. In an attempt to fill this gap, a new jumping genes operator has been recently developed for real-coded elitist non-dominated sorting genetic algorithm (RNSGA-II), namely, simulated binary jumping genes (SBJG). This work aims at exploring the utility of SBJG for solving real-life industrial optimization problems using a case study of multi-objective optimization (MOO) of an industrial phthalic anhydride (PA) catalytic reactor. The results obtained are found to be converging faster than RNSGA-II and its other existing jumping genes adaptations for both two- and three-objective formulations.

Key words: evolution, genetic algorithm, multi-objective optimization, jumping genes

1. INTRODUCTION

Real-life optimization problems often involve multiple objectives, complex process models and several variables. Since last two decades, EAs and particularly genetic algorithm (GA) (Holland, 1975) has been used efficiently to solve these problems. GA is inspired from Darwinian principle of evolution of species. Originally, it is a binary-coded algorithm which requires the coding of problem variables in binary (0 or 1) form (i.e. each variable represents a set of binaries and corresponding binaries of all variables are arranged one after another in a string to constitute a chromosome) since these symbolize the actual genes closely and facilitate mimicking of genetic operators while their real values are required to evaluate the fitness of chromosomes. Thus, binaries are decoded to the real values, repeatedly. Additionally, this approach has few more limitations (especially, for the optimization problems with continuous variables) such as lack of absolute precision in the solutions obtained, fixed mapping of the problem variables, presence of Hamming cliff, etc. These limitations often make it difficult to achieve the convergence and lead to increase in time and cost of computation. Since real-life optimization problems often involve continuous variables, real-coded algorithms are better suited for these. RNSGA-II is one such popular real-coded algorithm which is a modification of original binary-coded NSGA-II (Deb et al., 2002). It simulates the effect of binary crossover and mutation in real-coded
framework using simulated binary crossover (SBX) (Deb & Agrawal, 1995) and polynomial mutation (PM) (Deb & Agrawal, 1999). Further, these problems involve complex model equations and solving these requires inordinately large computational efforts. In RNSGA-II, these model equations are solved repeatedly for a fixed number of times (= population size × number of generation). Since computational resources are restricted, these problems are often attempted for a limited number of generations and population size. Thus, one has to rely on prematurely converged solutions and their quality depends on the convergence speed of the algorithm. Any improvement in convergence speed is thus highly recommended.

Recently, a new operator, SBJG (Ramteke et al., 2014), is developed to further improve the convergence speed of RNSGA-II. This operator simulates the concept of jumping genes (JG) (McClintock, 1987) from natural genetics and has been successfully applied to several benchmark problems (Ramteke et al., 2014). However, its potential for complex industrial problems is yet to be explored. In this study, two such complex MOO problems related to an important process of chemical industry i.e., phthalic anhydride catalytic reactor are solved using RNSGA-II-SBJG. The results obtained using RNSGA-II-SBJG are compared with that obtained using RNSGA and its other existing JG adaptations and are found to be quite promising with a considerable improvement in convergence speed.

2. SIMULATED BINARY JUMPING GENES (SBJG)

Kasat and Gupta (2003) introduced the concept of JG (or transposons) from natural genetics in the framework of binary-coded NSGA-II (NSGA-II-JG). Subsequently, several other researchers (Sankararao & Gupta, 2006; Bhat et al., 2006; Bhat, 2007; Ripon et al., 2007) proposed various JG adaptations of binary- and real-coded NSGA-II. In NSGA-II-JG, binary chromosomal strings, present in the offspring population, are selected randomly with a jumping gene probability, \( P_{JG} \). The jumping gene sites on these chromosomal strings are then randomly identified. For this, a random number between 0 and 1 is generated and multiplied by \( L_{bc} \) (length of chromosomal string); the result rounded-off to an integer represents one jumping gene site. Similarly, the other site is identified and binaries between these sites are replaced by randomly generated new ones. This procedure involves a macro – macro mutation, and helps in increasing the genetic diversity. A drawback of this adaptation is that several variables get perturbed simultaneously, if the length of jumping gene is too large, compromising the fitness of the chromosome. To improve upon this, Bhat et al. (2006) and Bhat (2007) proposed another adaptation (NSGA-II-aJG) with fixed length jumping genes. In this, one jumping gene site is identified randomly whereas the other is selected \( f_{JG} \) (a predefined constant) binaries later in the same chromosome. This often restricts the perturbation up to maximum two variables. A similar JG adaptation (say RJG1) in real-coded framework for multi-objective simulated annealing (MOSA) has been proposed by Sankararao and Gupta (2006). Here, real-coded variables instead of binaries are replaced by randomly generated new ones following the similar procedure as in NSGA-II-JG and NSGA-II-aJG. This complete perturbation of variables often results in the loss of original information contained in variables and subsequently to their inferior fitness values. In another real-coded adaptation (Ripon et al., 2007) (say RJG2), variables are re-operated using existing operators, SBX and PM, with a probability \( P_{JG} \). Thus, no new operation is added in this adaptation and existing operators are used for simulating the effect of jumping genes.

SBJG is an attempt towards improving the drawbacks of above given real-coded JG adaptations. It simulates the binary-coded JG operation for its implementation in RNSGA-II. In SBJG operation, a variable \( V_j \) (\( j \in [1,n] \)) of a real-coded chromosome (i.e. the array of \( n \) problem variables \( V \)), selected randomly with a probability \( P_{JG} \), is perturbed with following mathematical formulation:

\[
V_j = V_j + \left(V_j^{High} - V_j^{Low}\right) \alpha_j
\]

where, \( \alpha \) is a distribution function given by:

\[
\alpha_j = \begin{cases} 
\left[\left(2R_j\right)^{-\eta_{JG}}\right]^{-1} & \text{if } R_j < 0.5 \\
1 - \left[\left(2\left(1-R_j\right)\right)^{-\eta_{JG}}\right]^{-1} & \text{if } R_j \geq 0.5
\end{cases}
\]

where, \( R_j \) is a random number and \( \eta_{JG} \) (jumping gene index) is a integer random number (IR) between two jumping gene sites \( j_{g1} \) and \( j_{g2} \):

\[
\eta_{JG} = (IR)_j \in [j_{g1}, j_{g2}]
\]

The efficacy of SBJG in simulating the effect of binary JG operation is analyzed by Ramteke et al.
(2014) through a qualitative analysis. In this analysis, the distribution of variable perturbation is calculated for a large number of random instances. The pattern of distribution of local perturbation obtained using SBJG resembles that of binary JG operation whereas the same using RJG1 (i.e., large perturbation) differs completely from that of binary JG operation. As RJG2 utilizes SBX and PM operators, the functioning of SBJG is also compared with that of SBX and PM by calculating the distribution functions of these operators (i.e., $\alpha$, $\beta$ and $\delta$ respectively) for a large number of random instances. It is found that values of $\alpha$ corresponding to same random numbers are distributed for different instances whereas the values of $\beta$ and $\delta$ remain fixed. This uniqueness of SBJG is attributed to variable nature of $\eta_3$ which differs from fixed indices of SBX and PM. Thus, SBJG maintains higher genetic diversity while remaining in the local range of variable perturbation which is similar to binary JG operation.

3. PROBLEM FORMULATION

Phthalic anhydride (PA) is commonly used as a raw material in the manufacture of various polymers (i.e. polyester, plastics from vinyl chloride), insecticides, pesticides, dyes, etc. It is produced by gas-phase catalytic oxidation of o-xylene in a multitubular reactor with V$_2$O$_5$ and TiO$_2$ catalyst packed in many zones in each tube. The processed gas coming out of the reactor is treated in a pair of ‘switch condensers’ used alternately where the gas is cooled to separate the condensate PA. The treated gas is then scrubbed with water, or incinerated catalytically or thermally. Bhat and Gupta (2008) developed a mathematical model of this operation using a reaction network suggested by Skrzypek et al. (1985) (see figure 1a). They have assumed that the reactions take place in a reactor having several identical tubes with each tube comprising nine zones of catalyst and eight intermediate zones of inert packing (see figure 1b). Each of these inert packing zones acts as a cooling region for the high temperature processed gas coming from the previous catalyst zone. The coolant is continuously circulated in the outer side of the tubes to maintain the temperature within the acceptable limits. This model consists of nonlinear algebraic equations for steady state mass and energy balance between the bulk gas and the external surface of the impervious solid catalyst and differential equations for the steady state balance of the bulk gas phase over the length of a single tube. The nonlinear algebraic equations are solved using Powell’s hybrid algorithm (DNEQNF program) whereas the differential equations are solved using Gear’s algorithm (code DIVPAG) from IMSL library, simultaneously. The complete details of model such as model equations, rate expressions, adsorption and kinetic parameters are reported in Bhat and Gupta (2008).

\begin{figure}
\centering
\includegraphics[width=0.7\textwidth]{figure1.png}
\caption{Reaction network for the production of PA (Skrzypek et al., 1985) (a) and PA reactor with 9 zones of catalyst (Bhat & Gupta, 2008) (b).}
\end{figure}

Bhat and Gupta (2008) formulated a two-objective optimization problem involving maximization of the yield of PA which increases the productivity and minimization of the cumulative length of the catalyst beds which decreases the loading of expensive catalyst. The problem formulation involves 9 catalyst zones and 20 decision variables. The details are given as follows:

3.1. Problem 1

Objectives:

$$\max I_1(u) = \frac{X_{\text{PA}}}{1 - \frac{X_{\text{PA}}}{1.2}} + w_1 \left[ 1 - \frac{X_{\text{PA}}}{1.2} \right]^2 + w_2 \left[ 1 - \frac{L_{\text{catalyst}}}{3.6} \right]^2 + w_3$$

(4)
max \( L_1(u) = \left( \frac{1}{1 + L_{\text{cat}}} \right) + w_1 \left[ \left( 1 - \frac{X_{PA}}{1.2} \right)^2 \right] + w_2 \left[ \left( 1 - \frac{L_{\text{cat}}}{3.6} \right)^2 \right] + w_3 \)  

(5)

where, minimization of \( L_{\text{cat}} \) is converted (Deb, 2001) to maximization problem as

\[ u = [c_{in}, T_{F,\text{in}}, T_{c,\text{in}}, m, Z_1, Z_2, ..., Z_8, L_1, L_2, ..., L_8]^T \]

and \( L_{\text{cat}} = \sum_{i=1}^{9} L_i \)  

(6)

Constraints:

\[ T_{\text{max}} \leq 510 \, ^{\circ}\text{C} \]  

(7)

\[ L_i \geq 0.01 \, \text{m} \]  

(8)

Total length of the reactor tube,

\[ L_{\text{tube}} = 3.5 \, \text{m} \]  

(9)

\[ L_9 = \left( 3.5 - \sum_{i=1}^{8} L_i - \sum_{i=1}^{8} Z_i \right) \]  

(10)

Bounds:

\[ 65 \leq c_{in} \leq 85 \, \text{g OX/(m}^3\text{air at NTP)} \]  

(11)

\[ 147 \, ^{\circ}\text{C} \leq T_{F,\text{in}} \leq 287 \, ^{\circ}\text{C} \]  

(12)

\[ 337 \, ^{\circ}\text{C} \leq T_{c,\text{in}} \leq 447 \, ^{\circ}\text{C} \]  

(13)

\[ 0.001 \leq m \leq 0.005 \, \text{(kg coolant)/s} \]  

(14)

\[ 0.2 \leq Z_i \leq 0.45 \, \text{m}, \quad i = 1, 2, ..., 7 \]  

(15)

\[ 0.1 \leq Z_8 \leq 0.45 \, \text{m} \]  

(16)

\[ 0.05 \leq L_i \leq 0.9 \, \text{m} \]  

(17)

\[ 0.01 \leq L_i \leq 0.2 \, \text{m}, \quad i = 2, 3, ..., 8 \]  

(18)

Here, \( X_{PA} \) and \( L_{\text{cat}} \) are the yield of PA (i.e. mass of PA produced / mass OX consumed) and cumulative length of catalyst beds. The decision variables \( c_{in}, T_{F,\text{in}}, T_{c,\text{in}}, m, Z_i, \) and \( L_i \) represent concentration of OX in feed, inlet temperature of feed, inlet temperature of coolant, mass flow rate of coolant, length of \( i^{th} \) inert zone and length of \( i^{th} \) catalyst zone, respectively. In equations (4) and (5), the values of \( X_{PA} = 1.2 \) and \( L_{\text{cat}} = 3.6 \, \text{m} \) in the bracketed penalty functions are guessed to keep these variables in the range of real or expected values. The optimal values of \( L_1 \) – \( L_9 \) and \( Z_1 – Z_8 \) are chosen by the optimization algorithm whereas \( L_9 \) is calculated directly using length balance (see equations (9) – (10)). Equations (7) – (10) represent the constraints on temperature of gas and length of catalyst beds while equations (11) – (18) represent the bounds of the decision variables. The weighting functions for bracketed penalty terms used in equations (4) and (5) are given as:

\[ \text{If } X_{PA} \leq 1.1, w_1 = -500; \text{ else } w_1 = 0 \]  

(19)

\[ \text{If } L_{\text{cat}} \leq 0 \, \text{m}, w_2 = -3000; \text{ else } w_2 = 0 \]  

(20)

In addition to this, hard penalties, \( w_3 \), are imposed in equations (4) and (5) for violation of constraint on maximum allowable temperature of gas given in equation (7). Depending on the location of catalyst bed where the constraint is violated different penalty weight is imposed as given below:

\[ \text{If } T_{\text{max}} \leq 510 \, ^{\circ}\text{C} \text{ in bed } i, i = 1, 2, ..., 9; \]

\[ w_3 = 0; \text{ else } w_3 = -3000 + 250(i-1) \]  

(21)

Also, the hard penalties are imposed to maintain the length of each catalyst bed above a given limit of 0.01 m (see equation (8)) as given below:

\[ \text{If } L_i \geq 0.01 \, \text{m}, i = 1, 2, ..., 9; w_3 = 0; \text{ else } w_3 = 300 \]  

(22)

The diameter of each reactor tube used is 25 mm, mass flux \( G \) is 19.455 kg m \(^{-2}\) h \(^{-1}\) and diameter of catalyst particle used is 3 mm. More details can be obtained from Bhat and Gupta (2008).

Another critical factor in the production of PA is the formation of CO \(_x\). Excess formation of CO \(_x\) causes a severe decrease in the PA yield, reduces the catalyst life and increases the downstream processing cost. Hence, the minimization of CO \(_x\) is also an important objective for the optimal operation of PA reactor. Considering this fact, a three-objective optimization problem (problem 2) is formulated for the first time in this study where the first two objectives are same as problem 1 and the third objective is the minimization of mole fraction of CO \(_x\) in the processed gas.

### 3.2. Problem 2

Objectives:

\[ \text{max } L_1(u) = X_{PA} + w_1 \left( 1 - \frac{X_{PA}}{1.2} \right)^2 + w_2 \left( 1 - \frac{L_{\text{cat}}}{3.6} \right)^2 + w_3 \]  

(23)
\[
\max I_2(u) = \left( \frac{1}{1 + L_{\text{cat}}} \right) + w_1 \left[ 1 - \frac{X_{\text{PA}}}{1.2} \right]^2 + w_2 \left[ 1 - \frac{L_{\text{cat}}}{3.6} \right]^2 + w_3
\]

(24)

\[
\max I_1(u) = \left( \frac{1}{1 + \text{CO}_x} \right) + w_1 \left[ 1 - \frac{X_{\text{PA}}}{1.2} \right]^2 + w_2 \left[ 1 - \frac{L_{\text{cat}}}{3.6} \right]^2 + w_3
\]

(25)

where, minimization of \( \text{CO}_x \) is converted (Deb, 2001) to maximization problem as \( \left( \frac{1}{1 + \text{CO}_x} \right) \).

The constraints are defined by equations (7) – (10), bounds: equations (11) – (18). All other details are same as in the problem 1.

4. RESULTS AND DISCUSSION

The above explained MOO problems are solved using RNSGA-II-SBJG. The values of GA parameters: population size \( (N_p) \), crossover probability \( (P_{\text{cros}}) \), mutation probability \( (P_{\text{mut}}) \), jumping genes probability \( (P_{\text{JG}}) \), crossover index \( (\eta_{\text{cros}}) \), mutation index \( (\eta_{\text{mut}}) \) and jumping genes index \( (\eta_{\text{JG}}) \) are taken (Ramteke et al., 2014) as 100, 0.9, 0.5/total decision variables, 20, 20 and \( IR \in [5, 20] \), respectively. The algorithms are executed on a desktop computer (Intel Xeon E3 – 1225 v3@3.20 GHz processor, 8 GB RAM and Windows 7 operating system).

The Pareto optimal front for problem 1 obtained using RNSGA-II-SBJG for 200th generation is shown in figure 2a. The results are compared with that obtained using RNSGA-II, RNSGA-II-RJG1 and RNSGA-II-RJG2. Figure 2a shows that the results obtained using RNSGA-II-SBJG are superior to that obtained using RNSGA-II, RNSGA-II-RJG1 and RNSGA-II-RJG2. Also, the results with all JG adaptations are superior to that of RNSGA-II. This proves the usefulness of JG operation for real-life MOO problems. Among the JG adaptations, SBJG performs better than the other two adaptations. It is expected to perform better than RJG1 because it uses the distribution function for perturbing the variables instead of complete replacement used in the latter. The complete replacement in RJG1 causes the loss of useful improvement accumulated in the variables over the generations and in turn affects the results badly. Also, SBJG adaptation performs better than RJG2 since it adds a better perturbation process instead of repeating the existing operations of RNSGA-II. Also, the spread of Pareto optimal front for RNSGA-II-SBJG is better than that of other three algorithms. The range of Pareto optimal front, \([X_{\text{PA}}]_{\max}, [L_{\text{cat}}]_{\min}\), using RNSGA-II-SBJG is \([1.171, 0.407]\) which is superior to \([1.166, 0.543]\), \([1.166, 0.512]\) and \([1.169, 0.450]\) obtained using NSGA-II, RNSGA-II-RJG1 and RNSGA-II-RJG2, respectively. Further, a point A (1.168, 0.548) in figure 2a is identified on the Pareto optimal front obtained using RNSGA-II-SBJG beyond which the nature of the curve changes in such a way that marginal increase in yield has to be compromised with significant increase in catalyst loading. Such point can be selected as an operating point. However, more accurate selection requires the in hand experi-
ence and the knowledge of cost factors associated with individual objectives.

The CPU time required for solving 200 generations using RNSGA-II, RNSGA-II-RJG1, RNSGA-II-RJG2 and RNSGA-II-SBJG for problem 1 is 42, 35, 40 and 39 hrs respectively. It shows that CPU time taken by RNSGA-II-SBJG is less than that of RNSGA-II and RNSGA-II-RJG2 and thus performs better in terms of CPU time in addition to convergence as explained above. However, it takes higher CPU time than RNSGA-II-RJG1. In order to compare fairly, both RNSGA-II-SBJG and RNSGA-II-RJG1 are executed for same CPU time of 39 hrs. Also for this case, the results of RNSGA-II-SBJG are found to be considerably better than that of RNSGA-II-RJG1 (see figure 2b).

For problem 2, the Pareto optimal fronts using RNSGA-II, RNSGA-II-RJG1, RNSGA-II-RJG2 and RNSGA-II-SBJG are shown in figure 3a-d for 200th generation. Among all adaptations, the Pareto optimal front of RNSGA-II-SBJG is the best in terms of PA yield and cumulative length of catalyst beds whereas it is comparable with that of RNSGA-II and RNSGA-II-RJG1 in terms of the mole fraction of COx. Also, some additional points are captured by RNSGA-II-SBJG at the lower end of the Pareto optimal front shown by an ellipse in figure 3d. This is primarily due to the fact that in most of the instances, values of distribution function $\alpha$ for SBJG remain close to 0 maintaining the local perturbation. This local perturbation keeps the fitness of a perturbed solution under the acceptable limit. Thus, SBJG in RNSGA-II provides an effective exploration of search space unlike other JG adaptations. Further, the diversity of the solutions is maintained in all variants of RNSGA-II using the concept of crowding distance (Deb et al., 2002) in which the less crowded solutions in objective function space are assigned with higher fitness in the selection process.

For problem 2, the CPU time taken by RNSGA-II, RNSGA-II-RJG1, RNSGA-II-RJG2 and RNSGA-II-SBJG is 44, 36, 43 and 40 hrs, respectively. Again, the CPU time taken by RNSGA-II-SBJG is less than that of RNSGA-II and RNSGA-II-RJG2 whereas it is higher than that of RNSGA-II-RJG1. Thus, the results of RNSGA-II-SBJG are compared with that of RNSGA-II-RJG1 (see figure 3e) for same CPU time of 40 hrs. The results of RNSGA-II-SBJG show better convergence compared to that of RNSGA-II-RJG1 for same CPU time of 40 hrs. Thus, it can be inferred from the re-
sults of problems 1 and 2 that RNSGA-II-SBJG performs better than all other adaptations.

5. CONCLUSIONS

Two MOO problems related to PA reactor are solved by incorporating a recently developed jumping genes operator, SBIG, in RNSGA-II. The results are found to be better than RNSGA-II and its two JG adaptations i.e., RNSGA-II-RJG1 and RNSGA-II-RJG2. This indicates that SBIG operator could prove quite useful for solving other computationally intense MOO problems similar to those solved in this study. Moreover, the use of SBIG operator is not limited to RNSGA-II and it can be incorporated in other real-coded algorithms such as simulated annealing, particle swarm algorithm, etc. This could be a subject of future studies.

ACKNOWLEDGEMENTS

Partial financial support from the Department of Science and Technology, Government of India, New Delhi (through grant SERB/F/1510/2014-2015, dated June 5, 2014) is gratefully acknowledged.

REFERENCES


WIELOKRYTERIALNA OPTYMALIZACJA REAKTORA BEZWODNIKA FTALOWEGO KATALITYCZNEGO Z WYKORZYSTANIEM ALGORYTMU GENETYCZNEGO Z OPERATOREM PRZESUNIĘCIA GENÓW

Streszczenie

Wielokryterialna optymalizacja problemów przemysłu chemicznego rozwiązywana jest z powodzeniem z wykorzystaniem algorytmów ewolucyjnych (EAs). Jednak w związku z długimi czasami obliczeń, dotychczasowe algorytmy są modyfikowane dla poprawy szybkości zbieżności. Jedną z możliwości jest idea „przesuwania genów”, zaadoptowana dla algorytmu genetycznego z kodowaniem binarnym, która znacznie poprawiła działanie algorytmu. Jednakże wprowadzenie tej modyfikacji do algorytmu z kodowaniem zmiennoprzecinkowym nie przyniosło spodziewanych efektów. Rozwiązianiem tego problemu było wprowadzenie nowego operatora przesunięcia genów do algorytmu RNSGA-II, a mianowicie symulowanego binarnego przesunięcia genów (SBIG). Celem pracy była ocena możliwości zastosowania algorytmu SBIG do rozwiązania rzeczywistych problemów przemysłowych na przykładzie wielokryterialnej optymalizacji (MOO) reaktora bezwodnika ftalowego katalitycznego. Otrzymane wyniki pokazały lepszą zbieżność zastosowanego algorytmu w porównaniu z metodą RNSGA-II i innymi dotychczasowymi adaptacjami operatora przesunięcia genów zarówno dla dwu- jak i trój-kryterialnego sformułowania problemu.
DIFFERENTIAL EVOLUTION FOR SUSTAINABLE SUPPLIER SELECTION IN PULP AND PAPER INDUSTRY: A DEA BASED APPROACH

SUNIL KUMAR JAUHAR1*, MILLIE PANT1, MAHESH C. NAGAR2

1 Indian Institute of Technology, Roorkee-Haridwar Highway, Roorkee (UK), India, 247667
2 Swadesh International Ltd, Holly Building, Hong Kong Street, Singapore, 059675
*Corresponding author: suniljauhar.iitr@gmail.com

Abstract

Management of sustainable (ecologically green) supply chain is a vital concern for any industry around the world. The pulp and paper industry in INDIA is facing tremendous challenges like tumbling demands due to ecological concern, gloomy trend in technology advancement, and fierce worldwide competition. This highly competitive operating environment is of much concern and hence attention is focused towards the direction of climatic changes and additional ecological concerns. These concerns are voiced by activists, researchers, and the common man, which has compelled the industries to take suitable actions for sustainability. This paper aims to examine the challenges of sustainable supplier selection (SSS) and proposes a Differential Evolution (DE) based approach for selecting sustainable suppliers for a hypothetical north Indian pulp and paper industry’s supply chain management (SCM).

Key words: sustainable supplier selection, supply chain management, differential evaluation, data envelopment analysis, pulp and paper industry

1. INTRODUCTION

Nowadays, sustainable development has become a topic of discussion among almost all the segments of our society and is getting a favourable response from different units like manufacturing (Jayal et al., 2010), business development (Floridi et al., 2011), tourism (Luthe & Schuckert, 2011) and agriculture (Paoletti et al., 2011).

In SCM both academics and general practitioner contemplate the sustainable concerns in their workings. The significance of Sustainable SCM has emerged as a result of the present-days commercial conditions of worldwide competition, globalization of supply chains, small product life cycles, immediate modifications in technologies, the need to deliver greater levels of consumer service, and the continual

Fig. 1. Sustainable supply chain with stages and relationships field (Sarkis, 2012).
force to decrease prices, increase asset use while taking care of environment concerns. Figure 1 presents a flow diagram of sustainable supply chain with stages and relationships field.

In view of the growing awareness concerning sustainability in the pulp and paper industries, the SSS would be the vital element in the process of managing a sustainable supply chain. The nature of this decision usually is difficult and is many times unstructured. Now, it is the duty of the purchasing managers to identify most suitable clusters of sustainable suppliers for their product. Optimization practices might serve as useful tools for these types of decision-making problems.

During last few years, DE has arisen as a dominating tool used for solving a variety of problems arising in numerous fields. Proposed by Storn and Price (1995), a population set based evolutionary algorithm was applied successfully to wide-ranging issues (Plagianakos et al., 2008; Wang & Jang, 2000; Joshi & Sanderson, 1999; Ilonen et al., 2003; Ali et al., 2011).

The proposed approach is an extension of method proposed by Jauhar et al. (2014a), which focuses on developing suitable supplier clusters. In Jauhar et al. (2014a) the authors presented an approach for SSS problem where the input criteria are lead time, price and quality of the delivered goods and the output criteria are service quality and CO2 emission of the product and services. After successfully applying this approach with traditional output criteria for SSS, we further apply this approach for value-added SSS in pulp and paper industry. In this study we have taken reusability as output criteria with CO2 emission of parts reuse and material recycling has become an important business strategy in pulp and paper industry (Mazhar et al., 2005).

Materials can have an impact on the environment in a range of different ways and at different times during their life cycles. For example, the extraction, transportation and manufacturing of raw material consume energy and produces carbon. However, when a material is reused or recycled the wider environmental impact of the material is significantly reduced (http://www.ukgbc.org/content/materials).

The paper manufacturers are striving for new approaches and methods to embrace the extended cycle for supplied products. They need economically competitive as well as environmentally friendly end of life material recovery options that help to transform the classical cradle-to-grave economy into a cycle economy.

The main difference between the earlier paper of the authors (Jauhar et al., 2014a) and this paper is that, in the previous paper the authors have considered the traditional output criteria for supplier selection and have considered a model accordingly. While in this paper, the authors have discussed the problem considering the aspect of global sustainability, for this the authors have also considered the impact of reusability and CO2 emission. The output criteria are changed accordingly and modified model is considered in addition to the model considered in Jauhar et al. (2014a). Further, in this paper the study, including literature review, is done in more detail.

This paper is organized in eight sections. Subsequent to the introduction in section 1, the sustainable SCM in pulp and paper industry, sustainable supplier selection and methodology are briefed in 2, 3 and 4 sections respectively. Section 5 describes the mathematical model formulation with DEA used in this article. Section 6 describes the DE algorithm for SSS and a case on pulp and paper industry is discussed in section 7. Finally, discussion with conclusion of the current research is given in last section.

2. PULP AND PAPER INDUSTRY SSMC

The paper products manufacturing companies has a significant and complicated role in the worldwide carbon cycle. Pulp and paper are massive consumers of energy. In fact, the world’s fifth-leading consumer of energy is the pulp and paper industry.

The World Resources Institute, a body of experts, placed the industry’s CO2 emissions at around 500m tons worldwide in 2005 (Martel et al., 2005). Rather, European companies are reasonably green. The Confederation of European Paper Industries (CEPI), a trade association, states its associates’ emissions were 46m tons in 2011.Greenhouse gas (GHG) emissions from the pulp and paper source group are most of the part of CO2 with lesser quantity of CH4 as well as N2O (Philpott & Everett, 2001).

Indian pulp and paper industries produces a great amount of paper as well as cellulose based fibre products. Bulletin papers, copy papers, different kinds of tissue, bottle sticky label, cigarette papers plus coffee filter are just a small number of patterns of the products frequently used in our daily life.

At present in India, there are 759 pulp and paper industries with an installed capacity of 12.7 million tons producing around 10.11 million tons/annum of
paper/paper board and newsprint out of the total world production of around 402 million tons. The Indian pulp and paper industries structure consists of small, medium and large-sized paper mills having production capacities ranging from 10 to 1,150 tons per day (Chakraborty & Roy, 2014).

The industry employs wood, agro residues and recycled/waste paper as the major raw material for manufacturing different varieties of paper, paper board and newsprint. In 2000, the share in production of paper from wood-based raw materials, agro residues and recycled/waste paper had been 39, 31 and 30 %, respectively (Kulkarni, 2013).

Pulp and paper industries in India continue to face challenges with forest (wood)-based raw materials. The projected demand for paper by 2025 is 24 million tonnes leading to an estimated shortfall of 12 million tons of wood (Kulkarni, 2013).

There is a huge amount of activity involved in the chain behind these products; such a system of actions is acknowledged as supply chain in management as well as operation research works (https://www.cirrelt.ca/DocumentsTravail/2006/DT-2006-AM-3.pdf).

These industries are categorized by a huge and extremely incorporated supply chain. The entire supply chain starts in the procurement network, carry on through production network, distribution network and finishes by sales network. In this age, concern is usually for the sectors effect on external network of pulp and paper supply chain organization and its emissions of CO2. Figure 2 presents an easy illustration of the wood based pulp and paper industry supply chain organization with corresponding main SCM level.

3. SSS IN PULP AND PAPER INDUSTRY

Supplier selection is a significant part of SCM. Supplier selection and evaluation forms an integral part of a supply chain. A wrong choice or decision may lead to unpleasant circumstances and in worst case may even lead to the deterioration of the entire supply chain’s financial and operational position. The supplier selection practices are comprehensively studied in the literature with multi-criteria decision making models (MCDM). Different researchers have studied the works in the past relating the sup-

Fig. 2. The pulp and paper supply chain organization (Martel et al., 2005).

Indian business organizations are gradually identifying that the effective management of sustainable supply chains is a primary driver of value creation as well as environmental performance. Industries in the pulp and paper area have now started to identify the immense scope and potential future prospects that exist for sustainability in SCM.
Now days, a growing environmental consciousness has favoured the rise of the new sustainable supply chain paradigm; as a result, similarly in the suppliers selection issues, sustainable criteria have been integrated. Traditionally, only the management aspects like lead time, quality, price, late deliveries, rate of rejected parts and service quality of the supply chain were considered for selecting a potential supplier. However, with the growing environmental issues researchers are also paying attention to factors like greenhouse effect, reusability, carbon-di-oxide (CO₂) emission etc. The resulting problem is called ‘Sustainable supplier selection (SSS)’, where a balance is maintained between the management and environment concerns.

Recently, Klassen and Vachon (2003) examined the role of suppliers in plant level ecological enrichments in the Canadian printing industry in addition to describing the significance of cooperation for supplier investments in environmental friendly technologies. To select the potential suppliers, two focuses comprise significance: one is the degree of the selection criteria, and the other is the suppliers’ sustainable performance, these two focuses need to be verified with the appropriate decision makers (Buyukozkan & Çifçi, 2011).

Towards accomplishing a sustainable supply chain, entire associates in the chain from raw material suppliers to topmost administrators must have natural liking in relation to sustainability. Even now, comprehensive SCM study is yet to be accomplished on how corporations can contain suppliers in sustainable management practices and involve them into sustainable activities.

4. METHODOLOGY

To measure and analyse the relative efficiency of pulp and paper industries suppliers, we follow a four step methodology:

1) Design a criteria containing input and output criteria.
2) Select a problem.
3) Formulate the mathematical model of the SSS problem with the help of DEA.
4) Apply DE on mathematical model.

The present model can be carried out for any quantity of suppliers and there is no limitation, by using this model, the company can obtain a recommended combination of efficient suppliers.

5. MATHEMATICAL MODEL FORMULATION WITH DEA

DEA based method is used for determining the efficiencies of Decision-Making Units (DMUs) on the basis of multiple inputs and outputs (Dimitris et al., 2009). DMU can comprise of business firms, divisions of huge groups such as institution of higher education, schools, hospitals, power plants, police stations, tax offices, prisons, a set of organizations (Ramanathan, 2003; Wen & Chi, 2010; Dobos & Vorosmarty, 2012; Kumar et al., 2011). The DMU described in this research work using input as well as output criteria is as follows.

The performance of DMU is estimated in DEA by the concept of efficiency or productivity, which the proportion of weights sum of outputs \((o/p)\) to the weights sum of \((i/p)\) inputs (Srinivas, 2000) given by equation:

\[
\text{Efficiency} = \frac{\text{Weighted sum of } o}{\text{Weighted sum of } i} \quad (1)
\]

The two basic DEA models are the CCR (Charnes, Cooper and Rhodes) model (Charnes et al., 1978) and the BCC (Banker, Charnes and Cooper) model (Banker et al., 1984), these two models may be distinguished on the basis of returns to scale assumed. The former assumes constant returns-to-scale whereas the latter assumes variable returns-to-scale (Dimitris et al., 2009). In the current study we use CCR model which is well-defined further down: suppose that there are \(N\) DMUs and each unit have \(I\) input and \(O\) outputs then the efficiency of \(m^{th}\) unit is achieved by resolving the below model given by Charnes et al. (1978):

\[
\begin{align*}
\text{Max } E_m = & \sum_{k=1}^{O} w_k \cdot \text{Output}_{k,m} \\
\text{subject to } & \sum_{i=1}^{I} \text{Input}_{i,m} \cdot z_i \\
& \sum_{k=1}^{O} w_k \cdot \text{Output}_{k,m} \leq 1 \quad n = 1, 2, \ldots, m \ldots N \\
& \sum_{i=1}^{I} \text{Input}_{i,m} \cdot z_i \leq 1 \quad n = 1, 2, \ldots, m \ldots N \\
& w_k, z_i \geq 0 \quad \forall k, l \\
\end{align*}
\]  

(2)

where:

\(E_m\) is the efficiency of the \(m^{th}\) DMU, \(k=1 \text{ to } O, \ l=1 \text{ to } I\) and \(n = 1 \text{ to } N\); \text{Output}_{k,m} is the \(k^{th}\) output of the \(m^{th}\) DMU and \(w_k\) is weight of output \text{Output}_{k,m}; \text{Input}_{i,m} is the \(i^{th}\) input of \(m^{th}\) DMU and \(z_i\) is the weight
of $\text{Input}_{l,m}$; $\text{Output}_{k,n}$ and $\text{Input}_{l,n}$ are the $k^{th}$ output and $l^{th}$ input, respectively of the $n^{th}$ DMU, where $n=1, 2, \ldots m \ldots N$.

The fractional equation (2) can be converted in a linear equation (3) presented below. To calculate the efficiency score for each DMU we run the above program $N$ times. A DMU is considered efficient if the efficiency score is 1 otherwise it is considered inefficient.

$$\text{Max } \sum_{k=1}^{O} w_k \text{Output}_{k,n}$$
$$\text{s.t.}$$
$$\sum_{l=1}^{I} z_l \text{Input}_{l,m} = 1$$
$$\sum_{k=1}^{O} w_k \text{Output}_{k,n} - \sum_{l=1}^{I} z_l \text{Input}_{l,m} \leq 0 \quad \forall n$$
$$w_k, z_l \geq 0 \quad \forall k, l$$

6. DE ALGORITHM

DE algorithm is a type of evolutionary algorithm, most effectively used in optimization problems.

A buyer (decision maker) can effect an assessment (supplier evaluation) with the ability to choose weight system. For this purpose with the help of program implemented in DEV C++, we generate uniform random numbers (between 0 to 1) using the inbuilt function `rand()` in DEV C++, the fitness value is taken as the average fitness value in 30 runs and the program is terminated when Max-Iteration is reached, to assist the selection of the weights for input as well as output criteria in a manner to permit the control of the result for the sustainable supplier evaluation and assessment practice.

6.1. Experimental settings

The parameter settings for DE are given in table 1.

<table>
<thead>
<tr>
<th>Table 1. Parameter setting for DE.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pop size (NP)</td>
</tr>
<tr>
<td>Scale Factor (F)</td>
</tr>
<tr>
<td>Crossover rate (Cr)</td>
</tr>
<tr>
<td>Max Iteration</td>
</tr>
<tr>
<td>DE Scheme used</td>
</tr>
<tr>
<td>Constraint handling</td>
</tr>
</tbody>
</table>

7. CASE ON SUSTAINABLE SUPPLIER SELECTION

The case study presented in this paper stands a hypothetical data (Jauhar et al., 2013) for pulp and paper industry in Northern India (X Company). The Raw Material is assumed to be agro based. Around 150 Indian pulp and paper industries use agro residues like bagasse, wheat and rice straw etc. and produces 2.2 million tonnes of material which is 20% of total production. After verifying a group of criteria in a viewpoint of sustainable merits, some criteria including lead time, quality, price, reusability and CO2 emissions of the delivered products are derived for SSS problem.

7.1. Designing a criteria

In current study we split the criteria in two manners: the input and output criteria (shown in figure 3). The input criteria are the traditional supplier selection criteria, such as lead time, price and quality of the delivered goods. The output criteria are the reusability and CO2 emission of the product and services.

We assume that the reusability and CO2 emission are the output of the examined model. Reusability concept is taken from Mazhar et al. (2005) and for CO2 Emissions, LOCOG Guidelines on Carbon

**Table 2. Data for numerical example.**

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Management criteria (Inputs)</th>
<th>Environmental criteria (Outputs)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lead Time (L) (Day)</td>
<td>Quality (Q) (%)</td>
</tr>
<tr>
<td>Suppliers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>80</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>70</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>90</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>65</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>55</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>70</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>85</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>95</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>67</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>72</td>
</tr>
<tr>
<td>11</td>
<td>5</td>
<td>51</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>58</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>72</td>
</tr>
<tr>
<td>14</td>
<td>4</td>
<td>60</td>
</tr>
<tr>
<td>15</td>
<td>5</td>
<td>63</td>
</tr>
<tr>
<td>16</td>
<td>3</td>
<td>90</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>87</td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>82</td>
</tr>
</tbody>
</table>

**7.2. Selection of a problem**

The data is shown in table 2 with the supplier’s database covering management (input) as well as environmental (output) criteria of an item provided in the shipment of company.

**7.3. Mathematical model**

Based on the basis of the above data the DEA model of Kth DMU with the help of equation (3) will be as follows:

\[
\max \quad S\mathcal{Q}_m + CE_n \\
\text{s.t.} \\
\quad z_1L_m + z_2Q_m + z_3P_m = 1 \\
\quad w_1S\mathcal{Q}_n + w_2CE_n - z_1L_m - z_2Q_m - z_3P_m \leq 0 \\
\quad \forall \ n = 1, \ldots, m, \ldots, 18
\]

**7.4. Applying DE on mathematical model**

After applying DE on sustainable supplier selection problem, in table 3 average efficiency and weights results of all DMUs are given. In table 4 results of all DMUs are given and figure 4 shows the histogram of all suppliers with their efficiency score.

**Table 3. Average efficiency and weights in 30 runs.**

<table>
<thead>
<tr>
<th>Suppliers</th>
<th>Value of input and output weight</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Z_1)</td>
<td>(Z_2)</td>
</tr>
<tr>
<td>1</td>
<td>0.100129</td>
<td>0.00204427</td>
</tr>
<tr>
<td>2</td>
<td>0.9899</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0.0111122</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0.0153862</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0.0181836</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0.0142871</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0.71429e-018</td>
</tr>
<tr>
<td>8</td>
<td>0.442549</td>
<td>0.00121055</td>
</tr>
<tr>
<td>9</td>
<td>0.989</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0.0794687</td>
<td>0.00947534</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0.0196098</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0.0167855</td>
</tr>
<tr>
<td>13</td>
<td>0.0944849</td>
<td>0.0112657</td>
</tr>
<tr>
<td>14</td>
<td>0.0069216</td>
<td>0.0121538</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0.0158746</td>
</tr>
<tr>
<td>16</td>
<td>0.0728352</td>
<td>0.00868438</td>
</tr>
<tr>
<td>17</td>
<td>0.963</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>0.0117559</td>
</tr>
</tbody>
</table>
## Supplier Efficiency Ranking

<table>
<thead>
<tr>
<th>Supplier</th>
<th>Efficiency</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0.833417</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>0.744619</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>0.525436</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>0.850091</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>0.81318</td>
<td>14</td>
</tr>
<tr>
<td>7</td>
<td>0.909066</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>0.916572</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>0.855927</td>
<td>11</td>
</tr>
<tr>
<td>11</td>
<td>0.907594</td>
<td>10</td>
</tr>
<tr>
<td>12</td>
<td>0.958812</td>
<td>5</td>
</tr>
<tr>
<td>13</td>
<td>0.908605</td>
<td>9</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>0.390023</td>
<td>17</td>
</tr>
<tr>
<td>16</td>
<td>0.252156</td>
<td>18</td>
</tr>
<tr>
<td>17</td>
<td>0.937594</td>
<td>6</td>
</tr>
<tr>
<td>18</td>
<td>0.959331</td>
<td>4</td>
</tr>
<tr>
<td>Average</td>
<td>0.820134</td>
<td></td>
</tr>
</tbody>
</table>

From table 4 we can see that for suppliers 1, 9 and 14, the efficiency score is 1 so these supplier are assumed to be 100% efficient while efficiency score for all other supplier is less than 1 indicating that these suppliers are not as efficient and among these, supplier no. 16 is probably the most inefficient in comparison to all other suppliers.

### 8. DISCUSSION AND CONCLUSION

For performance assessment of the suppliers, three inputs and two outputs, as described in section 7.1, are taken and the results are shown in table 3. The research of efficient SSS practice can acquire a desirable cluster of competent sustainable suppliers 1, 9 and 14 using DE algorithm. The investigation shows that out of 18 suppliers, only 3 suppliers, namely, 1, 9 and 14 are most efficient. All the remaining suppliers are relatively less efficient as they have the efficiency scores less than one.

The average efficiency score using DE algorithm is 0.820134. Five suppliers 3, 4, 6, 15 and 16 scored lower than the average efficiency score. The lowest efficiency score (0.252156) is calculated for the 16. So the overall performance of 16th supplier is very poor.

For the current research conducted in 18 suppliers, the results are:

1. For suppliers 1, 9 and 14, the efficiency score is 1 so these suppliers are assumed to be 100% sustainable efficient.
2. Supplier 16 is probably most inefficient in comparison to all other suppliers.
3. Suppliers 1, 9 and 14 would be the most suitable set of suppliers (or key suppliers).
4. By using this DE algorithm, the business firms can acquire desirable clusters of competent sustainable suppliers.
5. Combination of suppliers 1, 9 and 14 would be the desirable clusters of competent sustainable suppliers set meanwhile the business firms requiring single-item sustainable suppliers.

In this study, we presented an approach to solve the multiple-criteria SSS problem with the application of DE, for data envelopment analysis (DEA) based mathematical model, and validated this approach with the help of a case taken from north Indian pulp and paper industry. Present study shows DE algorithm as a tool for selecting the optimal sustainable suppliers. Numerical results validate the efficiency of DE for dealing with such problems.

The key offerings of this research are précised as below:

1. SSS in SCM: to date, there are a small number of researches seeing sustainable concern in the supplier selection practice.
2. The selection criteria on the basis of sustainable concern are collected by means of the literature after that these are put in to the mathematical model for the SSS practice.
3. The present model can be used for analyzing any number of suppliers and criteria in the great size business firms.
4. In spite of the fact that lots of efforts have been made for the supplier selection, taking into consideration sustainable concern for this problem remains a demanding task.
5. In this study, the goal was the application of DE algorithm to the efficient SSS in the SCM.
6) A case on pulp and paper industry validates the application of the present approach.

Future research may explore the practice of the DE for finding a solution to more difficult problem with additional sustainable input and output criteria. Sensitivity analysis can also be used for the measuring the influence of each criteria for efficient sustainable suppliers. It is the hope of authors that the results of the DE and DEA model presented in this paper will stimulate further researches in the use of DE for sustainable supplier’s evaluation and selection.

REFERENCES


Zarządzanie zrównoważonym (ekologicznie zielonym) łańcuchem dostaw jest istotnym problemem dla każdej branży przemysłowej na całym świecie. Przemysł celulozowo-papierniczy w Indiach stoi przed wielkim wyzwaniem związanym z dbałością o ekologię, pozostającą w sprzeczności z postępem technologicznym, i ostrą konkurencją na świecie. Konkurencja i rynek wymuszają troskę o środowisko w aspekcie zmian klimatycznych i innych problemów ekologicznych. Swoje obawy zgłaszają działacze, naukowcy, a także zwykli ludzie, zmuszając przemysł do podejmowania właściwych działań na rzecz zrównoważonego rozwoju. W pracy badano problem zrównoważonego wyboru dostawców (SSS). Zaproponowano zastosowanie algorytmu ewolucji różnicowej (DE) do rozwiązania postawionego zagadnienia SSS dla przykładowego przemysłu celulozowo-papierniczego na północy Indii, jako elementu zarządzania łańcuchem dostawców (SCM).
OPTIMIZING THE PERFORMANCE OF ELECTRICALLY POLED POLYMERIC FILMS

RENU TYAGI1*, MILLIE PANT2, YUVRAJ SINGH NEGI1

1 Department of Polymer Science & Process Engineering, Indian Institute of Technology Roorkee, India
2 Department of Applied Science & Engineering, Indian Institute of Technology Roorkee, India

*Corresponding author: renutyagi80@gmail.com

Abstract

In this paper organic guest host system using 2-methyl-4-nitro aniline (2-MNA) as guest material and polyether sulfone (PES) as host material is considered for analysis. Thin and transparent film samples are prepared by using different concentration of 2-MNA. To align 2-MNA molecules in the electric field direction within polymer matrix, the films are poled for half an hour by contact electrode poling technique. Conductance and dissipation factor of films are measured at room temperature by Agilent Impedance Analyzer after poling the films. Wide frequency ranges varying from 100Hz to 10M Hz are kept for optimization. The effects of chromophoric group (2-MNA) concentration on the electrical conductance and dissipation factor is analyzed. The behavior of conductance and dielectric loss are optimized mathematically using FMINCON (a MATLAB tool) and multiobjective differential evolution algorithm (MODEA). To optimize the relation of conductance and dissipation factor with doping concentration of 2-MNA and applied frequency, the measured data is also modeled taking conductance and dissipation factor of films as dependent variable, which are affected by two independent variables namely frequency and dose of 2-MNA. The statistical validity and predictive capability of the obtained models is also checked by determining absolute average deviation and coefficients of determination.

Key words: 2-MNA, contact poling, polyether sulfone, guest host system, conductance, dissipation factor, multi objective optimization

1. INTRODUCTION

Since three decades, organic polymeric materials have been performing as promising candidates for researchers in advanced device applications (Dodabalapur, 2006). Organic polymers offer several advantages over inorganic material as they are clear plastic, weight less and have the possibilities to be molded in desired shapes, like spherical, aspheric and symmetric. These materials are utilized in various devices as both active and passive components. Researchers have described the wide use of organic polymeric materials in Lasers, LEDs, Solar Cells and TFTs. Bock et al. have shown the future of polymeric materials in (Bock, 2003). Significant research efforts have been directed to understand the engineering of these materials (Katz & Huang, 2009). From an optical point of view, the organic polymeric material possesses large non resonant electronic nonlinearities viz. second and third order nonlinearities. This is beneficial for all optical switches, signal processors and communication systems (Kuang et al., 2003). Also, these materials exhibits excellent transparency as well as optical damage threshold value (Rajasree et al., 1993). Organic materials can be structured in desired forms viz. thick and thin films, bulk crystals and multilayered films’ structures with necessary electrical and optical properties layer to layer (Avila-Niño et al., 2011; Piqué et al., 2003). Electrical and mechanical stability with temperature and humidity is the key charac-
teristics for practical application of these polymers in communication environment.

Organic conjugated polymers provide typical optical and electrical characteristics due to alternate presence of single and double bond in their backbone (Pron & Rannou, 2002). High energy orbital present in such type of configuration have loosely bounded electron corresponding to their atom. With the application of an external electric field, induced charge movement takes place within the material. There are two energy bands in the material (i) valance band which gives up the electrons and (ii) conduction band which attracts the electrons. Both energy bands are separated by the energy gap corresponding to their forbidden energy levels also known as the band gap. The distance between valance band and conduction band determines the characteristics of material either conducting, semiconducting or insulating (Moliton & Hiorns, 2004). The conduction mechanism is generated by the carrier movement (or jump) from valance band to conduction band which is faster if the band gap is small and slower if this band gap is large (Van Mullekom et al., 2001).

Considering the promising candidacy of organic polymers in various optoelectronics and microelectronics devices it becomes necessary to study the effect of external electric field on electrical properties of different composite films. Because of finding potential application in electronics, organic films are receiving unusual attention by researchers for last three decades (Klauck, 2006).

In the present study, we have developed polymeric film having different composition through guest - host system. We used poly ether sulfone (PES) as host material and 2-methyl-4-nitro aniline (2-MNA) as guest material. Although, the materials considered in the present study have been studied in the past but it has been done for the non linear optics. However, in this paper we have studied how molecular alignment of chromophoric group affects the electrical characteristics of the materials under consideration. In order to make a proper analysis we have considered different concentrations of chromophore (2-MNA). 2-MNA is incorporated in to PES as doping material from 2 to 18% by weight of PES with step size of 2. Thin and transparent films of developed composite are prepared by solution casting method. These films are poled by contact poling method under external field (6KV/cm) at 120°C temperature. The poled films are cooled down to room temperature by 4294A Agilent Impedance Analyzer over a wide frequency range (100Hz-10MHz). The electrical properties studied by Impedance Analyzer are conductance and dissipation factor considering as a function of frequency and 2-MNA concentration. It is observed that how these properties of films are affected by applying external DC electric field and by different doses of substituted aniline (2-mNA). A multi objective optimization model is considered. Multi Objective Differential Evolution Algorithm (MODEA) is employed for obtaining the solution to the mathematical model.

The paper is structured into 8 sections including introduction given in section 1. In the next section a description about different types of polymeric films is given. Preliminaries are presented in section 3. Preparation of polymeric films and their characterizations are briefly discussed in section 4. Optimization methods are given in section 5. Results and discussions are given in section 6. Mathematical model and multi objective optimization model of the problem are presented in section 7 and the paper finally in section 8 concludes the write up.

2. A SHORT NOTE ON ‘DIFFERENT TECHNIQUES FOR FABRICATING POLYMERIC THIN FILMS’

There are various fabricating processes which are frequently used to develop such type of noncentrosymmetric films. We are emphasizing on some techniques that are used for fabricating organic polymeric noncentrosymmetric films.

2.1. Langmuir Blodgett Films

This technique was first employed by Langmuir and Kathreen (Agarwal, 2008). The molecules which are hydrophobic at one end and hydrophilic at other end, known as linear amphiphiles molecules, are found most suitable for this technique. This method is applicable for both polymeric and monomeric systems. According to this process, monomolecular films of lipids are transfered layer by layer from water surface onto a solid substrate. The monolayers are structured containing required functions and molecular architectures. These monolayers are stacked to form desired simple or complex multilayer structures. These layers are then spread over a water surface and compressed for proper alignment of particles and thus the film deposits.
2.2. Covalent Self-Assembly

By this process, the treated surfaces of monolayers are brought in contact with such type of material that is capable to covalently bind to the surfaces. Highly organized multilayers structures can be constructed by this technique with mechanical, chemical, thermal and electrical stability (Decher & Schlenoff, 2003).

2.3. Ionically Self-Assembled Monolayer (ISAM) Thin Films

In this film forming process, ionic interactions and hydrogen bonding interactions are bounds the monolayers together. By this film forming process, material is adsorbed onto charged substrates in a self-limiting manner. Films fabricated by ISAM methods have enhanced optical properties.

2.4. Poled Polymers and Guest-Host Technique

Lastly we talk about poled polymers (Singer et al., 1986) and guest host system. In guest-host system NLO chromophoric material is doped as guest into an optically inactive host polymer. Guest-host material mixture is then deposited on solid surface. This system is known as guest-host system. Films fabricated by this method are easier to incorporate in devices and exhibit exceptional optical and nonlinear properties. In this paper we studied the characterization of films prepared by guest-host technique.

During poling process, NLO chromophores are aligned electrically. The films are heated to near their glass transition temperature. When these films are in rubbery state, films are subjected to a high electric field for a period of few minutes to several hours. So that chromophore or guest material is aligned to attain the necessary noncentrosymmetry. Then the heating is removed so that films are cooled to room temperature in the presence of the applied electric field. This step freezes the alignment of NLO molecules. Then electric field is removed, and below the glass transition temperature, restricted mobility of the polymers preserves the chromophore orientation even in the absence of the applied electric field. This system is known as ‘poled polymer’.

The degree of NLO molecule alignment (i) is directly proportional to ground state dipole moment ($\mu$) of the chromophore and applied electric field ($E$) (ii) is inversely proportional to Boltzmann’s constant ($k$) and poling temperature ($T$).

3. PRELIMINARIES

Every material has a unique set of electrical characteristics. Accurate information of these valuable properties can be beneficial for scientist and researchers to properly monitor the manufacturing process and incorporation the material in its intended application for more solid design.

Electrical conductance is how easily electric current flows along a certain path. It is the reciprocal of resistance. The conductance of an object is depends on its shape and size. Conductance ($G$) is directly proportional to the cross sectional area and inversely proportional to the length of given object. It can be represented as $G = \frac{A}{l}$.

If a field $V$ is applied across the material, the applied field accumulates the charges at the interface of material. If $d$ is the displacement of charges, the induced dipole moment is given by $\mu = Qd$, here $Q$ is the charges and is given as $Q = CV$. Where $C$ is the capacitance of material and can be written as $C = C_0 \epsilon/\epsilon_0$.

4. POLYMERIC FILM PREPARATION AND THEIR EXPERIMENTAL ANALYSIS

4.1. Materials Used to Make Films

Commercial grade polyethersulfone purchased from Solvay Chemicals is used as host material and optical grade 2-methyl4-nitroaniline (Merck) is used as guest material. The molecular structures of both materials are shown in figure 1. Polar solvent dimethylsulfoxide (Merck) is used to dissolve the materials.
4.2. Development of Polymeric Composite Films

Weighed amount of Polyethersulfone and 2-methyl-4-nitroaniline powder as reported in table 1 are dissolved in a polar solvent namely dimethylsulphoxide. The prepared solution is kept under constant stirring for three hours to make homogeneous solution. Resulted viscous solution is filtered to remove foreign particles. Viscous and homogenous solution is spread out on a glass plate with the help of glass rod having uniform thickness. Thus, the films prepared by solution casting method are dried in vacuum oven at 80°C temperature to evaporate the solvent.

![Fig. 1. Structure of Polyethersulfone (a) and 2-Methyl-4-Nitroaniline (b).](image)

Table 1. Composition of Polymeric Films.

<table>
<thead>
<tr>
<th>Film Sample No.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total solid weight – 1 gm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% of 2MNA</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td>14</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>Wt. of 2-MNA (g)</td>
<td>0.02</td>
<td>0.04</td>
<td>0.06</td>
<td>0.08</td>
<td>1.0</td>
<td>1.2</td>
<td>1.4</td>
<td>1.6</td>
<td>1.8</td>
</tr>
<tr>
<td>Wt. of PES (g)</td>
<td>0.98</td>
<td>0.96</td>
<td>0.94</td>
<td>0.92</td>
<td>0.90</td>
<td>0.88</td>
<td>0.86</td>
<td>0.84</td>
<td>0.82</td>
</tr>
</tbody>
</table>

4.3. Method of Contact Electrode Poling

The films are poled by contact electrode poling method (Dalton et al., 2009) to align the 2-MNA molecules in electric field direction. Poling is done for half an hour at 120°C temperature by applying DC electric field of 6KV/cm strength across the films. After half an hour poling the heating is switched off so that films are cooled down up to room temperature in the presence of DC electric field. The setup used to poled the films is shown in figure 2.

4.4. Electrical Characterization of Films

In recent years a trend has been developed to investigate the transport phenomena of charge carriers by Impedance Spectroscopy (Macdonald, 1992). Continuing in this direction, poled films are analyzed by 4294A Agilent Impedance Analyzer (Prabhakaran & Sullivan, 2002) at room temperature with OSC level 0.5 (V). The behaviors of conductance and dissipation factor are optimized in typical frequency range 100Hz-10M Hz as a function of frequency and 2-MNA concentration.

5. METHODS USED FOR OPTIMIZATION

5.1. Nonlinear Constrained Minimization (Fmincon)

The formulated optimization problem is solved using the MATLAB ‘fmincon’ optimization tool box (version 8.0). This function is based on local gradient method which implements sequential quadratic programming to obtain a local minimum of a con-
strained multivariable function. In ‘fmincon’ the nonlinear problem can be represented by using the syntax given as, \( x = fmincon(fun,x_0,lb,ub,options) \).

Where objective function ‘fun’ of a multidimensional design vector \( x \) is the scalar vector. \( x_0 \) stands for starting point (initial design point). ‘lb’ and ‘ub’ are the lower bounds and upper bounds of design variables.

5.2. Methods Used to Solve Multi-Objective Model/Multi Objective Differential Evolution Algorithm (MODEA)

MODEA was proposed by Ali et al. (2012) for solving MOOP and is a modified version of MDE. MODEA differs from MDE in the selection phase while the other operators viz. crossover and mutation are employed in the same manner as in MDE. While modifying DE for solving MOOP, selection plays a very crucial role because this phase decides a candidate solution for Pareto optimal front. Target solution is compared to the trial solution in MODEA, if target solution is found dominated then it replaces the target solution immediately in the form of current population and adds the advanced population into the target solution. In MODEA, current and advanced populations are combined after each generation. Thus total size is \( 2NP \) (where \( NP \) denotes the population size). The final population of size \( NP \) is determined with elite-preserving and using an explicit diversity-preserving strategy (Deb et al., 2002; Li & Zhang, 2009). MODEA facilitates convergence to the true Pareto front, the first goal of multi objective optimization.

6. RESULTS AND DISCUSSION


Films are characterized by Scanning Electron Microscopy and Atomic Force Microscopy before poling process (figure 3a and 4a, respectively) and after half an hour poling treatment (figure 3b and 4b, respectively). Figure 3 shows the SEM image of films doped with 18 wt. % 2-MNA. From figure 3b, it is observed that 2-MNA molecules aligned successfully in a centrosymmetric order by contact poling method. Image obtained from Atomic Force Microscopy has been shown in figure 4. Randomness of 2-mNA molecules into PES matrix can be seen in unpoled films (figure 4a) which divert as a uniform alignment of 2-MNA molecules after applying electric field (figure 4b). Effect of contact electrode poling on 2-MNA molecules can be seen clearly from figure 4b.

![Fig. 3. SEM Images of Films.](image1)

![Fig. 4. AFM Images of Films.](image2)
6.2. Studies from Fourier Transform Infrared Spectroscopy

The interaction between 2-MNA and PES molecules are recorded by Perkin Elmer FTIR Spectrophotometer and characteristics spectral bands for PES, 2-MNA and PES-2MNA composition are shown in figure 5. The characteristics FTIR spectra are obtained at room temperature over a spectral frequency range of 4000-500 cm\(^{-1}\). Infrared spectral band for PES corresponding to asymmetric stretching of C-O obtained at 1243 cm\(^{-1}\) which is observed at 1256 cm\(^{-1}\) in doped system. C-H stretching and symmetric stretching of S=O for PES is observed at 1578 cm\(^{-1}\) and 1150 cm\(^{-1}\) respectively. In doped films these characteristics vibrations observed at 1581 cm\(^{-1}\) and 1148 cm\(^{-1}\) respectively and confirms the presence of PES in doped films. In prepared films, characteristics vibration peaks of 2-MNA founds at 1485 cm\(^{-1}\), 2715 cm\(^{-1}\) and 3471 cm\(^{-1}\) corresponds to asymmetric stretching of NO\(_2\), CH\(_3\) and NH\(_2\) group respectively. Presences of N-H bending and N-H stretching modes are shown by peaks at 1633 cm\(^{-1}\) and 2945 cm\(^{-1}\) respectively. FTIR peaks of 2-MNA molecules correspond to symmetric modes of free NH\(_2\) group are obtained at 3464 cm\(^{-1}\) and 3363 cm\(^{-1}\). These characteristics modes show the presence of 2-MNA molecules in doped films. Obtained characteristics peaks of PES, 2-MNA and PES-2MNA composite are tabulated in table 2.

Table 2. FTIR modes of PES and 2-MNA.

<table>
<thead>
<tr>
<th>Types of vibrations</th>
<th>2-MNA</th>
<th>PES</th>
<th>PES-2MNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asymmetric stretching of C-O</td>
<td>-</td>
<td>1243</td>
<td>1256</td>
</tr>
<tr>
<td>C-H ring stretch</td>
<td>-</td>
<td>1578</td>
<td>1581</td>
</tr>
<tr>
<td>S=O symmetric stretch</td>
<td>-</td>
<td>1150</td>
<td>1148</td>
</tr>
<tr>
<td>Asymmetric stretching of NO(_2)</td>
<td>1509</td>
<td>-</td>
<td>1485</td>
</tr>
<tr>
<td>Asymmetric stretching of CH(_3)</td>
<td>2723</td>
<td>-</td>
<td>2715</td>
</tr>
<tr>
<td>Asymmetric stretching of NH(_2)</td>
<td>3472</td>
<td>-</td>
<td>3471</td>
</tr>
<tr>
<td>Symmetric stretching of NH(_2)</td>
<td>3364</td>
<td>-</td>
<td>3363</td>
</tr>
<tr>
<td>N-H Bending</td>
<td>1634</td>
<td>-</td>
<td>1633</td>
</tr>
<tr>
<td>N-H Stretching</td>
<td>2947</td>
<td>-</td>
<td>2945</td>
</tr>
</tbody>
</table>

6.3. Studies from Conductance

Conductance is the ability of a material to conduct electricity and it is the ratio of the current flowing through the material to the potential difference across it. Polymers which have conjugated double bonds yielding π-conjugation, exhibits good electronic transport properties due to facile delocalization of electrons. Figure 6 shows the variation of conductance with frequency and 2-MNA concentration. In low frequency region the periodic reversal of electric field is slow which causes the accumulation of ions and results the low value of conductance. It is observed that, in mid frequency region and high frequency region conductance obeys the universal power law, \(\sigma(\omega) \propto \omega^n\), (Bowen & Almond, 2006; Papathanassiou et al., 2007) so conductance increases sharply with frequency. Conductance are plotted with different concentration of 2-MNA, to analyse
the effect of 2-MNA on it. Conductance is observed to increase with 2-MNA content.

6.4. Studies from Dissipation Factor

Dissipation factor is the indication of power loss and is represented by \(\tan\delta\). In any dissipative system, dissipation factor is a measure of loss-rate of energy for one mode of oscillation. In other words, reciprocal of dissipation factor represent the quality of oscillation. It is defined as the ratio of the power loss in a dielectric material to the total power transmitted through the dielectric. Dissipation factor of poled films are shown in figure 7, as a function of frequency and doping concentration of 2-MNA. It can be seen that, dissipation factor increases gradually with frequency. At higher frequency region it shows a relation peaks which indicates the breakdowns of centrosymmetry of dipoles, afterward the value of dissipation factor (\(\tan\delta\)) tends to goes down. However, addition of chromophoric material (2-MNA) generates the power loss in form of thermal energy.

\[
Y_k = a_0^k + \sum_{i=1}^{2} a_i^k X_i + \sum_{i=1}^{2} d_i^k X_i^2 + \sum_{i=1}^{2} \sum_{j=1}^{2} a_{ij}^k X_i X_j
\]

(1)

where \(Y_k\) represents the \(k^{th}\) electrical property such that \(k=1, 2, 3\); \(a_0^k\) is a constant, \(a_i^k\), \(a_{ij}^k\) and \(a_{ij}^k\) are the linear, quadratic and interactive coefficients respectively. \(X_i\) and \(X_j\) are the levels of the independent variables. Two objective functions viz. \(F_1\) and \(F_2\) are constructed as follows:

\[
F_1 = \min \sum_{i=0}^{n} (y_{i1}^{fit} - y_{i1}^{obs})^2
\]

(2)

\[
F_2 = \min \sum_{i=0}^{n} (y_{i2}^{fit} - y_{i2}^{obs})^2
\]

(3)

where \(n\) indicates the total number of observations, \(y_{i1}^{fit}\) and \(y_{i1}^{obs}\), \(y_{i2}^{fit}\) and \(y_{i2}^{obs}\) are the fitted value and the observed value of the \(i^{th}\) observation for conductance and dissipation factor respectively. Since the behavior of the system is not known \(a\ priori\), therefore we have to check whether the model fits well to the experimental data. Absolute average deviation (AAD) is calculated for statistical validation

\[
\text{SSM} = \sum (\hat{y}_i - \bar{y})^2
\]

(4)

\[
\text{SSE} = \sum (y_i - \hat{y}_i)^2
\]

(5)

where \(y_i - \hat{y}_i\) observation, \(\hat{y}_i\) - predicted \(i^{th}\) observations, \(\bar{y}\) - mean of all observations.

The resulting value of the constants \(a_0^k\), and the coefficients \(a_i^k\), \(a_{ij}^k\) and \(a_{ij}^k\) corresponding to Conductance (\(Y_1\)) and Dissipation Factor (\(Y_2\)) are presented in table 3. In table 4, a comparison in terms of absolute average deviation (AAD) and coefficient of determination (COD) is given. The initial solutions are obtained using Fmincon. The result indicates that, Fmincon optimization technique produces satisfactory performance in terms of COD and AAD. The resultant equations for Conductance (\(Y_1\)) and dissipation factor (\(Y_2\)) are given below:

\[
Y_1 = 11.343 + 0.000432X_1 + 6.267X_2 + 2.97E-11X_1^2 + 2.70E-05X_1X_2 + 0.1474X_2^2
\]

\[
Y_2 = 5.32 + 8.314E-06X_1 + 0.9112X_2 - 8.112E-13X_1^2 - 1.72E-09X_1X_2 + 0.000227X_2^2
\]
Table 3. Coefficients of fitted polynomials for Conductance (\(Y_1\)) and Dissipation Factor (\(Y_2\)).

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>(a_0)</th>
<th>(a_1)</th>
<th>(a_2)</th>
<th>(a_{11})</th>
<th>(a_{12})</th>
<th>(a_{22})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value obtained by fmincon ((Y_1))</td>
<td>11.343</td>
<td>0.000432</td>
<td>6.267</td>
<td>2.97E-11</td>
<td>2.70E-05</td>
<td>0.1474</td>
</tr>
<tr>
<td>Value obtained by fmincon ((Y_2))</td>
<td>5.32</td>
<td>8.314E-06</td>
<td>0.9112</td>
<td>-8.112E-13</td>
<td>-1.72E-09</td>
<td>0.000227</td>
</tr>
</tbody>
</table>

Table 4. Value of \(R^2\) and Absolute Average Deviation (AAD) for the Both Responses.

<table>
<thead>
<tr>
<th>(R^2)</th>
<th>AAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Y_1)</td>
<td>0.970555</td>
</tr>
<tr>
<td>(Y_2)</td>
<td>1.75892</td>
</tr>
</tbody>
</table>

Table 5. Conductance Vs Dissipation Factor obtained by MODEA.

<table>
<thead>
<tr>
<th>Conductance (ns)</th>
<th>Dissipation Factor (mU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.09954768E15</td>
<td>1.61955113E9</td>
</tr>
<tr>
<td>-8.95935461E14</td>
<td>1.46182353E9</td>
</tr>
<tr>
<td>-5.49676439E14</td>
<td>1.1450998E9</td>
</tr>
<tr>
<td>-3.24608365E14</td>
<td>879943502</td>
</tr>
<tr>
<td>-3.8338988E13</td>
<td>302341089</td>
</tr>
<tr>
<td>-1.13400329E9</td>
<td>1644811.17</td>
</tr>
</tbody>
</table>

7.2. Regression Model

The independent variables, frequency and doping concentrations of 2-MNA gave a positive effect on both responses viz. conductance and dissipation factor. Responses and factors may not always have a linear relationship. It is well known that simultaneously changes in more than one factor produce different degree of responses. From developed regression model it is clear that interaction effect of \(X_1\) with itself is favorable for conductance while it is unfavorable for dissipation factor. Interaction effect of \(X_2\) with itself is favorable for all electrical characteristics. Significant values of \(R^2\) and AAD indicates the high predictive ability of the model (table 5).

7.3. Multi Objective Optimization Model

Consideration of multi objective optimization model is quite obvious because in real life scenarios we often have to consider several conflicting objectives simultaneously. An efficient solution methodology is the one that is able to provide a satisfactory solution to all the objectives while satisfying the constraints. The objectives considered here are maximization of Conductance and minimization of Dissipation Factor. Frequency and concentration are taken as box constraints. The three equations (1), (2) and (3) given in section 7, are taken together to construct an MOOP as: \(\text{Max} \{\text{Conductance} (Y_1)\}; \text{Min} \{\text{Dissipation Factor} (Y_2)\}; \text{St.} \ 100 \ \text{Hz} \leq x_1 \leq 10MHz; \ 2 \leq x_2 \leq 18; \) where \(x_1\) and \(x_2\) represent frequency and concentration, respectively.

7.4. Multi Objective Results

The Pareto front obtained by MODEA is depicted in figure 8. The graph indicates the relation between both parameters viz. Conductance and Dissipation Factor. From the figure we observe that MODEA is able to provide a smooth Pareto curve. This provides a trade-off between conductance and dissipation factor. In table 5 some suitable combinations of both parameters are given. This may be beneficial in electronics applications in real life scenarios, where the above mentioned parameters play a significant role.

8. CONCLUSIONS

In the present study an attempt is made to observe and analyze the performance of films in maximum possible aspects for electrical field. The work is conducted in three phases. In the first phase we
performed laboratory experiments for analyzing the electrical properties of thin polymeric films; in the second phase we developed a mathematical model in the form of MOOP and finally in the third phase we optimized the model through MODEA. The Pareto optimal front generated using this technique helps in providing a trade-off between two parameters viz. Conductance and Dissipation Factor. This may be very useful in electronics applications in real life scenarios, where the abovementioned parameters play a significant role.

From the electronics point of view, Conductance of PES-2-MNA composite films increases with frequency as frequency rises from lower to higher region. Behavior of conductance with frequency follows the universal power law for all composite films. Addition of 2-MNA in films increases the conductivity of films. Dissipation factor of all composite films increases with frequency from lower to higher range of frequency and doping of 2-MNA causes the more power loss in the films. The peak appears at a particular higher frequency range in loss tangent graph and shows the presence of relaxation of dipoles in all film samples. Considering the obtained characteristic it can be said that the composite films can be utilized in various microelectronics and optoelectronics application according to desired requirements.

REFERENCES


OPTYMALIZACJA WYDAJNOŚCI SPOLARYZOWANYCH ELEKTRYCZNIE POWŁOK POLIMEROWYCH

Streszczenie

W pracy analizowano układ organiczny gospodarz-góść, składający się z aniliny 2-metylo-4-nitro (2-MNA) będącej gościem oraz polieterosulfonu (PES) będącego gospodarzem. Przygotowano cienkie i transparentne próbki filmów dla różnej koncentracji aniliny 2-MNA. W celu ułożenia molekuł 2-MNA w kierunku

...
pola elektrycznego w matrycy polimeru, filmy pozostawały 
w polu elektrycznym przez pół godziny wykorzystując polaryzację 
elektrody. Przewodność elektryczną i czynnik rozproszenia 
powłok mierzono w temperaturze pokojowej analizatorem impe-
dancji Agilent po procesie polaryzacji. Do optymalizacji procesu 
polaryzacji stosowano szeroki zakres częstotliwości od 100 Hz do 
10 MHz. Analizowano wpływ koncentracji grupy chromoforowej 
(2_MNA) na przewodność elektryczną i czynnik rozproszenia. 
Zachowanie przewodności i straty dielektryczne było matema-
tycznie optymalizowane z wykorzystaniem FMINCON (narzędzia 
MATLABa) i wielokryterialnego algorytmu ewolucji różnicowej 
(MODEA). Do optymalizacji relacji przewodności elektrycznej 
i czynnika rozproszenia z ilością substancji 2-MNA i zastosowaną 
częstotliwością, dane pomiarowe również były modelowane, 
przyjmując przewodność i czynnik rozproszenia jako wielkości 
załężne, będące funkcją dwóch zmiennych niezależnych: częstot-
lliwości i ilości 2-MNA. Statystyczna istotność i możliwości 
przewidywania opracowanych modeli zostały także zweryfikowa-
nie poprzez określenie średniego odchylenia bezwzględnego 
i współczynników determinacji.

Received: October 1, 2014
Received in a revised form: November 18, 2014
Accepted: December 10, 2014
THE NUMERICAL ANALYSIS OF A TITANIUM SHEETS WELDING PROCESS AND WELDING JOINT TENSILE BEHAVIOR

KONRAD ADAMUS*, PIOTR LACKI

Częstochowa University of Technology, ul. Akademicka 3, 42 200 Częstochowa, Poland
*Corresponding author: konrad.adamus@gmail.com

Abstract

The paper analyses joining of titanium Grade 2 and Grade 5 sheets using electron beam welding, EBW, technology. The joined sheets undergo further processing to produce final aircraft component.

EBW is a fusion welding process that utilizes electrons to produce heat and join materials. It is characterized by low heat input, small distortions, precise and repeatable welding parameters and vacuum operating conditions. Titanium is a unique material that offers high specific strength, heat resistance and corrosion resistance. In aerospace applications it is used for components of rotors, engines and airframes. Tailor Welding Blanks is a technology of joining sheets of different properties in order to obtain components that have mixed properties. Titanium Grade 2 has better formability while titanium Grade 5 has higher strength. The welded sheets are further formed in order to obtain final shape.

The numerical simulation of the welding process investigates the impact of thermal load produced by a moving electron beam on the deformations occurring in the joined sheets. The model is based on finite element method and it takes into account elastic, plastic and thermal strains. The analyzed sheets have thickness of 0.8 mm, their small cross-section contributes to significant post-welding bending both in transverse and longitudinal directions. The paper analyzes the influence of mesh structure and welding gap on the degree of deformation. The obtained numerical results were compared with actual welding deformations. The research into modeling of titanium sheets welding was extended with analysis of joint behavior during tensile tests. Specimen was cut out from the joint and was subject to uniaxial tensile test. The strains occurring during tension were recorded using optical non contact measurement system that uses two cameras. The experimentally measured strains were compared with numerical results from a separate finite element method model.

Key words: electron beam welding, welding distortions, welding gap, thin sheets, finite element method

1. INTRODUCTION

In the work a finite element method model of thin sheets electron beam welding was presented. The problem of welding thin sheets was selected as it produces high distortions in the material and can validate the welding model’s accuracy. According to Lindgren (2007) accurate predicting of post-welding deformations is more difficult than predicting residual stresses. A number of factors impact on calculated deformations and residual stresses, among others: welding gap, mesh structure and weld bead geometry.

Interface elements can be used for modeling of welding gap between sheets (Wang et al., 2013; Murakawa et al., 2012; Deng et al., 2012). Interface element technique consist in applying nonlinear spring elements joining the welded sheets. The definition of relationship between displacement and bonding force allows for taking into account bonding strength, gaps and misalignments occurring at different stages of a welding process.

Moein and Sattari-Far (2014) investigated the application of different element techniques to modeling of a welding gap and their impact on magnitude of calculated residual stresses in a simulation of
aluminum welding process. Three cases were analyzed: welded sheets were simulated as a single block of material, birth element option was used to represent moving welding pool, interaction technique was used to describe the surface-to-surface thermal and structural contact between welded sheets. The study showed that element birth technique yielded most accurate results for residual stresses. Arai (2010) analyzed the case of welding thin sheets fixed with clamping device and noted that there are significant discrepancies in calculated deformations between models taking into account and not taking into account a welding gap. Shan et al. (2007) compared element birth technique with sheets as single block technique. It was found out that element birth technique produces more accurate temperature field. Koch et al. (2013) modelled welding of curved plates. One of the plates was at the top of the another. In order to simulate the air gap between the overlapping plates the conductive and radiative heat exchange between the gap surfaces was taken into account.

2. EXPERIMENT

During experimental research two titanium sheets of different grades were joined using electron beam welding, EBW, technology. EBW is a fusion welding technology that utilizes a beam of electrons to melt metals which after crystallization form a joint. Extensive description of EBW can be found in (Schultz, 2000). During welding the sheets were fixed using a clamping device. Both sheets had the same dimensions: 0.8 x 130 x 125 mm.

One of the sheets was made of titanium Grade 2 that is characterized by good formability and the other from titanium Grade 5 that is characterized by high strength. The purpose of joining titanium sheets of different grades is to produce Tailor Welded Blanks, TWB, that will undergo further forming operations to produce final aircraft component. Application of titanium to TWB was described in (Winowiecka et al., 2013; Adamus & Lacki, 2014).

Schenk et al. (2009) analyzed the dependency of distortion amplitude on finite element mesh for welding simulations. The authors showed that adequate prediction of distortions requires fine mesh resolution everywhere where bending occurs and it is not enough to use high element densities only where high stress and temperature gradients occur. Two kinds of mesh were investigated: coarsening towards sheet sides (T-meshes) and non coarsening towards sheets sides (L-meshes). T-meshes don’t produce acceptable results and the best results were obtained for L-meshes.

According to (Sun et al., 2014; Matsuoka et al., 2013) accurate description of the heat source is essential to the simulation of thin-sheet welded structures and has significant impact on the final deformation of thin sheets.

The welding process introduces deformations into the sheets: transverse bending and longitudinal bending. Figure 1 presents the character of sheet distortions and distortions measures. Transverse bending is measured by butterfly angle $\alpha$, longitudinal bending is measured by displacement vector of the end of a welding trajectory.

3. THEORY AND CALCULATIONS

Computational welding mechanics models usually fall into one of the two categories: computational fluid mechanics, CFD, or computational solid mechanics, CSM. CFD models describe fluid motion inside weld pool and aim at predicting fusion zone, FZ, shape based on welding parameters (Piekarska et al., 2010; Sloma et al., 2011). CSM models pre-
dict stresses and displacements occurring as a result of thermal load (Lacki et al., 2014).

This study uses CSM model to describe deformations of the welded sheets. Thermo-mechanical coupled analysis was applied. The model was built using ADINA program. Finite element procedures used in ADINA can be found in (Bathe, 2006).

Transient heat transfer was described by Fourier-Kirchoff equation:

$$\frac{\partial T}{\partial t} = a \nabla^2 T + \frac{q_v}{\rho c_p}$$  \hspace{1cm} (1)

where:  
- $a$ – the thermal diffusivity, 
- $\rho$ – the density, 
- $c_p$ – the specific heat, 
- $q_v$ – the efficiency of internal volume heat source.

The equation defining thermo-elasto-plastic material has the following form:

$$\sigma_{ij} = c_{ijrs}(\varepsilon_{rs} - \varepsilon_{Prs} - \varepsilon_{Thrs})$$  \hspace{1cm} (2)

where: 
- $\sigma_{ij}$ – stress tensor at time $t$, 
- $c_{ijrs}$ – elasticity tensor at temperature corresponding to time $t$, 
- $\varepsilon_{rs}$ – total strain tensor at time $t$, 
- $\varepsilon_{Prs}$ – time independent plastic strain tensor at time $t$, 
- $\varepsilon_{Thrs}$ – thermal strain tensor at time $t$.

Plastic strains were calculated using von Mises model with isotropic hardening. Thermal strains were calculated based on temperatures in nodes and thermal expansion coefficient. Relationship between stresses and strains was described using bilinear model. The computational model takes into account dependency of physical properties on temperature. The detailed information about applied values of material properties can be found in (Adamus et al., 2013).

The welding model uses combined surface and volumetric heat source model. The outline of the heat source is presented in figure 2.

White elements denote sheets and light gray elements denote the volume of material that is filling the initial welding gap. For clarity, the part of mesh corresponding to the sheet in direction of positive values on X axis was removed. The black dot denotes the current position of the heat source. As the heat source moves along the welding trajectory (Y axis) the consecutive elements fill the gap between sheets. The occurrence of the welding gap elements is implemented using element birth option. If this options is set elements will become present in the simulation after simulation time exceeds their birth time. Surface heat source in the shape of a disk was denoted with striped pattern. Below the disk heat source there is volumetric heat source consisting of a single element denoted with dark gray color. Both heat sources move together, along the welding trajectory, with step equal to the length of element edge in Y axis direction. During calculation 90% of power was assigned to the surface heat source and 10% of power was assigned to volumetric heat source.

Figure 3 presents the comparison of FZ calculated by the welding model and the macrostructure of the actual FZ. The macrostructure of the specimen cross-section shows that even though the sheet has small thickness, there is significant change of FZ geometry with increasing depth.

Figure 4a presents the outline of the welding model geometry. Gray planes show the interface between welded sheets and clamping device. The clamping device was modeled using spring elements with high stiffness attached to appropriate nodes. Figures 4b and 4c present fragments of mesh structures used in the study.
During research 5 different models were analyzed:
- Model A: the number of elements in direction of Z and Y axes was constant, figure 4b. Element birth option was used to simulate initial welding gap.
- Model B: the number of elements in direction of Y axis was decreasing and in direction of Z axis was constant. Element birth option was used to simulate initial welding gap.
- Model C: the number of elements in direction of Y axis was constant and in direction of Z axis was decreasing. Element birth option was used to simulate initial welding gap.
- Model D: the number of elements in direction of Y and Z axes was decreasing, figure 4c. Element birth option was used to simulate initial welding gap.
- Model E: the number of elements in direction of Z and Y axes was constant, figure 4b. There was no initial welding gap and sheets were treated as single block of material.

4. RESULTS

Table 1 presents the summary of calculated deformations for model A-E. A discrepancy is defined as:

$$ d = \frac{v_c - v_a}{v_a} $$

where: $v_a$ - actual value, $v_c$ - calculated value. It can be seen that best results were obtained for model A that has constant number of elements in directions of Y and Z axes and utilizes element birth option to represent initial welding gap. The reduction of the element count in Y direction causes significant increase in the discrepancy of longitudinal bending, from 18 to 50 %. The reduction of the element count in Z direction increases the discrepancy of longitudinal bending to a lower extent, from 18 to 30 %. For models A-D the worst results were obtained for model reducing element count in Y and Z directions. The analysis of results obtained for model E, which has the same mesh as model A, shows that the lack of initial gap in the welding model contributes to the greater discrepancy than the reduction of mesh density, both for transverse and longitudinal bending.
Table 1. The discrepancy between actual and calculated bending for different mesh structures.

<table>
<thead>
<tr>
<th>Welding model</th>
<th>Transverse bending discrepancy, %</th>
<th>Longitudinal bending discrepancy, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>A no mesh reduction along Y and Z axes, element birth</td>
<td>-18.5</td>
<td>-23.7</td>
</tr>
<tr>
<td>B mesh reduction along Y axis, element birth</td>
<td>-18.7</td>
<td>-50.4</td>
</tr>
<tr>
<td>C mesh reduction along Z axis, element birth</td>
<td>-17.7</td>
<td>-29.2</td>
</tr>
<tr>
<td>D mesh reduction along Y and Z axes, element birth</td>
<td>-24.3</td>
<td>-53.0</td>
</tr>
<tr>
<td>E no mesh reduction along Y and Z axes, sheets as single block of material</td>
<td>-77.7</td>
<td>-57.7</td>
</tr>
</tbody>
</table>

Figure 5 presents the distribution of displacement along Z axis i.e. in the direction of sheet thickness. Z-displacement equal to 0 corresponds to the initial position of the sheet top surfaces. Part (a) corresponds to model A that takes into account the initial welding gap, part (b) corresponds to model E that treats sheets as single block of material. The character of deformation in both cases is similar. There is convex longitudinal bending and concave transverse bending. However, there is large difference in the minimal value of Z-displacement, about 4.1 mm for model A and about 0.8 mm for model E. For model A there is distinct shift in the displacement values along welding trajectory. The shift occurs to a smaller extent also for model E. In case of model E the shift corresponds to slightly concave weld face. In case of model A the additional factor is the difference in displacement between Grade 2 and Grade 5 sheet edges. The shift increases toward the end of the welding trajectory, corresponding to higher Y axis values. The shift occurred also in the actual welded sheets, as difference between sheet edges near the welding trajectory end was about 25% of the sheet thickness.

Figure 5 presents the distribution of logarithmic principal strain during uniaxial tensile tests in specimen cut out from welded sheets. The specimen in the shape of a strip was cut out in the direction perpendicular to the welding trajectory. Part (a) corresponds to results obtained from numerical model. During calculations the same material model was applied as in case of the welding simulation. Part (b) corresponds to experimental measurements at time before the onset of the crack. The measurements were performed using optical system Aramis. The system uses two cameras to track the motion of stochastic pattern, in the form of black dots with white background, spray painted at the sheets’ external surface, figure 6c. During numerical calculations of uniaxial tension necking occurred in the specimen at the side of titanium Grade 2 sheets, which has lower strength. The strains at the side of titanium Grade 5 were minimal. In the case of experimental measurements the necking also occurred at the side of titanium.
Grade 2 sheet. There is similarity between the strain fields calculated by the numerical model and measured by the optical system.

5. CONCLUSIONS

In the work thermo-mechanical model of thin sheets welding process was described. The goal of the model was to predict deformations occurring due to the welding process. Different finite element meshes and welding gap models were analyzed. Based on the performed analysis the following conclusions can be drawn:

- Computational welding models predicting distortions of thin sheets are very sensitive to mesh structure and welding gap model.
- The best results with regard to accuracy of transverse and longitudinal bending prediction were achieved for meshes that avoid reduction of element count in directions corresponding to sheet thickness and welding trajectory, in the volume of sheets away from welding line.
- The discrepancy in the calculated longitudinal bending was much more significantly affected by element reduction than the discrepancy in the calculated transverse bending.
- The computational model treating sheets as a single block of material gave significantly larger discrepancies in longitudinal and transverse bending than the computational model using element birth option and taking into account the initial welding gap.

ACKNOWLEDGEMENTS

Financial support of Structural Funds in the Operational Programme - Innovative Economy (IE OP) financed from the European Regional Development Fund - Project "Modern material technologies in aerospace industry", Nr POIG.01.01.02-00-015/08-00 is gratefully acknowledged.

REFERENCES


Schultz, H., 2000, Elektronenstrahlschweifen, DVS Verlag, Düsseldorf.
ANALIZA NUMERYCZNA PROCESU SPAWANIA BLACH TYTANOWYCH I ZACHOWANIA SPOINY PODCZAS ROZCIĄGANIA

Streszczenie

W pracy badano zagadnienie spawania blach tytanowych Grade 2 i Grade 5 za pomocą wiązki elektronów. Zespawane blachy stanowią wad do tłoczenia, w wyniku którego powstaje finalny komponent samolotu.

W procesie spawania wiązka elektronów jest wykorzystywana do stopienia i łączenia materiałów. Spawanie wiązki elektronów charakteryzuje się małą ilością wprowadzanego ciepła, małymi deformacjami, dużą precyzją i powtarzalnością ustawienia parametrów oraz zastosowaniem komory próżniowej. Tytan jest materiałem o wysokim stosunku wytrzymałości do masy, dużej odporności na podwyższone temperatury i środowisko korozjnym. W lotnictwie tytan znajduje zastosowanie w produkcji wirników, silników oraz kadłubów. Tłoczenie wad spawanych z różnych materiałów pozwala na uzyskanie komponentów z mieszanych właściwościach składowych materiałów. Tytan Grade 2 zapewnia dobrą tłoczność, natomiast tytan Grade 5 oferuje wysoką wytrzymałość.

Analiza numeryczna procesu spawania koncentrowała się na badaniu wpływu obciążenia termicznego wywołanego przez poruszającą się wiązkę elektronów na deformacje pojawiające się w łączonych blachach. Analizowane blachy miały grubość 0,8 mm. Ich mały przekrój umożliwia powstawanie znacznych deformacji spawalniczych zarówno w kierunku poprzecznym jak i podłużnym. Jako miarę deformacji przyjęto kąt ugięcia poprzecznego oraz przemieszczenie końca linii spawania. Model numeryczny wykorzystywał Metodę Elementów Skośnych. W symulacji uwzględniono odkształcenia sprężyste, plastyczne oraz termiczne. Do symulacji oddziaływania wiązki elektronów wykorzystano połączone powierzchniowe i objętościowe źródło ciepła. W pracy przedstawiono wpływ struktury siaki elementów skośnych oraz modelu odstępu pomiędzy łączonymi blachami na stopień ugięcia poprzecznego i podłużnego. Uwzględniono siatki o jednorodnej i zmiennej liczbie elementów w kierunku grubości blach i w kierunku linii spawania. Porównano model, w którym blachy stanowiły pojedynczy blok materiału, z modelem, w którym do symulacji odstępu wykorzystano opcję narodzin elementu. Uzyskane wyniki zostały porównane z rzeczywistymi deformacjami zespawanych blach. Analiza procesu spawania została rozszerzona o model opisujący zachowanie próbki wyciętej z połączonych blach podczas statycznej próby rozciągania. Odkształcenia powstające w wyniku rozciągania zostały zmierzone za pomocą systemu optycznego wyposażonego w dwie kamery. Porównano odkształcenia obliczone numerycznie i zmierzone eksperymentalnie.

Received: September 26, 2014
Received in a revised form: November 26, 2014
Accepted: November 29, 2014
NUMERICAL INVESTIGATION OF REFILL FRICTION STIR SPOT WELDING JOINTS

ANNA DERLATKA*, PIOTR LACKI
Częstochowa University of Technology, Dąbrowskiego 69, 42-202 Częstochowa, Poland
*Corresponding author: aderlatka@bud.pcz.czest.pl

Abstract
The paper presents an analysis of tensile tests for welded specimens made of 6061-T6 aluminium alloy. Three kinds of lap joins were made by Refill Friction Spot Stir Welding (RFSSW). The each specimen has one joint, but they were varied in the position of the sheets. In the first, the angle between sheets axes was 0°, to determine the tensile capacity of the joint. In the second and third, the angle between sheets axes was 20°, and -20°. That position of sheets allow determine the maximum load and displacement of stretched and twisted structure. The numerical calculations were performed using the ADINA System based on the Finite Element Method (FEM). The sheets and joints were modelled with 3D-solid elements. The experimental investigations were carried out using a testing machine and a non-contact and material independent measuring system providing, for loaded test objects, accurate 3D displacements and surface strain values. The stress, strain distribution and displacements were analysed. The numerical and experimental results were compared. The structures were assessed in respect of strength and the possibility of applying in the aircraft industry.

Key words: finite element method, aluminium 6061-T6, Refill Friction Spot Stir Welding (RFSSW)

1. INTRODUCTION

Friction Stir Welding (FSW) is used in automobile and aircraft structures in order to reduce weight and improve performance in relative to the rivets. The process is characterized by reliability and low time-consuming preparation of materials for welding and making the joint. FSW competes with traditional linear welding methods such as TIG-welding (Lacki & Derlatka, 2013; Lacki et al., 2012) Refill Friction Stir Spot Welding (RFSSW) and other methods (e.g. riveting and hybrid joints presented by Sadowski et al. (2013), Sadowski and Golewski (2013), Sadowski et al. (2011) belong to a group of modern spot joining techniques.

RFSSW is used for welding aluminium alloys i.e. AA 6061-T6 (Venukumar et al., 2013), AA2024-T3 and AA5754-H22 analysed by Bozkurt and Bilici (2013) because of the good quality of the joints, especially in comparison to traditional welding techniques. It is also possible to join other materials e.g. titanium (Kudla et al., 2009), steel (Sun et al., 2014; Ghosh et al., 2011) or steel with aluminium alloy (Kundu et al., 2013). The continuous and smooth face of the RFSSW joint is formed by a specific tool consisting of a probe, sleeve and clamping ring (figure 1). In Mishra and Mahoney (2007) is presented,
that the sleeve and probe are the moving parts which can rotate and protrude for heating and plasticization of the base material. Yuan et al. (2011) present the analyze of the spots making by two tools: a conventional tool with a center pin and an off-center feature tool.

**Table 1. Chemical composition of 6061 aluminium alloy.**

<table>
<thead>
<tr>
<th>Element</th>
<th>% wt.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cr</td>
<td>0.04-0.35</td>
</tr>
<tr>
<td>Cu</td>
<td>0.15-0.40</td>
</tr>
<tr>
<td>Fe</td>
<td>0.70</td>
</tr>
<tr>
<td>Mg</td>
<td>0.80-1.20</td>
</tr>
<tr>
<td>Mn</td>
<td>0.15</td>
</tr>
<tr>
<td>Si</td>
<td>0.40-0.80</td>
</tr>
<tr>
<td>Ti</td>
<td>0.15</td>
</tr>
<tr>
<td>Zn</td>
<td>0.25</td>
</tr>
<tr>
<td>other</td>
<td>0.15</td>
</tr>
<tr>
<td>Al</td>
<td>rest</td>
</tr>
</tbody>
</table>

The thermo-mechanical models are used to analyze the welding process. Thermal analyses use a heat source distributed over the tool surface, with a heat flux per unit area. In another approach, the heat input may be distributed over the probe volume. The consequence of the effect of metal flow on the distribution of heat generation is ignored. Assidi et al. (2010) simulate FSW for the Al 6061 aluminum using an Arbitrary Lagrangian Eulerian (ALE) numerical model and a Eulerian formulation, which provides a faster way to reach a steady welding state. To calculate the plastic energy and to account for the change in the yield strength with welding temperature, the model based on welding energy is proposed by Hamilton et al. (2009). A fully coupled thermo-mechanical model presented by Zhang and Zhang (2009) is adopted to analyse the effect of welding parameters: the rotating speed, the welding speed and material flow patterns. To analyse the structures made using RFSSW, the mechanical model is used. Derlatka et al. (2014) present the comparison of three types of model with different kinds of finite elements: full shell elements model, full 3D-solid elements model and 3D-solid and shell elements model for analysis the structures during tensile testing. The full 3D-solid model is the most exact, but it requires a long total solution time. The alternative for shorter computational and modelling times are 3D-shell elements. In Fanelli et al. (2012) the RFSSW joints and heat affected zone are modelled as 3D-sold elements and base material as shell elements. Derlatka and Kasza (2014a, 2014b) present the analysis of bending beams. Complex aluminium structures are made from sheets welded with spot joints. The joints are meshed by 3D-sold elements and sheets by shell elements.

**2. GOAL AND SCOPE OF WORK**

The analysis of the tensile test results for different configurations of sheets in RFSSW joints is the goal of the paper. The structures were assessed according to tensile strength, strains and displacements. Two sheets of 6061-T6 aluminium alloy are welded as lap joints by RFSSW. The chemical composition of the base material is given in table 1. Three variants of sheet position relative to each other are analysed – figure 2. In each model, the same finite elements are used: 3D-solid for the RFSSW joint and shell elements for the sheets. The numerical results are compared with experimental investigations. The tensile tests were carried out...
using a testing machine and a non-contact optical 3D deformation system, Aramis, which enables measuring displacements and strains.

The RFSSW processes were carried out using the following parameters: joint thickness 2.2 mm, tool rotational speed 2000 rpm, tool input speed 0.7 s, tool output speed 0.5 s.

3. EXPERIMENTAL INVESTIGATION

Figure 3 shows the analyzed samples. Figure 4 shows the force - displacement diagram during tensile tests. The maximum force for the sample with the angle between sheets axes of 0º is 9.8 kN, for the sample with the angle between sheets axes of 20º is 9.2 kN, while for the sample with the angle between sheets axes of -20º is 7.8 kN. The displacements are respectively 1.36 mm, 2.70 mm and 3.58 mm.

The strain distributions for the analysed samples, during peak force, are presented in figure 5. In each sample, the area of maximum values concentration is located on the edge of the joint and base material. In the first sample, the distributions are symmetrical with respect to the vertical axis (Y-axis). In the samples for the angle between sheets axes of 20º and -20º, the distributions reflect the shape of a screw. The same dependencies are observed in the joint shapes after breaking (figure 6).
4. COMPUTATIONAL MODELS

The ADINA System based on the Finite Element Method was used for the numerical studies. The geometry of the models is identical to the real samples. The model for the samples with the angle between sheets of 1600 do not include the differences in the direction of welding. The sheets and joints are modelled as 8-node 3D-solid elements. Between the 3D-solid element surfaces of the sheets, contact conditions are assumed. The sheet and joint elements are connected in nodes.

The boundary conditions and type of load are the same for each sample and they are added to the surfaces. All the degrees of freedom on the first external edge of the sample are fixed, the second external edge has one free degree of freedom (X-translation). This edge is loaded by the displacement of the experimental value. The meshes of the samples are shown in figures 7, 8.

Plastic orthotropic material with the following parameters: modulus of elasticity 68.9 GPa, yield strength 276 MPa, Poisson's ratio 0.33 and density 2700 kg/m$^3$, is used for each model.

The model with the angle between sheets of 1800 has 11648 3D solid elements and 15041 nodes. The model with the angle between sheets of 1600 has 9704 finite elements and 12185 nodes. The calculations are performed in 100 time steps.

5. RESULTS

The effective plastic strains for the sample with the angle between sheets axes of 0º are presented in figure 9. Figure 10 shows the effective plastic strains for the sample with the angle between sheets axes of 20º. The results for the sample with the angle be-
between sheets axes of -20º are not shown. The strain distributions are similar in the sample with the angle between sheets axes of 20º. Only the extreme results are different.

In the sample with the angle between sheets axes of 0º, the effective plastic strain distributions are symmetrical with respect to the X-axis and similar in both sheets. The extreme results are located on the inner sides of the sheets, which is difficult to specifically identify. The maximums are on the fixed part of the sheet, the minimums are on the free part of the sheet. The free parts of the sheets are deflected with respect to the Y-axis.

In the sample with the angle between sheets axes of 0º the strain distributions are not symmetrical, but they are similar in both sheets. The extremes are on the area surrounding the RFSSW joint. The precise location of the extreme results is difficult, because extremes are also located on the inner sides of the sheets and the sheets are asymmetrically deflected.

6. DISCUSSION

Based on the displacement-force diagram, the highest breaking force is for the sample with the angle between sheets axes of 0º. The shape of the curve suggests the sample strains during pure tensile. The samples with the angle between sheets axes of 20º are stretched and twisted. Therefore, the curves have a complex shape and the displacements reach higher values.

The structures behaviour affect the places of strain concentration. The numerical studies show that in each sample the maximum value of strains are on the inner part of the joint. In the tensile test, the samples fracture in the RFSSW joint but they do not separate from the sheets.

![Fig. 10. Effective plastic strain for sample with angle between sheets axes of 20º: X-Z view (a), axonometric view, mag. 5x (b-c).](image)

7. CONCLUSION

The analysis of RFSSW joints with a different arrangement of sheets was conducted. The following conclusions were made:

- The strength of the sample with the angle between sheets axes of 0º is higher than the sample with the angle between sheets axes of 20º and -20º.
- In each sample, the maximum strains are observed on the joint area, on the inner surfaces of the sheets. The concentration of strains on the outer surface are observed on the edge of the joint.
- In each case, the fractures are in the joint. The spot welds do not separate from the sheets.
- The sheets of the sample with the angle between axes of two parts of 0º are deflected with respect to the Y-axis. The sheets of the sample with the angle between axes of two parts of 20º are deflected with respect to the Y-axis and they are twisted.
ACKNOWLEDGMENT

Financial support of Structural Funds in the Operational Programme - Innovative Economy (IE OP) financed from the European Regional Development Fund - Project "Modern material technologies in aerospace industry", Nr POIG.01.01.02-00-015/08-00 is gratefully acknowledged.

REFERENCES


ANALIZA NUMERYCZNA PUNKTOWYCH POŁĄCZEŃ ZGRZEWANYCH TARCIOWO Z MIESZANIEM MATERIALU

Streszczenie


Received: September 30, 2014
Received in a revised form: October 27, 2014
Accepted: December 21, 2014
THE MODELLING OF RING TESTS AT ELEVATED TEMPERATURES FOR THE DETERMINATION OF FRICTION IN Ti-6Al-4V FORGINGS

JAMES D. POLLARD1*, ANDREW WATFORD2, MARTIN JACKSON3, BRADLEY P. WYNNE3

1 EPSRC Centre for Doctoral Training in Advanced Metallic Systems, Dept. of Materials Science and Engineering, The University of Sheffield, Mappin Street, Sheffield S1 3JD, UK.
2 RTI Extrusions Europe, Ltd., Brighouse Road, Low Moor, Bradford, BD120QL, UK.
3 Dept. of Materials Science and Engineering, The University of Sheffield, Mappin Street, Sheffield S1 3JD, UK
*Corresponding author: jpollard1@sheffield.ac.uk

Abstract

Ring compression tests and finite element modelling were used to explore the friction conditions present in high temperature Ti-6Al-4V forgings where glass is used as the lubricant. The work explored the use of isothermal and non-isothermal simulations as a means of modelling non-isothermal test conditions. The friction factor is determined by comparison of the deformation of the internal diameter of the experimental compression rings and the simulated compression rings. It was determined that the heat transfer coefficients (HTCs) used in the simulations have a significant result on the friction factor predicted by the simulation results. It was found that it is possible to predict similar deformations through combinations of low HTC/high friction factor and high HTC/low friction factor. Consequently, it is considered critical that the heat transfer conditions for non-isothermal work where there is a high temperature gradient between workpiece and its surroundings be correctly modelled in order to determine the correct friction factor to be used in later simulations.

Key words: Ti-6Al-4V, friction, heat transfer, finite element modelling, thermomechanical processing

1. INTRODUCTION

The friction between the tool and workpiece in a hot metal forming process can have a significant influence on the applied loads required for shape change and on the strains, strain rates and stress distributions within the workpiece and thus its final microstructure and mechanical properties. For numerical simulations of metal forming a number of different friction models can be applied, for example the Coulomb friction law and the shear friction law. Regardless of which model is used it is necessary to identify the correct value for the relevant frictional values to accurately simulate the interaction at the workpiece/tool interface. A common method of experimentally determining these friction values is the ring compression test (Male & Cockroft, 1964). This method relies on comparison of the dimensions of experimentally deformed ring specimens with deformed dimensions of rings predicted by analytical techniques; traditionally this was accomplished via slip-line field theory, lower and upper boundary methods or slab theory however this has been superseded by finite element methods (FEM) as noted elsewhere (Wu et al., 2003; Zhu et al., 2011). In this work we examine one of the key issues that is often overlooked when performing FEM analysis of the ring compression test: the heat transfer between the
tool and the workpiece, particularly when there is a significant temperature difference between them.

Investigations into the impact of the heat-transfer coefficient (HTC) on non-isothermal ring compression tests by Andersson et al. (1996) using aluminium alloys suggests HTC has a definite effect on metal flow and on the results of the compression test but the overall effect is small compared to inaccuracies in the calibration curves and the effects of varying friction. However, Zhu et al. (2011) using Ti-6Al-4V and glass lubricants with a sample temperature of 940°C and tooling at 220°C found that the friction calibration curves generated by FEM using DEFORM 3D varied significantly depending on the HTC values. The authors conclude that an increase in the HTC value produces an increased metal flow velocity and that the resulting deformation is similar to that produced by assuming an increased friction factor (Zhu et al., 2011). The work conducted in this paper has similarities to that done by Zhu et al. (2011) in terms of the use of Ti-6Al-4V at elevated temperatures in conjunction with tooling at lower temperatures, however the temperature is higher at 1100°C such that deformation is undertaken in the β phase field where the material is considerably softer and any cooling of the specimens during testing may result in a microstructural gradient due to phase changes.

2. METHODOLOGY OF EXPERIMENTAL WORK AND FE SIMULATIONS

2.1. Experimental Procedure

Rings of Ti-6Al-4V (β transus ~1005°C) were prepared with an outer diameter of 19.05 mm, inner diameter 9.525 mm and a height of 6.35 mm giving a ratio of 6:3:2. Two different glass lubricants were tested; one with a borosilicate chemistry and one with a silica based chemistry. The borosilicate glass was in powder form, the silica glass was tested as both a powder and in the form of glass fibres. The specimens were coated prior to testing by means of heating them to 1000°C for 90 seconds and then immersing them in a container of the appropriate glass powder, this resulted in the powders coating all surfaces of the rings. To apply the glass fibres, the fibres were teased out into flat mats and the heated rings were placed on top of one mat with a second mat placed on top of them; a ceramic heat tile was placed on top of the second mat to ensure there was contact between the surfaces of the rings and the fibres. Once cooled, the tile was removed and any excess fibres were trimmed from the edges of the rings. This method of applying the fibres produced coatings on the flat surfaces of the rings and left the inner and outer diameter surfaces uncoated, the results are shown in figure 1.

![Fig. 1. Ring specimen with glass fibres applied.](image)

Compression testing was performed using the University of Sheffield’s Thermo-Mechanical Treatment Simulator (TMTS). This set up allowed a programmed sequence to control the entire experiment from heating to compression to post-compression quenching. The coated rings were placed in the machine’s sample holding apparatus, an induction furnace was used to heat the specimens to 1100°C and hold them at temperature for approximately 30 s; temperature was monitored throughout testing by the use of thermocouples embedded in each specimen. Once the temperature of the specimens was stable they were inserted directly into the test furnace (held at 200°C) and compressed at a strain-rate of 1 s⁻¹. The height reductions for each batch of specimens are given in table 1. Initial tests were conducted using specimens coated with a boron nitride aerosol; these tests served to verify experimental procedure prior to testing with the glass coated specimens. At the end of each test the specimen was removed from the test furnace and water quenched. The final dimensions of the deformed ring specimens were measured at 12 locations around the diameter of the ring and the average results plotted as the percentage reduction in height against the percentage reduction in internal diameter.
Table 1. Height reduction of the test specimens given as true strain values.

<table>
<thead>
<tr>
<th>Lubricant</th>
<th>True Strain Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boron Nitride</td>
<td>0.2 0.5 0.7 0.8</td>
</tr>
<tr>
<td>Glass Fibre</td>
<td>0.2 0.3 0.5 0.7 0.8</td>
</tr>
<tr>
<td>Powdered Glass Fibre</td>
<td>0.2 0.3 0.5 0.7 0.8</td>
</tr>
<tr>
<td>Borosilicate Powder</td>
<td>0.2 0.3 0.5 0.7 0.8</td>
</tr>
</tbody>
</table>

2.2. FE Simulation Work and Experimental Results

The ring compression tests were initially modelled as an isothermal 2D axisymmetric process using DEFORM 2D V11 in order to determine if isothermal modelling would be sufficient to identify friction values. Simulations were performed using the shear friction law and friction factors $0 \leq m \leq 1$ in increments of 0.1 at the contact surfaces between the rings and the compression platens; a plastic flow model was used taking the Sellars-Tegart flow stress rule with constants generated from previous in house work (Sellars & Tegart, 1972). All simulations were performed using the same base simulation file and mesh (figure 2). The mesh was comprised of 2087 elements with an average element size of 0.168 mm. Mesh refinement was performed at the contact surfaces and the inner diameter to accurately capture heat transfer contact behaviour and the internal deformation of the ring, the element size in these regions varied between 0.06 and 0.09 mm. The temperature of the isothermal simulations was set to 1100°C and strain rates were kept the same as in the experimental work. The isothermal friction factor calibration curves generated using DEFORM 2D are presented in figure 3 and figure 4 shows the experimental calibration curves on the same plot as the calibration curves.

Figure 4 indicates that several of the lubricants tested would have a friction factor of $m > 1$, this is not a possible value for $m$ and indicates a problem with the simulation curves. The key difference between experimental and simulation work was non-isothermal testing compared to isothermal simulations. Analysis of temperature data logged during the compression test indicated that significant heat was lost during the compression of the specimen with temperature losses of up to 200°C recorded in some tests. The simulations were repeated with a contact HTC of 2.188 N/mm/s/oC (as used in Ti-6Al-4V extrusion simulation work using glass lubricant previously performed by Li et al. (2002)) in an attempt to account for the heat transfer effects observed. These simulation results gave temperature decreases of approximately 70°C. Performing simulations with higher HTC values (20 N/s/mm/oC) produced additional increases in the deformation of the internal diameter of the rings and temperature losses similar to those recorded experimentally (225°C). A comparison of calibration curves for isothermal simulations and simulations with HTC values of 2.188 N/s/mm/oC and 20 N/mm/s/oC is given in figure 5.

The results indicate that modelling heat transfer will result in an increased reduction of the internal diameter of the rings and produce significantly higher friction factor curves. For the high HTC simulations the curve corresponding to a friction factor value of $m = 0.5$ is above the curve for $m = 1$ that was generated with an HTC of 2.188 N/s/mm/oC. The assertion that heat transfer is the cause of discrepancies between isothermal simulation results and experimental measurements is reinforced by comparison of the shape of simulation specimens with experimental specimens as shown in figure 6, figure 7 and figure 8.
The shape of the workpiece from low HTC simulation results and that of the compressed specimen are noticeably different (figure 6 and figure 8). The use of a higher HTC value provides greater heat transfer from the specimen to the platens; this results in rapid chilling of the ring at the contact surfaces and consequently an increase in flow stress in this region. Deformation then becomes concentrated in the hotter regions of the compression specimen and there is a localised increase in flow velocity; the results can be seen in figure 7 where the shape produced has a deformation pattern closer to that of the experimental specimen than that produced by the low HTC simulations.

3. DISCUSSION

The results obtained from experimental work produced calibration curves that varied significantly from what was expected. When plotted against the calibration curves generated via isothermal simulation work, the apparent friction factor values for some of the glass lubricants exceeded the sticking condition with values of \( m > 1 \). It was determined that the use of isothermal simulations to generate the calibration curves was not suitable for the testing conditions modelled. Non-isothermal simulations performed using HTC values found in literature...
produced calibration curves with an increased degree of reduction of the internal diameter of the compression rings. The result was an increase in the gradient of the calibration curves compared to the isothermal curves at the same height reduction; increasing the HTC used in the simulations further moves the curves changing the position of the experimental results in relation to them such that experimentally determined values of $m$ no longer exceed 1. Comparison of images of the cross-sections of the experimental specimens to the simulations revealed that simulations with high HTC values produced a similar cross-sectional shape to that of the experimental results. The areas where deformation was concentrated in the experimental specimens was observed in macrographs; these areas of deformation show similarities to strain maps produced by high HTC simulations. However, these results indicate that different values of friction factor could be assigned to the experimental data by adjusting the HTC used in the simulation. As the friction factors and the HTC$s$ present at various interfaces in the experiment are unknown it is impossible to estimate the HTC$s$ by fixing the friction factor in the simulation and varying the HTC value until the curves become aligned with the experimental data as the same deformation may be produced by:
- Low friction paired with a high HTC
- High friction paired with a low HTC

Fig. 5. Isothermal calibration curves and non-isothermal calibration curves with a HTC at contact surfaces of 2.188 N/mm/s/°C.

Fig. 6. Simulation image of the strain distribution and shape at $\varepsilon_T = 0.5$ with friction 0.5 and HTC=2.188 N/mm/s/°C.

Fig. 7. Simulation image of the strain distribution and shape at $\varepsilon_T = 0.5$ with friction 0.5 and HTC=20 N/mm/s/°C.

In addition to this it is possible for the heat transfer of the system to be relatively complex as it may involve combinations of heat transfer through lubricants to the test platens, from bare metal to the environment on the sides of the rings as well through
lubricant to the environment for specimens that were entirely covered by lubricant.

A further complication is that die chilling produces not only a localised increase in flow stress but may also result in phase changes in the material. In the case of Ti-6Al-4V this would result in a transformation from the BCC β phase where flow stresses are typically low to the α+β phase (HCP+BCC) where flow stresses are significantly higher. These two phases are usually modelled in FE simulations using different flow stress formulations, so it is necessary to ensure that the material model used to simulate the process is robust enough to account for this possibility.

Fig. 8. Composite micrograph of the cross-section of a compression ring deformed to $\varepsilon_T = 0.5$ (actual $\varepsilon_T = 0.46$).

4. CONCLUSIONS

In order to accurately determine the friction factor via a combination of ring compression tests and simulation work it is necessary to correctly model the full experimental conditions in the simulation. For non-isothermal conditions where there is a large temperature gradient between the specimen and the test environment the heat transfer will significantly alter the deformation behaviour of the ring compression specimens compared to isothermal conditions. The impact of correctly modelling thermal conditions may vary with the magnitude of the thermal gradient between the workpiece and the environment. The extent of the heat loss recorded during experimental testing implies that a robust flow stress model is required to model the test in order to reflect any phase changes that may occur during the process and accurately predict the flow of material as a result of phase changes and temperature variations.

ACKNOWLEDGEMENTS

JDP gratefully acknowledges the financial support of the Engineering and Physical Sciences Research Council, UK. Thank you also to Mr Jacob Mawby for his assistance with the experimental trials and Mr Luke Marshall for his assistance in operating the TMTS.

REFERENCES


MODELOWANIE PRÓBY ŚPĘCZANIA PIERŚCIEŃ W WYSOKICH TEMPERATURACH DLA ODKUWEK ZE STOPU TI-6AL-4V

Streszczenie

W pracy badano warunki tarcia w wysokich temperaturach pomiędzy odkuwkami ze stopu Ti-6Al-4V i szkłem, będącym smarem, w próbie śpęczania pierscieni przy wykorzystaniu modelowania metodą elementów skończonych. Sprawdzono możliwość zastosowania symulacji zarówno przy zachowaniu warunków izotermicznych, jak i nieizotermicznych, do modelowania nieizotermicznych warunków prób doświadczalnych. Czynnik tarcia oszacowano porównując zmiany wewnętrznej średnicy pierscienia w eksperymencie i symulacji numerycznej. W trakcie badań zaobserwowano, że wartość współczynnika wymiany ciepła (HTCs) przyjęta w modelowaniu ma istotny wpływ na wyznaczony w oparciu o symulacje czynnik tarcia. Zauważono, że możliwe jest otrzymanie zbliżonych wyników dla kombinacji: mały HTCs-duża wartość czynnika tarcia oraz duża wartość HTCs/małych czynnik tarcia. W związku z powyższym stwierdzono, że warunki wymiane ciepła w procesie nieizotermicznym, charakteryzującym się znacznym gradientem temperatury pomiędzy próbką a otoczeniem, powinny być prawidłowo uwzględnione dla poprawnego wyznaczenia czynnika tarcia stosowanego w kolejnych symulacjach.

Received: October 1, 2014
Received in a revised form: October 9, 2014
Accepted: November 14, 2014
DETERMINATION OF FRICTION FACTOR BY RING COMPRESSION TESTING AND FE ANALYSIS

MICHAL GZYL1*, ANDRZEJ ROSOCHOWSKI2, LECH OLEJNIK3, KAMIL SIKORA4, MUHAMMAD JAWAD QARNI1

1 Advanced Forming Research Centre, University of Strathclyde, 85 Inchinnan Drive, Renfrew PA4 9LJ, United Kingdom
2 Design, Manufacture and Engineering Management, University of Strathclyde, James Weir Building, 75 Montrose Street, Glasgow G1 1XJ, United Kingdom
3 Institute of Manufacturing Technology, Warsaw University of Technology, ul. Narbutta 85, 02-524 Warsaw, Poland
4 Department of Applied Computer Science and Modelling, AGH University of Science and Technology, al. Mickiewicza 30, 30-059 Krakow, Poland

*Corresponding author: michal.gzyl@strath.ac.uk

Abstract

The goal of this study was to examine performance of various lubricants for aluminium alloy AA5083. Conventional ring compression tests were conducted at 200°C. Samples were compressed to 50% of the initial height with a constant ram velocity 0.5 mm/s using a servo-controlled hydraulic press. The optimization procedure was implemented in self-developed software to identify friction factors from experiments. The application launches remotely finite element (FE) simulations of ring compression with a changing friction factor until a difference between experiment and numerical prediction of the internal diameter of the sample is smaller than 0.5%. FE simulations were run using Forge3 commercial software. The obtained friction factor quantitatively describes performance of a lubricant and can be used as an input parameter in FE simulation of other processes. It was shown that application of calcium aluminate conversion coating as pre-lubrication surface treatment reduced friction factor from 0.28 to 0.18 for MoS2 paste. It was also revealed that commercially available graphite-based lubricant with an addition of calcium fluoride applied on conversion coating of calcium aluminate had even lower friction factor of 0.11

Key words: friction, FE simulation, ring compression

1. INTRODUCTION

In most metal forming operations, friction between dies and a workpiece has a significant effect on material flow, deformation force, surface quality of a final product and die wear. Additionally, properly defined friction is an important input parameter for commonly used numerical simulations of metal forming processes. Coulomb and shear friction laws (the latter is also known as Tresca law) are traditionally used in finite element (FE) modeling; they are given by equations (1) and (2), respectively (Joun et al., 2009):

\[ \sigma_t = \mu \sigma_N, \quad (1) \]
\[ \sigma_f = mk, \quad (2) \]

where: \( \sigma_t \) – frictional stress, \( \sigma_N \) – normal stress, \( k \) – shear flow stress, \( \mu \) – Coulomb’s coefficient of friction, \( m \) – friction factor.

Prediction of frictional stress made by the two laws can be significantly different and selection of
the way of describing friction should be made carefully with respect to the nature of a particular forming operation. Both laws give rather an estimation of frictional effects than an accurate prediction and are not suitable for some applications, e.g. Coulomb law significantly overestimates friction in extrusion process while shear friction law cannot predict the sticking region near the neutral point in rolling (Joun et al., 2009).

The most common method of identifying either friction factor or friction coefficient is uniaxial compression of ring samples. The method was proposed for the first time by Kunogi (1956) and the idea was to establish a relation between friction on tool-ring interface and a change of the internal diameter of a ring during compression. It was shown that the internal diameter of the compressed ring was decreasing along with increase of friction, as shown in figure 1. Then, Male and Cockroft (1965) experimentally obtained calibration curves for various friction coefficients, which display a relation between the reduction of inner diameter and the reduction of height of the ring. Those curves were widely used to identify friction coefficients for various materials. However, it was shown more recently that the shape of a friction calibration curve could be also affected by material properties and it was recommended to generate individual curves for each material (Sofuoglu et al., 2000; Camacho et al., 2013).

2. EXPERIMENTAL PROCEDURE

Ring compression tests were conducted on specimens machined from a rolled plate of AA5083. Dimensions of the rings were as follows: outer diameter – 16 mm, inner diameter – 8 mm, height – 5.3 mm. Tests were carried out with a constant ram velocity, \( v = 0.5 \text{ mm/s} \), on a 250 kN servo-controlled hydraulic press with a Zwick’s Cube control system. Cylindrical dies used to deform specimens had 25 mm diameter and were made of a Vanadis® 23 tooling steel. They were heated up to 205°C by band heaters with a power of 250 W each, temperature of the dies was slightly higher than targeted 200°C in order to compensate eventual heat loss on the surface. The ring sample was kept between the dies for 1 minute before testing to heat it up. During this time, the upper die was compressing the sample with a force of 0.5 kN to enhance conductive heat transfer and to obtain similar thermal conditions on both sides. Two controlling thermocouples, with 0.5 mm diameter, were located in the centre of each die, 1 mm below its surface, as shown in figure 2.

![Fig. 2. The sample placed between dies with attached thermocouples just before compression.](image)

Each ring was lubricated carefully (all surfaces) before putting it on the die, heating for 1 minute and compressing to 50±3% of the initial height. Various lubricants were tested, including conventional greases (molybdenum disulphide and graphite) and more sophisticated pre-lubrication surface treatment (conversion coating); details of lubricants and experimental plan are shown in table 1. Each compression test was repeated to obtain more reliable experimental data. Rocol, Molykote, Loctite 8008 and 8009 were applied using a brush, Formkote was sprayed to obtain uniform solid film. Pre-lubrication layer of calcium aluminate conversion coating, as shown in figure 3, was applied on the surface of selected rings in order to investigate its effectiveness in reducing friction. The same surface treatment was already shown to be effective for I-ECAP of 1050 and 5083 aluminium alloys (Rosochowski &
In this case, sample preparation procedure included: (1) surface cleaning; (2) application of calcium aluminate layer by keeping the sample in a bath of commercially available Gardobond Z3900 solution; and (3) lubrication with one of the pastes (Rocol, Molykote, Loctite 8008 and 8009). Height of ring was measured with a micrometer while a calliper was used to measure its inner diameter. In each case, an average from three measurements performed in different points was taken as a final result, which was especially important for compressed samples with shapes close to ellipsoidal.

Fig. 3. Ring samples from aluminium alloy 5083: as-machined (left) and after conversion coating (right).

Table 1. Details of lubricants used in the current study and experimental plan

<table>
<thead>
<tr>
<th>Lubricant name</th>
<th>Main ingredients</th>
<th>Comments</th>
<th>Conversion coating</th>
<th>No conversion coating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rocol® dry moly paste</td>
<td>MoS$_2$</td>
<td>grey oil-based paste</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Molykote® HTP solid paste</td>
<td>unknown</td>
<td>white oil-based paste</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Loctite® 8008</td>
<td>copper</td>
<td>paste based on copper particles</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Loctite® 8009</td>
<td>graphite, calcium fluoride</td>
<td>grey paste based on graphite and calcium fluoride</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Formkote® T-50</td>
<td>graphite</td>
<td>black solid film lubricant</td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>

3. FE ANALYSIS

3.1. Details of simulations

Three-dimensional elastic-plastic FE simulations of ring compression were performed in this study using Forge3 ver. 7.3 commercial package. Dimensions of the sample were the same as in the experiment. Flow stress for AA5083 was introduced using an analytical equation:

$$\sigma = A \exp(m_1 T) \varepsilon^n \dot{\varepsilon}^{m_2} \exp\left(\frac{m_3}{\varepsilon}\right)$$  \hspace{1cm} (3)$$

where: $T$ – temperature, $\varepsilon$ – strain, $\dot{\varepsilon}$ – strain rate, $A$, $m_1$, $m_2$, $m_3$ – material coefficients; $A = 445.93$, $m_1 = -0.00119$, $m_2 = 0.17172$, $m_3 = 0.01604$, $m_4 = -0.00953$.

Coefficients of the flow stress model and thermal properties of the material were taken from the Forge3 database. Although the material data were taken from the database, they exhibit a reasonably good agreement with literature results after rolling (Naka & Yoshida, 1999; Clausen et al., 2004). Dies had the same initial temperature as the ring placed between them, heat exchange coefficient between dies and the ring was 20 kW/m$^2$K. Thermal conductivity, specific heat capacity and density of AA5083 were 250 W/mK, 1230 J/kgK and 2800 kg/m$^3$, respectively. Simulations were run in order to investigate: (1) thermal conditions in the ring during heating and deformation; (2) an effect of material properties on friction calibration curves.

3.2. Thermal conditions

Temperature has a significant influence on material flow stress and performance of a lubricant; therefore, it is important to have a good control on temperature stability. FE modelling was performed in this work to simulate sample heating before testing and heat generation due to plastic deformation. It was shown in figure 4a that temperature within the ring was stable at 205°C after 1 minute of heating from the dies. Side walls of the ring are cooler than the interior; however, the difference of 0.1°C is considered negligible. Predicted temperature within the sample was the same as temperature recorded by the thermocouple placed in the die before deformation began. Then, temperature increase due to plastic work was also calculated by FE model and compared with experimental readings. It was predicted
that temperature exceeded 214°C in the sample compressed to 50% of the initial height, as shown in figure 4b. At the same time, the thermocouple placed in the die recorded 212°C, which gives a reasonably good agreement between simulation and experiment.

3.3. Effect of material properties on friction calibration curves

The classical approach to obtain friction factor, \( m \), is to compare experimental results with standard calibration curves plotted for various values of \( m \). However, it has been shown recently for different alloys that shape of calibration curves is dependent on a flow stress curve so the recommendation for using friction factor curves generated individually for each material was made by various authors (Sofuoglu et al., 2000, Camacho et al., 2013). To investigate this issue, the friction curves for AA5083 were compared with another aluminium alloy, AA1050, as shown in figure 5a. The former simulation was performed at 200°C while the latter was run at room temperature, corresponding to usual forming conditions for those materials. The flow stress curves for both materials were plotted in figure 5b; they were taken from Forge3 material database. The obtained results show that yield stress of a material can influence friction curves. The biggest difference was observed for low friction factors, the investigated curves tended to bring closer to each other along with increasing friction factor. The comparison of AA5083 and AA1050 confirmed that more accurate results can be obtained when friction factor curves are generated for individual materials.

\[ \Phi = \frac{|D_{\text{exp}} - D_{\text{FE}}|}{D_{\text{exp}}} \cdot 100\% < 0.5\% . \]
where: \(D_{\text{exp}}\) – experimental inner diameter of the compressed ring measured after testing, \(D_{\text{FE}}\) – prediction of inner diameter of the ring obtained from FE simulation.

The goal function, \(\Phi\), was minimised by changing friction factor, \(m\), in FE simulation at each iteration step, \(n\). The value of friction factor was modified according to the following equation:

\[
m_n = m_{n-1} + d_n k, \quad k = \begin{cases} 1, & \text{if } D_{\text{exp}} < D_{\text{FE}} \\ -1, & \text{if } D_{\text{exp}} > D_{\text{FE}} \end{cases}
\]

where \(d\) is a step size, reduced at each iteration step. Initial friction factor and step size were \(m_0=0.5\) and \(d_0=0.25\).

Figure 6 shows a schematic diagram of the optimisation procedure.

---

**5. EXPERIMENTAL RESULTS**

The tested samples are displayed in figure 7. It is apparent that the biggest internal diameter among them was obtained when combination of conversion coating and Loctite 8009 was used (sample 4 in figure 7). It is also clearly visible that sample coated only with Rocol paste (MoS\(_2\)) has the smallest internal diameter (sample 6 in figure 7), which indicates high frictional effect. It is worth noting that thin layer of calcium aluminate applied during conversion coating still remained after testing, which was revealed after cleaning the samples with white spirit. It suggests that the coating is resistant to the elevated temperature of 200°C in these working conditions.

Identification of friction factors was performed for all tested samples to allow quantitative comparison between the lubricants. The calculated friction factors for the tested lubricants are shown in figure 8. The lowest friction, \(m = 0.104\), was reported for a conversion coated sample lubricated with Loctite 8009. Combination of conversion coating and other lubricants resulted in friction factor varying from 0.181, measured for copper-based Loctite 8008, to 0.235, obtained after brushing with Molykote. Two variants of Rocol (MoS\(_2\)) were used, with and without pre-lubrication surface treatment by conversion coating. It was revealed that applied surface treatment decreased friction factor from 0.28 to 0.183. Graphite-based solid film lubricant (Formkote) gave the second worst result among conducted experiments; therefore, it is not recommended for aluminium alloys processed at 200°C.

---

**Fig. 6. Schematic diagram of the optimisation procedure.**

**Fig. 7. Samples after compression tests with various lubricants: 1 – Formkote (without conversion coating), 2 – Molykote, 3 – Loctite 8008, 4 – Loctite 8009, 5 – Rocol, 6 – Rocol (without conversion coating). Samples 2-5 were coated with a layer of calcium aluminate before applying a lubricant.**
REFERENCES


6. SUMMARY

The goal of the current study was to determine friction factors for various lubricants using a ring compression test. The material tested was aluminium alloy AA5083 deformed at ~200°C. It was shown through FE modelling that friction calibration curves depend on the flow stress of a material; therefore, new curves were generated in this work instead of using the standard ones. The friction factors were determined using the optimisation procedure implemented in the self-developed software. The presented approach enables obtaining accurate value of the sought parameter in an automated way. Moreover, the developed software can be modified to solve other optimization problems requiring FE simulation results as an input data at each iteration step.

In the experimental part of this work, it was shown that friction factor can vary from 0.1 to 0.3 for different commercially available lubricants. It was revealed that application of pre-lubrication conversion coating can reduce friction factor from 0.28 to 0.18 for MoS2-based paste. Further reduction of friction factor to 0.11 was achieved for a mixture of graphite and calcium fluoride (Loctite 8009) applied on conversion coating of calcium aluminate. Additionally, the coating was shown to be resistant to heating up to 200°C. Future work will be focused on simultaneous identification of flow stress curves and friction factors from uniaxial compression tests of cylinders and rings.

Fig. 8. Friction factors obtained after running the optimisation procedure.

OSZACOWANIE CZYNNIKA TARCIA NA PODSTAWIE PRÓB SPĘCZANIA PIERŚCIONI I ANALIZY METODĄ ELEMENTÓW SKOŃCZONYCH

Streszczenie

Celem niniejszej pracy było zbadanie przydatności różnych smarów dla obróbki plastycznej stopu aluminium AA5083. Standardowe testy spęczania próbek pierścieniowych zostały wykonane w temperaturze 200°C. Próbki zostały spęczone do 50% początkowej wysokości ze stałą prędkością przesuwu narzędzia 0.5 mm/s. Testy zostały przeprowadzone na prasie hydraulicznej z serwomąppedem. Procedura optymalizacyjna została zaimplementowana w samodzielnie opracowanej aplikacji w celu zidentyfikowania współczynników tarcia na podstawie wyników prób spęczania. Aplikacja zdalnie uruchamia symulację metodą elementów skończonych (MES) procesu spęczania próbek pierścieniowej ze zmieniającym się współczynnikiem tarcia dopóki różnica pomiędzy wewnętrzną średnią próbki otrzymaną z eksperymentu oraz z symulacji numerycznej nie jest mniejsza niż 0.5%. Symulacje MES zostały wykonane w komercyjnym pakiecie Forge3. Otrzymane współczynniki tarcia pozwalają na ilościową ocenę przydatności użytchnego smaru oraz mogą być bezpośrednio użyte w symulacji MES. Wykazano, że dla smaru na bazie MoS2 zastosowanie powłoki konwersowej gliniianą wpactwa jako warstwy podsmarnej zmniejsza wartość współczynnika tarcia z 0.28 do 0.18. Zostało również pokazane, że dostępny na rynku smar grafitowy z domieszką fluoru wpactwa nałożony na powłoce podsmarnej gliniianą wpactwa dał jeszcze mniejszy współczynnik tarcia, równy 0.11.

Received: September 26, 2014
Received in a revised form: November 14, 2014
Accepted: November 20, 2014

INFORMATYKA W TECHNOLOGII MATERIAŁÓW
EXPERIMENTAL CHARACTERIZATION AND NUMERICAL MODELING OF THE MECHANICAL BEHAVIOR OF HALF SANDWICH LAMINATE IN THE CONTEXT OF BLANKING

L. CHEN*, T. CLAUSMEYER, A. E. TEKKAYA

Institute of Forming Technology and Lightweight Construction, TU Dortmund University, Baroper Str. 303, 44227 Dortmund, Germany
*Corresponding author: lin.chen@iul.tu-dortmund.de

Abstract

In this short paper the complex mechanical behavior of an aluminum/low density polyethylene (LDPE) half sandwich sheet in the context of blanking is investigated. Suitable mechanical tests for the polymer and metal layer as well as the delamination behavior of the adhesive between these two layers were conducted. The main focus of this study is to create a finite-element (FE) model for the blanking process of sandwich structures. Material parameters for a Lemaitre-type damage model, a Drucker-Prager and a cohesive zone model are identified for the metal, the polymer and the adhesive, respectively. The experimental force-displacement curves obtained in a blanking process of a half sandwich sheet are compared with the predicted results of the FE model. The qualitative agreement of the predicted force-displacement curves with the experimental results is good. Recommendations concerning the improvement of the FE model are given based on the obtained results.

Key words: sandwich plates, blanking, CDM, Drucker-Prager model, cohesive zone model

1. INTRODUCTION

Composite plates are increasingly used, e.g. in the automobile industry or structure applications because they offer advantages compared to solid metal sheets. These advantages include weight savings and noise-reduction. In manufacturing the forming behavior of these sandwich laminates is especially the focus of interest (Burchitz et al., 2005; Harhash et al., 2014). Composite plates exhibit unique failure modes during cold forming process compared to monolithic metal sheets. In the blanking process of two layer metal-polymer composites, the possible failure modes include the delamination of the two layers, the fracture of the surface metallic layer as well as the fracture of the polymer. Previous studies by Hambli (2001) and Steinbach et al. (2014) investigated the behavior of monolithic metal sheets during blanking with the help of simulations. However, to the best of the authors’ knowledge, there is not yet any research giving the numerical analysis of metal/polymer/metal sandwich structure in blanking process. Although some experimental studies exist in this field (Liewald, 2013). In order to predict the formability of the LDPE/EN AW 5005A laminate in blanking process, in this paper a fully characterized FE model is constructed.

For this purpose half sandwich laminates manufactured from commercial Dibond sandwich material are investigated. For the mechanical characterization of the metallic layer uniaxial tensile tests and notched tensile tests with varying radii are conducted to obtain the ductile fracture properties. To characterize the polymer layer tensile tests and disk compression test are conducted. The interface behavior of the laminate is studied using delamination
tests under classical mode I and mode II failure modes. Material parameters of suitable material models for the polymer and the metallic layer are determined by inverse identification. The mechanical behavior of the metallic layers is described by an enhanced version of Lemaitre’s damage finite strain elasto-plasticity model (Lemaitre, 1985). The validation of this model in the context of sheet metal blanking process was reported in a previous study (Steinbach et al., 2014). Since delamination in blanking is mainly governed by mode I and mode II failure, a corresponding cohesive zone model is chosen to consider this behavior in the finite-element simulations. Parameters for the cohesive zone model are successfully identified based on the experimental delamination tests. A FE-model of a blanking process is set-up in Abaqus to analyze the effect of blanking process using the identified material models for half sandwich laminate. The paper is structured as follows: The chosen methodology is explained. After that the results of mechanical tests, parameter identification and the blanking experiments are presented. Then the results of the blanking experiments are compared to the predictions of the FE model. The paper ends with a discussion and conclusion.

2. MATERIAL MODELS

For the analysis of the mechanical behavior of the composite sheets with the help of FE method it is essential to describe the material behavior of the polymer and metal layer as well as the interfacing adhesive. The selected models for the three layers are briefly presented in the following. The metal and the polymer are subjected to large inelastic deformation before rupture. Consequently the chosen constitutive models fall into the context of elasto-plasticity and damage at large deformation. For both materials the elastic behavior is modelled as isotropic. Assuming an additive split of the symmetric part $\mathbf{d} = \text{sym}(\mathbf{I})$ of the velocity gradient $\mathbf{I} = \mathbf{F} \cdot \mathbf{F}^T$ into an elastic part $(\mathbf{d}^e)$ and a plastic part $(\mathbf{d}^p)$ $d = d^e + d^p$, one obtains the Jaumann-type stress update

$$\dot{\mathbf{T}} = (1-D) \left[ \frac{E \nu}{(1+\nu)(1-2\nu)} \text{tr}(\mathbf{d}^e) \mathbf{I} + \frac{E \nu}{(1+\nu)} \mathbf{d}^e \right]$$

for the general case of fully-coupled elasto-plasticity, with $\mathbf{F}$ the deformation gradient, $D$ the damage variable as well as $E$ the Young’s modulus and $\nu$ the Poisson’s ratio. $\text{tr}(\mathbf{\cdot})$ denotes the trace operator. The general form of the yield condition for coupled damage and plasticity with isotropic hardening is

$$\Phi = \frac{1}{(1-D)} \sigma_{\text{eqv}} - \sigma_j (\alpha) \leq 0$$

where $\sigma_j$ and $\alpha$ represent the deformation-dependent yield stress and the equivalent plastic strain, respectively. For the metal $\sigma_{\text{eqv}} = \sigma_{\text{eqv}}^\text{pl} = \sqrt{3/2} \| \text{dev}(\mathbf{T}) \|$ the pressure-independent von Mises equivalent stress is chosen in terms of the norm $\| \cdot \|$. For the polymer the pressure-dependent equivalent Drucker-Prager (Drucker et al., 1952) stress

$$\sigma_{\text{eqv}} = \sigma_{\text{eqv}}^\text{vp} := \sigma_{\text{eqv}}^\text{pl} + 1/3 \text{tr}(\mathbf{T}) \tan(\beta)$$

with material-specific friction angle $\beta$, which considers the pressure-dependency of plastic yielding, is selected. The rate of plastic deformation $\mathbf{d}^p = \lambda \left( \frac{\partial \Phi}{\partial \mathbf{T}} \right)$ is determined by an associated flow rule in terms of the plastic multiplicator $\lambda$. The evolution of damage is given by

$$\dot{D} = \lambda \left( \frac{Y - Y_0}{S} \right)^\kappa \frac{1}{(1-D)\beta}$$

where $Y_0$, $\beta$, $S$ and $\kappa$ are required damage material parameters.

The driving force for damage $Y$ which is here stated in terms of the principal stresses $T_i$, considers the pressure-dependence of the propagation of damage by the material parameter $h$ and was previously used by Soyarslan et al. (2008)

$$Y = \frac{1}{2E(1-D)} \left[ -(1+\nu) \sum_{i=1}^{3} \left( T_i^2 - h(T_i^2) \right) + \nu \left( \text{tr}(\mathbf{T})^2 - h(\text{tr}(\mathbf{T})^2) \right) \right]$$

All FE-simulations are carried out in Abaqus using a VUMAT routine of the implemented Lemaitre coupled damage model. For the polymer the Abaqus implementation of Drucker-Prager plasticity is used. The contribution of a material point to the total element stiffness is set to zero as soon as the damage variable $D$ reaches the critical value $D_c$ in the metal. For the polymer a linear relation between the damage variable $D$ and the equivalent plastic strain $\alpha$ is used, after the $\alpha$ has reached the critical value $\alpha_c$.

In this research, the mixed mode delamination of the half sandwich is simulated by the commercial software ABAQUS built-in cohesive zone model.
The constitutive behavior of this model is formulated as a traction-separation law (TSL), which relates the traction, \( T \), to the separation, \( \delta \). The elastic behavior is given by

\[
\begin{bmatrix}
T_n \\
T_s \\
T_t
\end{bmatrix} = \begin{bmatrix}
K_{nn} & 0 & 0 \\
0 & K_{ss} & 0 \\
0 & 0 & K_{tt}
\end{bmatrix} \begin{bmatrix}
\varepsilon_n \\
\varepsilon_s \\
\varepsilon_t
\end{bmatrix} = K \varepsilon,
\]

(6)

where the nominal strain \( \varepsilon \) is defined as the corresponding displacement \( \delta \) divided by the original thickness of the cohesive element. \( K_{nn} \) corresponds to the stiffness in the normal direction while \( K_{ss} \) and \( K_{tt} \) correspond to the two shear stiffness in the 3D case. \( T_n, T_s \) and \( T_t \) are the corresponding cohesive tractions. For the damage initiation, the maximum separation criterion is used and is given by

\[
MAX = \left( \frac{\delta_n}{\delta_{n_{\text{max}}}}, \frac{\delta_s}{\delta_{s_{\text{max}}}}, \frac{\delta_t}{\delta_{t_{\text{max}}}} \right) = 1,
\]

(7)

where \( \delta_n, \delta_s \) and \( \delta_t \) are the separation in normal, first shear and second shear direction. \( \delta_{n_{\text{max}}}, \delta_{s_{\text{max}}} \) and \( \delta_{t_{\text{max}}} \) are the critical separation values for the three strain components. The damage evolution is using the table data from the delamination tests as shown in table 2.

Table 1. Material parameters of Al layer in Lemaitre damage model.

<table>
<thead>
<tr>
<th>( E ), GPa</th>
<th>( \nu )</th>
<th>( \sigma_0 ), MPa</th>
<th>( \beta )</th>
<th>( S ), MPa</th>
<th>( \kappa )</th>
<th>( Y_0 ), MPa</th>
<th>( D_{cr} )</th>
<th>( h )</th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>0.29</td>
<td>116</td>
<td>0.15</td>
<td>10.2</td>
<td>0.25</td>
<td>0.3</td>
<td>0.18</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2. Parameters in cohesive zone model.

<table>
<thead>
<tr>
<th>( K_{nn} ) MPa mm(^{-2} )</th>
<th>( \delta_{n_{\text{max}}} ), mm</th>
<th>( K_{ss}=K_{tt} ) MPa mm(^{-2} )</th>
<th>( \delta_{s_{\text{max}}} = \delta_{t_{\text{max}}} ), mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>297.7</td>
<td>0.053</td>
<td>17.9</td>
<td>0.21</td>
</tr>
</tbody>
</table>

3. METHODS

3.1. Testing methods

A commercial Dibond sandwich laminate (EN AW 5005A/LDPE/ EN AW 5005A) is decomposed by means of pliers after moderate heating of the adhesive in order to study the mechanical behavior of each component separately. The thickness of a single aluminum layer is 0.3 mm, while the thickness of the polymer core is 2.4 mm. Uniaxial tension tests (DIN EN ISO 527) are performed for characterization of the mechanical behavior of LDPE. The disk compression test sample is composed of four layers of circular (12 mm) LDPE sheet. The height of the disk sample is around 9.23 mm. The amount of glue between each layer is kept to minimum to limit the influence of the glue on the mechanical test. The cross head speed of the universal testing machine Zwick 250 is set to 5 mm/s for all mechanical tests.

The mechanical behavior of the interfacing adhesive between the aluminum and LDPE layer is studied with the help of two delamination tests. The traction separation relationship of the aluminum layer and LDPE layer under tensile mode (mode I) are tested with the help of specially-designed sample holder as shown in figure 1. For the investigation of delamination under tensile mode, the top and bottom free sides of Al/LDPE half sandwich plate are glued onto two square columns made of steel. The area of the interface between Al and LDPE is reduced to around 1/9 of the half sandwich plate’s area, such that the delamination occurs under mode I during the separation of the Al layer and LDPE layer. Shear mode delamination (mode II) is realized with the help of an additionally designed device as shown in figure 1. Two grooves are cut on both sides of the sample. The distance between two parallel grooves is 5 mm. This sample is mounted in the universal testing machine such that fracture occurs along a shear delamination path when the crosshead is moving.

The blanking tool is also mounted in the universal test machine Zwick 250. The commercially available punch is manufactured of hard alloy. For the presented experiments the diameter is set to 8 mm and the clearance is 0.08 mm. All the blanking tests are conducted at a constant speed of 5 mm/s such that the strain rate in the blanking experiment is of the same order of magnitude as in the mechanical tests.
3.2. Simulation methods

FE-models for the mechanical tests are set-up. 8 node trilinear elements with reduced integration (C3D8R) are employed in the 3D FE-model of the notched tensile tests and in the separation tests for Al. The minimum mesh size is set to 0.02 mm in the center of the minimum cross section of the notched tensile test samples; the same mesh size is used in the blanking simulation in the area between the punch and die of the sandwich sheet. The delamination test is simulated using a traction separation law. The polymer side is considered as elastic body for simplicity, while the Lemaitre model is used for Al. The same mesh size is used at both sides of the interface. For the simulation of the blanking process an axisymmetric finite element model using CAX4R elements for the half sandwich laminate is set up. Punch and die are simulated as rigid body.

4. RESULTS OF MECHANICAL TESTS AND PARAMETER IDENTIFICATION

4.1. Mechanical tests and parameter identification of Al layer

In this section the experimental results of the notched tensile tests for the Al layer, the delamination tests as well as the results of the tension and disk compression tests of for LDPE are presented. These results provide the basis for the identification of material parameters.

The force-displacement curves of the notched tensile tests (notch radii 5 and 10 mm) are shown in figure 2. Necking happens after considerable plastic deformation and partially leads to an obvious force decrease following the point of maximum force in EN AW 5005A aluminum sheet. The elastic and Swift-type hardening parameters are directly determined from the tensile test. The material parameters are given in Table 1. The parameter \( h \) governing the pressure-dependency of damage evolution was set to 0.

The damage parameters in the Lemaitre model are obtained by fitting the simulated force-displacement curves of notched tensile tests to the
experimental ones. The identified model accounts for the softening and fracture observed in the experiments. Since damage is fully coupled with elasticity and plasticity, the decrease of the force is affected by the plastic softening due to damage as well as necking. The end point of the force-displacement curve is determined by the deletion of the critical element at the center of the minimum cross section in the notched specimen. From there fracture initiates and rapidly propagates through the cross section of the sample. It shows that the predicted force-displacement curves for the tensile test with radius 5 mm and radius 10 mm, and in particular the onset of fracture agrees well with the experimentally determined ones.

4.2. Mechanical tests and parameter identification of interface

The traction separation behavior of the interface is specified by determining the stiffness $K_{nm}$ and $K_{ss}$ which equals to $K_{tt}$, the damage initiation criterion and the damage evolution. The stiffness of the interface is achieved from the initial linear part of the traction-separation curve from tensile and shear-mode delamination tests. From this curve the critical displacement and the following damage evolution behavior are also used in simulation.

![Fig. 3. Comparison of traction separation curve from delamination test and simulation under tensile mode (a) and shear mode (b).](image)

The delamination tests under tensile mode and shear mode are successfully conducted by means of the designed device on uniaxial tensile test machine. The traction separation curve of tensile delamination is shown in figure 3a. The traction increases linearly at the first. The maximum traction is 26.0 MPa. Figure 3a depicts the contour plot of polymer layer in tensile mode delamination test when the traction is close to the peak value. The color of the contour represents the level of damage of the interface. The interface only extends at the center square of the polymer sheet, therefore the rest margin of the polymer sheet remains undamaged during the whole test. It could be deduced from the contour plot that the delamination initiates from the four corners of the square and propagates towards the center.

Figure 3b shows the traction separation curve of delamination test under mode II. The contour plot of the FE model shown in figure 3b illustrates the shear model delamination. The maximum traction 7.2 MPa under shear mode is smaller than that of tensile mode. The separation of the two surfaces happens within a longer path until the maximum traction achieves compared with that under tensile mode. The energy per area needed for the total separation is 8.5 kJm$^{-2}$, which is larger than that of tensile mode (2.8 kJm$^{-2}$).

4.3. Mechanical tests and parameter identification of LDPE layer

Parameters for the Drucker-Prager model described in section 3.1 are identified for LDPE. The strain rate dependency and temperature dependency are not taken into consideration. The pressure sensitivity parameter $\beta$ is calculated to be $32^\circ$ from equation (3) using the experimentally determined yield strength in tension and compression. Figure 4a shows the stress strain curve from tensile test and compression test in terms of the engineering strain and the nominal stress. The yield strength $\sigma_t = 12.2$ MPa in tension and $\sigma_c = 20.1$ MPa in compression are determined according to DIN EN ISO 527. It means that plastic yielding depends on the pres-
sure. The isotropic hardening behavior for LDPE is determined from the disk compression test. Assuming an isochoric motion, the obtained true stress-true strain curve is input as tabulated hardening data for the polymer model.

5. APPLICATION

Finally, the experimental results of the blanking process of the half sandwich sheet are presented. With the identified models at hand, the blanking process is simulated and the predictions of the simulation are compared to the experimental results. Figure 4b shows the force-displacement curve of blanking process. The first peak refers to the blanking of the aluminum layer, the following peak is corresponding to the cutting of the polymer layer. The maximum force is overestimated by the simulation by about 20%, while the level of the force related to separation of the polymer agree well in simulation and experiment. Rupture occurs at a smaller displacement according to the simulation than in the experiment for the Al layer and the polymer.

The cutting edge of half sandwich sheets show different characteristics compared to blanking of monolithic sheets. Partial delamination occurs between the polymer and the metal. Furthermore, the rollover is larger due to the low stiffness of LDPE. In a top view of the circular workpiece the area affected by rollover appears as a ring. In this ring the aluminum layer bends towards the polymer layer.

The workpiece on top of the die is studied while the removed part under punch is considered as scrap.

Figure 5 shows the cutting edge of the workpiece. Then along with the penetration of the punch the aluminum layer is cut and exhibits typical roll over and burnish area. After the scrap of the aluminum layer separates the workpiece, it is pushed by the punch towards the polymer layer. The polymer layer is then pushed by the aluminum scrap towards the die. At this time the interface close to the punch is under the combination of large tension and shear load. This leads to the partial delamination of the workpiece. Finally the polymer layer is cut by the decreasing distance of the punch and die. The fully characterized FE model for half sandwich plate is applied in blanking process.

The damage variable contour of the cohesive element and the polymer layer is shown in figure 5b. From this figure it could be seen that the simulated
cutting edge is similar to that of the experimental one, in particular for Al and the upper part of LDPE. The roll over behavior of the half sandwich sheet is qualitatively captured by the simulation. The applied cohesive zone model gives a good prediction of the position and the length of delamination in the region close to the cutting edge. On the other hand the simulation results shed light to the other possible failure of delamination which could not be directly revealed from experiment. In figure 5b the elements are highlighted in terms of the interface far away from the cutting edge but under the position where aluminum layer starts to bend towards the polymer layer. The interface at this region is under large tension stress due to the lack of blank holder force. If the interface strength of the sandwich product is not strong or homogeneous enough, delamination could also happen at this region.

6. DISCUSSION AND CONCLUSION

This study is probably one of the first to compare a FE-model with experimental data for the blanking of half sandwich sheet. The complex mechanical behavior involving elasto-plasticity and damage of a composite sheet was described in a comprehensive, yet efficient approach. The results indicate that the simulation model can be modified to improve the agreement between simulation and experiment. It is expected that more elaborate consideration of the strain rate sensitivity of the polymer will improve the prediction quality of the simulation. A better agreement of the predicted maximum force in blanking may be obtained by identifying the parameter $h$ with a suitable experiment. However, the presented method of setting up a simulation model for the blanking of sandwich presents an important contribution to establishing proper simulation methods for this industrial application.

ACKNOWLEDGEMENTS

Funding of Lin Chen by German Academic Exchange Service (DAAD) is appreciated.

REFERENCES


HOMOGENIZATION OF FIBER METAL LAMINATE STRUCTURES CHARACTERIZED BY ORTHOTROPIC AND ELASTIC-PLASTIC MATERIAL MODELS

TOMASZ NOWAK

ABB Corporate Research Center, Starowiślna 13A Street, 31-138 Krakow, Poland
*Corresponding author: tomasz.nowak@pl.abb.com

Abstract

This paper gives a theoretical background and provides numerical calculations for the non-linear mechanical behavior of the panel structures consisting of fiber reinforced composite and aluminum laminates. Such structures offer high performance-to-weight ratio, therefore they are widely used in aerospace, subsea and high-pressure applications. The Classical Lamination Theory with orthotropic material properties is recalled and extended about elastic-plastic model for metal layers. The simplified stress-strain relation, as proposed by Hencky and Ilyushin, was applied to capture the influence of metal’s plasticity on the mechanical performance of the hybrid structure. The numerical example showed, that at higher loads, the composite reinforcement provides a strong support for the aluminum layers when the metal approaches plastic deformation. In case of plastic flow within the aluminum, the bigger percentage of the external load is safely transferred to the composite fibers having much higher elastic limit. It prevents deformations of the aluminum laminate from being too large, and ensures the reliable operation of the Fiber Metal Laminates (FML) structure. Since the aluminum layers do not exhibit extensive strains, the application of Hencky-Ilyushin deformation theory seems to be reasonable for FMLs, even if aluminum layers are subjected to the anisotropic plastic flow. The proposed calculation method allows for very fast, but yet accurate, optimization of the Fiber Metal Laminate designs.

Key words: composite-reinforced metal structures, elastic-plastic orthotropic material, fiber-metal laminates, FML, Hencky-Ilyushin deformation theory

1. INTRODUCTION

Metal alloys are frequently used as the engineering materials for various industrial applications because they offer good mechanical strength, facilitate easy machining and assembling, and are relatively cheap. At high loads metals may exhibit a strong elastic-plastic behavior, thus the theory of plastic flow has been intensively investigated over last decades, (Hill, 1950; Chakrabarty, 1987). In contrast, the fiber-reinforced composites are generally expensive and difficult in manufacturing, however they offer the linear elastic behavior, and primarily - the superior strength-to-weight ratio (Jones, 1998; Gay et al., 2002). In the last three decades the Fiber Metal Laminates (FML), comprising the best properties of both materials, were studied extensively. The first FML was Arall®, a combination of aluminum and aramid/epoxy. In the 1980s, Delft University began developing a glass/epoxy FML called Glare®, which finally has been commercialized when Airbus decided to use it on the A380 aircraft. The mechanical properties of FML are typically studied by the Metal Volume Fraction, MVF, method, similar to the rule of mixtures for traditional composites (Vlot & Gunnik, 2001). Nowak (2013) and Nowak and Schmidt (2014) presented the analysis of equivalent mechanical properties for the FML tubes used in subsea and pressurized applications, respectively.

The work described in this paper considers a hybrid type of metal-composite laminate, aiming to
form a high-performance structure. A theoretical model for elasto-plastic composite reinforced aluminum panel is proposed, and validated by FEM calculations. The proposed approach assumes an orthotropic material model for the composite, and anisotropic plasticity model for aluminum layers. In the contrast to the earlier studies, author do not homogenize aluminum material with the composite structure into a single solid before the constitutive equation is applied. The elastic behavior of composite and the elastic-plastic behavior of metal are treated separately since the solution is found. It doubles the number of constitutive equations to be solved, but allows to isolate the anisotropic plastic flow in aluminum from purely elastic deformation of the composite layer. In addition, it was proved that a simple deformation theory of Hencky-Ilyushin may be successfully applied to FML structures, and high quality results may be found much faster. This facilitates an easy optimization of the hybrid structures if the fibers’ angles and layers’ thicknesses should be adjusted to the given load conditions.

2. ANALYTICAL CALCULATION MODEL

2.1. Basic assumptions

The analytical model used to study a composite-reinforced aluminum panel assumes a general orthotropic laminate, subjected to in-plane loading. In the classical approach (Herakovitch, 1998; Crawford, 2002), the stress-strain relation is characterized by classical approach (Herakovitch, 1998; Crawford, 2002), the stress-strain relation is characterized by

\[ \begin{bmatrix} N \\ M \end{bmatrix} = \begin{bmatrix} A & B \\ B & D \end{bmatrix} \begin{bmatrix} \varepsilon^0 \\ \kappa \end{bmatrix} \]

where:
- \( N, M \) are vectors of forces and moments, respectively
- \( \varepsilon^0, \kappa \) are vectors of strains due to forces and moments, respectively
- \( A, D \) and \( B \) are called tension stiffness, bending stiffness and coupling stiffness matrices respectively. They are calculated as follows:

\[ [A, B, D] = \frac{1}{2} \int_{-t/2}^{t/2} [K_i] (1, z, z^2) dz \]

where \( [K_i] \) is the stiffness matrix of a single \( i \) lamina, which is spaced from the neutral axis of the laminate by distance \( z \), and \( t \) is the total thickness of the laminate.

The matrix \( [B] \) plays an important role in the lamination theory, since it causes the complex interaction between the in-plane loads and the bending effects (out-of-plane strains). However, composite structures are typically designed in such a way that all components of \( [B] \) are zero, therefore generated stresses are only the result of in-plane strains, \( \varepsilon \), driven by in-plane forces, \( N \). In this case, the analysis simplifies with the following assumptions:

- The object under study is constructed out of two different materials, forming alternately several flat layers, which are perfectly bonded. The parts made of different materials are shown schematically in figure 1:
  - the inner part, marked as „c”, represents the composite shell (having thickness \( e_c \))
  - the outer parts, marked as „m”, symbolize metal layers (having the total thickness \( e_m \)).
- The composite part consists of several layers of fibers (basalt, or carbon, e-glass, etc.), oriented at different angles, and embodied in epoxy resin matrix.
- The orientation, number, and thickness, \( t_o \), of fiber layers are unrestricted, but practically it should form a symmetrical and balanced structure, e.g.: \([\alpha/\beta/\ldots]_s\), and the entire thickness of all plies constitutes the inner part (\( t_i = e_c \)).
- For the metal part, the non-linear elastic-plastic material properties are assumed with anisotropic plasticity, while for the composite part – the orthotropic elastic model is used.

2.2. Derivation of the model

To determine the stress and strain fields in both the steel layers and the composite structure, one can apply equilibrium, constitutive, and continuity equations. On the basis of the first relation, shown in figure 1, it is possible to state:

\[ \begin{align*}
\varepsilon_f \sigma_{mn} + \varepsilon_f \sigma_{ox} &= f_X \\
\varepsilon_f \sigma_{my} + \varepsilon_f \sigma_{oy} &= f_Y \\
\varepsilon_f \tau_{mxy} + \varepsilon_f \tau_{exy} &= f_{xy}
\end{align*} \]

where: \( \varepsilon_f \) are the wall thicknesses, \( \sigma_f \) are the stresses, and \( f_i \) are the external distributed forces [N/m]; indices \( m \) and \( c \) correspond to the metal and composite, respectively.

The constitutive equations are based on Hooke’s law for orthotropic material (Lekhnitskii, 1981), but they must be provided for the metal part and the
composite parts independently, with respective indices, \( i = m, c \):

\[
\begin{bmatrix}
\varepsilon_{ix} \\
\varepsilon_{iy} \\
\gamma_{iXY}
\end{bmatrix} = \begin{bmatrix}
1 / E_{ix} \\
-v_{iYX} / E_{iy} \\
\eta_{ix} / E_{ix}
\end{bmatrix} \sigma_{ix} + \begin{bmatrix}
1 / E_{iy} \\
1 / E_{iy} \\
\mu_{iy} / E_{iy}
\end{bmatrix} \sigma_{iy} + \begin{bmatrix}
1 / G_{iXY} \\
1 / G_{iXY} \\
1 / G_{iXY}
\end{bmatrix} \tau_{iXY}
\]

where: \( \varepsilon_{ij} \) is the elastic strain; \( E_{ij} \) and \( G_{ij} \) are the values of Young and Kirchhoff modules, respectively; \( \eta_{ij} \) and \( \mu_{ij} \) are coupling factors (known as Rabinovic’s and Chentsov’s) in corresponding directions.

If the isotropic material is considered \( v_{XY} = v_{YX} = v \), \( E_X = E_Y = E \), and \( \eta_{XY} = \mu_{XY} = \eta_X = \mu_Y = 0 \).

The continuity equations refer to the assumption, that metal and composite parts are perfectly bonded. Therefore one should also postulate that the strains in respective directions must be equal each other. While the composite can be considered as the linear elastic body, the elastic-plastic behavior of the aluminum should be assumed. Thus, the metal strains comprise the elastic parts \( \varepsilon^E \), as given by the equation (4), and the plastic components, \( \varepsilon^P \):

\[
\begin{bmatrix}
\varepsilon^E_{ix} \\
\varepsilon^E_{iy} \\
\gamma^E_{iXY}
\end{bmatrix} = \begin{bmatrix}
\varepsilon_{ax}^E + \varepsilon_{ax}^P \\
\varepsilon_{ay}^E + \varepsilon_{ay}^P \\
\gamma_{ax}^E + \gamma_{ax}^P
\end{bmatrix}
\]

Based on Prandtl-Reuss material model (Hill, 1950), the plastic strain increment is expressed by theflow rule as:

\[
de^P_{ij} = \frac{\partial f}{\partial \sigma_{ij}} d\lambda
\]

where: \( \varepsilon_{ij} \) and \( \sigma_{ij} \) are the strain and stress tensors respectively, \( f \) is the plastic potential, and \( \lambda \) is the scaling factor. In most cases of the anisotropic elastic-plastic behavior of the material the Hill’s yield criterion is assumed as the plastic potential:

\[
f : F(\sigma_Y - \sigma_Z)^2 + G(\sigma_Z - \sigma_X)^2 + H(\sigma_X - \sigma_Y)^2 + 2L \tau_{YZ} + 2M \tau_{ZX} + 2N \tau_{XY} - 1 = 0
\]

where: \( F, G, H, L, M, N \) are constants proposed by Hill and obtained by tests of the anisotropic material in different orientations.

Based on equations (6) and (7), an exemplary, \( X \) component of the plastic strain increment may be expressed as:

\[
de^P_{ix} = \frac{\partial f}{\partial \sigma_{ix}} d\lambda = [-2G(\sigma_Z - \sigma_X) + 2H(\sigma_X - \sigma_Y)] d\lambda
\]

Deriving other plastic strain components, focusing on plane stress conditions only \((\sigma_Z = \tau_{YZ} = \tau_{ZX} = 0)\), and using the matrix notation one can write:

\[
\begin{bmatrix}
de^P_{ix} \\
de^P_{iy} \\
de^P_{iYX}
\end{bmatrix} = \begin{bmatrix}
G + H & -H & 0 \\
-H & F + H & 0 \\
-2N & 0 & 2N
\end{bmatrix} \begin{bmatrix}
\sigma_X \\
\sigma_Y \\
\tau_{iXY}
\end{bmatrix} d\lambda
\]

Note 1: the number “2” which appeared in the equation (8) is not introduced into the equation (9) since Hill’s constants are typically multiplied by the factor “1/2”. Note 2: The Hill’s function reduces to the Mises stress potential in case of the isotropic material \((F = G = H = 1/2 = 3L = 3M = 3N)\).
The scaling factor increment \( d\lambda \) involves the plastic strain intensity increment \( \Delta \varepsilon_{pl} \) and the stress intensity, \( \sigma \), and may be expressed as:

\[
d\lambda = \frac{1}{C} \frac{\Delta \varepsilon_{pl}}{\sigma}
\]  \hspace{1cm} (10)

where: \( \sigma \) is the Hill’s yield criterion, which extends the Mises function to allow anisotropic behavior:

\[
\sigma = \sigma_{Hill} =
\]

\[
\frac{1}{\sqrt{C}} \sqrt{G \sigma_X^2 + F \sigma_Y^2 + H (\sigma_X - \sigma_Y)^2 + 2N \tau_{XY}^2}
\]  \hspace{1cm} (11)

and the effective plastic strain intensity increment reads:

\[
d\varepsilon_{pl} = \sqrt{C} \left( Fd\varepsilon_X^2 + Gd\varepsilon_Y^2 + Hd\varepsilon_{XY}^2 \right) / (HF + GF + HG) + 2d\varepsilon_{XY}^2
\]  \hspace{1cm} (12)

The constant \( C \) can be assigned quite freely, and it basically should simplify calculations. Gabryszewski and Gronostajski (1991), based on Hill’s proposal, use \( C = 2(H+F+G)/3 \), while in ABAQUS (2013) the constant \( C \) is equal to one. Please note also that whichever factor \( C \) is used, the plastic work, corresponding to the product of strain and stress, is exactly the same.

Thus, based on equations (4) and (9), remembering \( \varepsilon_{XY} = 2\varepsilon_{XY} \), and assuming that flow rule (6) could be replaced by the deformation theory of Hencky-Lilyushin (Chakrabarty, 1987), if the plastic strain increments developed at the loading path constitute the total plastic deformation without unloading, one can write in the matrix notation:

\[
\left[ \begin{array}{ccc}
\frac{1}{E_X} & -\frac{\nu_{XY}}{E_Y} & \frac{\nu_{XY}}{G_{XY}} \\
\frac{\nu_{XY}}{E_X} & \frac{1}{E_Y} & \frac{1}{G_{XY}} \\
\frac{1}{E_X} & \frac{\nu_{XY}}{E_Y} & \frac{1}{G_{XY}} \\
\end{array} \right]
\left[ \begin{array}{c}
\sigma_{X} \\
\sigma_{Y} \\
\tau_{XY} \\
\end{array} \right]
\]

\[=
\]

\[
\left[ \begin{array}{ccc}
\frac{1}{E_X} & 0 & 0 \\
0 & \frac{1}{E_Y} & 0 \\
0 & 0 & \frac{1}{G_{XY}} \\
\end{array} \right]
\left[ \begin{array}{c}
\sigma_{X} \\
\sigma_{Y} \\
\tau_{XY} \\
\end{array} \right] =
\]

\[
\left[ \begin{array}{c}
\sigma_{X} \\
\sigma_{Y} \\
\tau_{XY} \\
\end{array} \right]
\]  \hspace{1cm} (13)

where: \( \phi \) is the plastic deformation potential, which is defined similarly to \( d\lambda \), but the strain incremental forms in relations (9) and (12), \( d\varepsilon \) take now their total values, \( \varepsilon \), and:

\[
\phi = \frac{1}{C} \frac{\varepsilon_{pl}}{\sigma}
\]  \hspace{1cm} (14)

Please note, that indices \( c \) and \( m \) in the equation (13) represent the composite and metal, respectively. The “bar” sign points out that the equivalent properties are being used for the composite, since it consists of several different plies. The material properties without “bar” sign refer to the metal.

Finally, relations (3) and (13) constitute a system of six equations with seven formal unknowns (\( \sigma_{mX}, \sigma_{mY}, \tau_{mXY}, \sigma_{cX}, \sigma_{cY}, \tau_{cXY}, \) and \( \phi \)):

\[
\left[ \begin{array}{c}
\sigma_{X} \\
\sigma_{Y} \\
\tau_{XY} \\
\sigma_{cX} \\
\sigma_{cY} \\
\tau_{cXY} \\
\end{array} \right]
\]

\[
= \left[ \begin{array}{c}
\varepsilon_{c} \\
\varepsilon_{c} \\
\varepsilon_{c} \\
\varepsilon_{c} \\
\varepsilon_{c} \\
\varepsilon_{c} \\
\end{array} \right]
\]

\[
\left[ \begin{array}{c}
1 + \frac{\phi(G+H)}{E_X} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{E_Y} & 0 & -\frac{\tau_{cXY}}{G_{XY}} & -\frac{\tau_{cXY}}{G_{XY}} & 0 & 0 \\
0 & -\frac{\tau_{cXY}}{G_{XY}} & \frac{1}{E_Y} & -\frac{\tau_{cXY}}{G_{XY}} & -\frac{\tau_{cXY}}{G_{XY}} & 0 & 0 \\
0 & 0 & 0 & 1 + \frac{\phi(G+H)}{E_X} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{E_Y} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{E_Y} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{G_{XY}} \\
\end{array} \right]
\]  \hspace{1cm} (15)

The closed-form solution of the equation (15) is generally not possible, since the plastic deformation potential \( \phi \) is a function of unknown stresses (\( \sigma_{mX}, \sigma_{mY}, \tau_{mXY} \)). However, it can be easily found numerically. Having the Hill’s material constants determined, one has to select a test point (\( \varepsilon_{pl}, \sigma_{pl} \)) on the true stress – plastic true strain curve managed for unidirectional tensile experiment, and calculate a “test plastic deformation potential”, \( \phi^* \) defined as:

\[
\phi^* = \frac{1}{G+H} \frac{\varepsilon_{pl}}{\sigma_{pl}}
\]  \hspace{1cm} (16)

If the test plastic deformation potential \( \phi^* \) is introduced into the equation (15) instead of \( \phi \) the set of equations may be solved for stresses, and finally, the actual value of \( \phi \) can be found by the equation (14). The calculation procedure with selection of the next “test point” should be repeated, until the difference between plastic deformation potential \( \phi \) and its test value \( \phi^* \) reaches the assumed tolerance.

The solution of the equation (15) can be successfully found if the respective material properties for
anisotropic plasticity material and the orthotropic composite layer are determined. Since the former is quite well elaborated in the metal industry (e.g. Gabryszewski & Gronostajski, 1991; ABAQUS, 2013), we will focus on the later in the next paragraph.

2.3. Equivalent material properties for the composite

Derivation of the equivalent material properties for the composite part refers basically to the Classical Lamination Theory. The mechanical behavior of a single ply is characterized by its local stiffness matrix, \([k_i]\), which is oriented along the roving fibers, according to the local co-ordinate system, (Herakovitch, 1998; Crawford, 2002):

\[
[k_i] = \begin{bmatrix}
E_i & V_{il}E_c & 0 \\
V_{il}E_l & 1 & 0 \\
1 & 0 & G_{LT}
\end{bmatrix}
\] (17)

Indices \(L\), \(T\) and \(LT\), \(TL\) refer to the longitudinal direction, transverse direction and shear, respectively. The corresponding values of the Young modules, \(E\), Poisson ratios, \(\nu\), and Kirchhoff modulus, \(G\), should be provided by the composite manufacturer, but if not given – they can be easily derived by the commonly used Rule of Mixtures, which incorporates the fiber volume fraction (Jones, 1998). Furthermore, because the longitudinal direction of a single \(i\)-th ply may be arbitrarily oriented in the global co-ordinate system - its \(Z\)-axis is rotated by the \(\alpha\) angle, therefore it is also needed to transform the local stiffness matrix into a global one, \([K_i]\):

\[
[K_i] = T^{-1}[k_i]T^{−1T}
\] (18)

where \([K_i]\) is the stiffness matrix in the global co-ordinate system, GCS, and \(T\) – is the transformation matrix, given as:

\[
T = \begin{bmatrix}
c^2 & s^2 & 2sc \\
s^2 & c^2 & -2sc \\
- sc & sc & c^2 - s^2
\end{bmatrix}
\] (19)

where: \(s = \sin \alpha\), \(c = \cos \alpha\), \(\alpha\) – rotation angle about \(Z\) axis from LCS to GCS.

If the composite consists of several plies (which is normally the case), then the global stiffness matrix \([K_c]\) for the whole composite structure is calculated with respect to the thickness of individual plies:

\[
[K_c] = \frac{1}{t} \sum_{i=1}^{N} t_i [K_i]
\] (20)

where: \(t_i\) – thickness of \(i\)-th ply, \(t\) – total thickness of all \(N\) plies \((t\) is equal to \(c_s\), according to the figure 1).

Finally, one can determine the equivalent mechanical properties of the multi-layer composite, as used in the equation (4). It is necessary to calculate the inverse of the global stiffness matrix, \([K_c]^{-1}\) first, and to perform some simple operations on its components:

\[
[K_c]^{-1} = \left[ S_c \right] = \begin{bmatrix}
\hat{E}_X & \hat{E}_Y & \hat{E}_{XY} \\
\hat{E}_Y & \hat{E}_X & -\hat{E}_{XY} \\
\hat{E}_{XY} & -\hat{E}_{XY} & \hat{E}_{XY}
\end{bmatrix}
\] (21)

3. HOMOGENIZATION OF FIBER METAL LAMINATE

The solution of the equation (15) gives the values of stresses in both: metal layers and the composite structure directly (without homogenization). However, the same result can be achieved using more common notation of the homogenization formalism. One may consider equation (20) as a simpler form of the generalized definition for an effective material properties \(<h>\) represented as the ensemble average of variable, \(h\), (Kroner, 1972):

\[
\langle h \rangle = \langle h(x) \rangle = \frac{1}{V} \int h(x) dV = \frac{1}{N} \sum_{i=1}^{N} h(x)
\] (22)

Applying the equation (22) to the layered structures of composite and metal, one may define the global stiffness matrix as:

\[
[K_G] = \frac{[K_M] e_M + [K_C] e_C}{e_M + e_C}
\] (23)

where the metal stiffness matrix \([K_M]\) reads:

\[
[K_M] = \left[ S_M \right]^{-1} =
\begin{bmatrix}
\frac{1}{E_x} + \varphi(G + H) & -\frac{V_{xx}}{E_y} - \varphi H & 0 \\
-\frac{V_{xx}}{E_y} - \varphi H & \frac{1}{E_y} + \varphi(F + H) & 0 \\
0 & 0 & \frac{1}{G_{xx}} + 4\varphi N
\end{bmatrix}^{-1}
\] (24)
and the composite stiffness matrix \([K_c]\) is defined by the equation (20). Naturally, the metal stiffness matrix \([K_m]\) is state-dependent and relates to the actual value of the plastic deformation potential, \(\varphi\).

Since the global stiffness matrix of the whole composite-metal structure is found by the equation (23), the respective equivalent, or homogenized material properties (effective Young and Kirchhoff modules, Poisson ratios, and coupling factors) may be calculated by the procedure described by the relation (21).

4. NUMERICAL EXAMPLE

The numerical example was conducted for the flat plate (100x100 mm) consisting of two external layers of 2024 T3 aluminum (having the thickness of 1.0 mm, each) and the [0/45/-45/0] composite structure made of carbon fibers reinforced polymer (having the fiber content at the level of 60% vol., and each of four plies is 0.5 mm thick). The material properties in the elastic range were assumed to be the typical values, as provided by material suppliers for aluminum \((E = 67.5 \text{ GPa})\), carbon fibers \((E = 240 \text{ GPa})\), and epoxy resin \((E = 4.5 \text{ GPa})\). Important to note, that 2024 T3 aluminum shows the plastic behavior starting from the yield level of about 337.5 MPa \((\sigma = 337.5+1107.3e_{pl}^{0.7285})\), and exhibits anisotropic flows (the following Hill’s factors were assumed: \(F = 0.53\), \(G = 0.295\), \(H = 0.705\), \(N = 1.5\)).

The composite structure, made of plies at different orientation of fibers, forms the orthotropic material \((E_x = 81 \text{ GPa}; E_y = 24 \text{ GPa}; G_{xy} = 20.8 \text{ GPa}; v_{xy} = .72; v_{yx} = .21)\). The in-plane distributed forces were set at full load as 2 kN/mm and -1 kN/mm in \(X\) and \(Y\) directions, respectively, as shown in figure 1. Introducing above material properties into the equation (15) one could solve it for unknown stress components in composite and aluminum layers. The only unidentified factor - the plastic deformation potential \(\varphi\) can be found by a trial-and-error approach, or any other more structured iterative method, aiming to minimize the difference between \(\varphi\) and its test value \(\varphi^*\), as described by equations (14) and (16), respectively.

The calculated values of stresses [MPa], plastic strains [-], and deformation potential [-], as well as composite-to-aluminum stress ratio for different load levels were compared with the numerical simulation results delivered by ABAQUS software package, table 1.

Table 1. Comparison of the results provided by the proposed procedure and ABAQUS.

<table>
<thead>
<tr>
<th>Load [%]</th>
<th>(f_x)</th>
<th>(f_y)</th>
<th>Aluminum</th>
<th>Composite</th>
<th>Comp./AL. Mises ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>600</td>
<td>-300</td>
<td>215.6</td>
<td>215.5</td>
<td>0.89</td>
</tr>
<tr>
<td>60</td>
<td>1200</td>
<td>-600</td>
<td>317.0</td>
<td>316.9</td>
<td>1.01E-05</td>
</tr>
<tr>
<td>100</td>
<td>2000</td>
<td>-1000</td>
<td>340.1</td>
<td>339.5</td>
<td>3.24E-05</td>
</tr>
</tbody>
</table>

Interesting to note that the increased plastic flow within aluminum, caused by the higher load level, affects the stress ratio between composite and aluminum layers. Since the composite structure protects the aluminum strains from being too large, the application of Hencky-Ilyushin theory of small elastic-plastic deformations into FML plates is reasonable.

5. CONCLUSIONS

This paper introduced a structured approach for the design of composite reinforced metal plates used in many industrial applications. The provided theory captured the elastic-plastic behavior of the aluminum layers, by use of the Hencky-Ilyushin theory. It was proved, that this theory can be successfully applied to the materials exhibiting the anisotropic plastic flow, if small elastic-plastic deformations are ensured. Since it is the case for FML structures, the proposed calculation method allows for very fast, but yet accurate, optimization of the analyzed structures in the industrial conditions.

REFERENCES


HOMOGENIZACJA STRUKTUR KOMPOZYTOWO-METALOWYCH Z WYKORZYSTANIEM MODELI MATERIAŁU ORTOTROPOWEGO ORAZ SPREŻYSTO-PLASTYCZNEGO

Streszczenie

W artykule przedstawiono podstawy teoretyczne, zilustrowane przykładem numerycznym, opisujące model laminatu metalowo-włóknistego składającego się z kompozytu polimerowego wzmocnionego włóknami węglowymi oraz warstw aluminium. Przedstawione struktury hybrydowe (Fiber Metal Laminates, FML) charakteryzują się wysokimi własnościami mechanicznymi w stosunku do ciężaru, dlatego są chętnie używane w przemysłach lotniczych oraz stosowane w aplikacjach wysokościennych. W artykule klasyczną teorią laminatów (Classical Lamination Theory) uzupełniono o model sprężysto-plastyczny dla warstw metalowych. Zastosowano uproszczony model plastyczności zaproponowany przez Hencky’ego i Iljuszyna. Podano rozwiązanie numeryczne dla przypadku panelu FML poddanego płaskiemu obciążeniu dwuosobowemu. Wykazano, że w przypadku znacznego wytężenia, gdy dochodzi do uplastycznienia aluminium, znaczna część obciążeń zewnętrznych przekazywana jest do warstw kompozytu włóknistego, który charakteryzuje się znacząco wyższą granicą sprężystości. Takie zachowanie zabezpiecza warstwy aluminium przed nadmiernym płynięciem i umożliwia bezpieczną eksploatację struktury hybrydowej, nawet w przypadku wysokich obciążeń. Wykazano, że przyjęta metoda obliczeniowa charakteryzuje się wystarczającą dokładnością, a dzięki swojej szybkości umożliwia przemysłowy optymalizację hybrydowych struktur FML.

Received: October 1, 2014
Received in a revised form: November 5, 2014
Accepted: November 19, 2014
COMPETITION BETWEEN KIRKENDALL SHIFT AND FRENKEL EFFECT DURING 2D DIFFUSION PROCESS

BARTEK WIERZBA*, PATRYCJA WIERZBA

Rzeszow University of Technology, al. Powstańców Warszawy 12, 35-959, Rzeszów, Poland
*Corresponding author: bwierzba@prz.edu.pl

Abstract

In this paper numerical description of the interdiffusion process where the competition between the Kirkendall shift and Frenkel effect is showed. The vacancy generation and voids evolution (Frenkel effect) is discussed in terms of numerical simulations in 2 dimensional space. The proposed approach based on the generalized Darken approach where the volume velocity is essential in defining the local material velocity at non-equilibrium.

Key words: Kirkendall shift, Frenkel effect, vacancy distribution, void formation

1. INTRODUCTION

The Kirkendall experiment (Smigelskas & Kirkendall, 1947) has focused new attention on the mechanism of diffusion in metallic systems. After his experiment there is no longer doubt that a marker shift occurs in diffusion couples and indicate the vacancy mechanism of the diffusion (Seitz, 1953).

The first experimental evidence that the holes form from a supersaturated solution by a process of heterogeneous nucleation was presented by Balluffi (Balluffi, 1954). He showed, that required relative excess vacancy concentration for hole formation in most specimens is probably <0.01. Balluffi suggested that the holes will form in all systems when vacancies are pumped into any small region of the diffusion zone at the rate of about $10^{16} / \text{sec/cc}$.

Control of voiding in alloys is an important practical problem of materials science. As a rule, the voiding should be suppressed since they lead to failure of microelectronic circuits (Tu, 2007; Tu, 2011). Voiding is a result of relaxation of pure material or alloy supersaturated with vacancies. The relaxation of vacancy subsystem can proceed by joining of vacancies into voids.

In this paper the method for the voids grow during diffusion in 2 dimensional space is presented. The method allows for the estimation of the radius of the voids during mass transfer.

2. MODEL

The model was developed by Wierzba, 2014. The bi-velocity method is based on the Darken model for interdiffusion. The main law is the mass conservation law for each component. This law in 2D can be written as:

$$\frac{\partial N_i}{\partial t} + \nabla \cdot (-D_i \nabla N_i + N_i \nu^{\text{drift}}) = 0$$

where: $N_i$ denote the molar fraction of the component, $D_i$ is the intrinsic diffusion coefficient and $\nu^{\text{drift}}$ denote the drift velocity. The drift velocity, after Darken, is defined from known diffusional fluxes of the components:
Above equation is called the Poisson equation and can be solved by iteration methods. The model take into account the vacancy conservation law:

\[
\frac{\partial N_v}{\partial t} + \nabla \cdot j_v + \frac{N_v - N_v^{eq}}{\tau_v} = 0
\]

(3)

Where \(N_v\) is the vacancy molar fraction, \(j_v\) is the vacancy flux. The \(N_v^{eq}\) and \(\tau_v\) denote the vacancy equilibrium molar fraction and relaxation time, respectively. The vacancy flux is a sum of the fluxes of the components, mainly:

\[
\sum_{i=1}^{r} j_i + j_v = 0 \quad \Rightarrow \quad j_v = \sum_{i=1}^{r} D_i^v \nabla N_i
\]

(4)

Above equations should be supplemented by boundary conditions. The concentration of the vacancies growth around the initially inserted void(s) \(|\Omega_v|\). This growth is due to the drift velocity. The following boundary condition should be implemented around the void(s):

\[
\frac{\partial N_i}{\partial t} = \frac{\partial N_v}{\partial t} = 0 \quad i=1,...,r, \text{ on } |\Omega_v|\]

(5)

The void radius can be approximated from the analytical expression:

\[
R = D_v \left( N_v - N_v^{eq} \right) \frac{1}{L_v} t + R_{min} \]

(6)

where: \(R_{min}\) denote the minimal void radius \(R_{min} = 10^{-8}\) m and \(t\) is the experimental time in seconds. \(L_v\) denote the mean vacancy migration length \(L_v = \sqrt{D_v \tau_v}\).

Finally for the binary A-B system the following set of the equations should be solved:

1. The components conservation law, Equation (1) describes the redistribution of components with account of lattice drift. This equation for A component can be rewritten in the following form:

\[
\frac{\partial N_A}{\partial t} - \nabla \cdot \left( D_A^\nu \nabla N_A \right) + \nabla \cdot \left( N_A \mathbf{u}^{\text{drift}} \right) = 0
\]

(7)

2. The redistribution of vacancies, Equation (3), taking the sinks and sources of vacancies into account in relaxation approximation (for simplicity we assumed that the vacancy fraction is small and does not influence the drift velocity):

\[
\frac{\partial N_v}{\partial t} + \nabla \cdot \left( \sum_{i=1}^{r} D_i^N \nabla N_i \right) + \frac{N_v - N_v^{eq}}{\tau_v} = 0
\]

(8)

3. The drift velocity, Eq. (2) - the Poisson equation - should be solved by e.g. iterative numerical schema. For the binary A-B system this equation reduces to the following form:

\[
\nabla \cdot \mathbf{u}^{\text{drift}} = \nabla \cdot \left( \nabla \mathbf{u}^{\text{drift}} \right) = \nabla \cdot \left( \left( D_A^\nu - D_B^\nu \right) \nabla N_A \right)
\]

(9)

where \(\mathbf{u}^{\text{drift}}\) is the drift potential.

4. The radius of the void, Equation, (6), in the system:

\[
R = D_v \left( N_v - N_v^{eq} \right) \frac{1}{L_v} t + R_{min}
\]

(10)

3. RESULTS

In this section the voids formation in binary diffusion couple will be simulated. We have introduced 5 voids into the simulation as the initial conditions. During the process these voids are growing. We assume that the mean migration length for vacancy was \(L_v = 10^{-8}\) m and the diffusion coefficients \(D_A = 10^{-9}\) and \(D_B = 10^{-10}\) cm\(^2\)s\(^{-1}\). Figures 1 - 4 show the two dimensional diffusion profiles over the time. Figure 1 presents the initial conditions (initial distribution over the components and initial void radius). Figures 2, 3 and 4 present the evolution of the components and voids. It is presented that during the time the diffusion flow from the faster to the slower diffusion part of the diffusion couple. From the figures the voids radius can be estimated. It is presented that the voids grow during the diffusion.
4. CONCLUSIONS

The analysis of the voids formation was presented and discussed. The two dimensional calculations for interdiffusion and voids formation processes was presented. It was showed that the voids grow during the diffusion process. The voids are formed on the faster diffusion side of the diffusion couple. The method can be extended to calculate the void grow during the diffusion and electric field in solder joints.

ACKNOWLEDGEMENTS

This work has been supported by the National Science Centre (NCN) in Poland, decision number 2013/09/B/ST8/00150.

REFERENCES


WPŁYW POŁOŻENIA PLASZCZYZNY KIRKENDALLA ORAZ EFEKTU FRENKLA PODCZAS PROCESU DYFUZJI WZAJEMNEJ

Streszczenie

W artykule zaprezentowana została metoda pozwalająca na wykonanie obliczeń oraz pokazanie wpływu efektów Kirkendalla oraz Frenkla podczas procesu dyfuzji wzajemnej. Metoda dwuprędkości jest uogólnieniem metody Darkena. Pokazano, że metoda pozwala na poprawne wyznaczenie położenia płaszczyzny Kirkendalla jak również określenie promieniowania pustki powstającej podczas procesu transportu masy.

Received: August 7, 2014
Received in a revised form: December 5, 2014
Accepted: December 11, 2014
CALCULATION OF THE Fe-Fe\textsubscript{3}C PHASE EQUILIBRIUM DIAGRAM

HENRYK ADRIAN*, PRZEMYSŁAW MARYNOWSKI, DARIUSZ JĘDRZEJCZYK

AGH University of Science and Technology, al. Mickiewicza 30, 30-059 Kraków
*Corresponding author: adrian@agh.edu.pl

Abstract

In heat treatment processes of steel the very important role plays the Fe-Fe\textsubscript{3}C phase equilibrium diagram. It enables the selection of the temperature of austenitisation in respect to carbon content in steel as well as to predict the microstructure composition of annealed steel. For numerical calculation of the phase equilibrium diagrams the CALPHAD method is applied, based on the knowledge of Gibbs energy of phases existing in alloy as functions of temperature and chemical composition of alloy, \( G = f(T, C) \). The existing phases in Fe-Fe\textsubscript{3}C system are: liquid, austenite, ferrite and cementite. The functions of Gibbs free energy, \( G(T, C) \), for these phases were published elsewhere. The basic rule for the phase equilibrium diagrams calculations is to calculate the minimum free energy, \( G \), conditions in system for actual parameters: temperature, \( T \), pressure, \( p \), and chemical composition of alloy.

In presented paper the algorithm for Fe-Fe\textsubscript{3}C phase diagram calculation and developed computer program, PD\textsubscript{Fe-Fe\textsubscript{3}C} are presented. Results of calculations using this program were compared with date obtained using commercial program, Thermo-Calc and both results were identical.

Developed computer program is open and can take into account the alloy elements applied in steel in order to improve their physical properties. The data base for the effect of alloying elements on the free energy of system was prepared, enabling to calculate the effect of common alloying elements, such as Mn, Cr, Si or Ni, on critical temperatures, \( A_{1}, A_{3}, A_{cm} \). The program may be used for calculation the driving force for phase transformations which occur during heat treatment of steel process as well as in modeling the image of microstructure formed during heat treatment using such numerical methods as cellular automata, CA, or phase field method, PFM.

Key words: equilibrium phase diagram, Fe-Fe\textsubscript{3}C diagram, Gibbs free energy

1. INTRODUCTION

The phase diagrams are very important tools in study of alloys for solving problems in practical metallurgy. These diagrams define the regions of phases stability that can exist in an alloy system under the condition of constant atmospheric pressure. For binary system the coordinates of these diagrams are temperature and composition. The interrelationships between the phases, the temperature and composition in an alloy system are presented by phase diagram only under equilibrium conditions. Such conditions occur during slow heating and cooling rates of alloys, when the kinetics of transformations do not play important role.

The phase diagrams are developed using different technical methods: metallographic, dilatometry, magnetics, rentgenographic and – most frequently – the thermal analysis. It is necessary to emphasize that increasing role plays the computer methods of phase diagrams calculations, based on the knowledge of the parameters of free energies for different phases.

The iron-carbon phase diagram plays the very important role in technology of carbon steels because it enables to predict their microstructure
after annealing as well as their mechanical properties. It also enables to select the temperature for heat treatment processes. Currently as the standard iron-carbon the Chipman (1972) diagram is accepted.

Thermodynamic calculation of Fe-Fe₃C phase diagram was carried out by Gustafson (1985). In literature there are several works for evaluation of complex systems containing Fe, C and alloying elements (Andersson, 1988; Beyong-Joo, 2001; Hillert & Qiu, 1991, 1992; Huang, 1987; Lacaze & Sundman, 1991). In these works there are thermodynamic data used for calculation Fe-Fe₃C diagram with taking into account additions of alloying elements on the critical temperatures of transformations in this system.

The aim of present work was to develop the computer program for calculation of Fe-Fe₃C phase diagram.

2. DESCRIPTION OF THE THERMODYNAMIC EQUILIBRIUM IN Fe-M-Fe₃C SYSTEM

The method of equilibrium phase diagram calculations is based on the knowledge of Gibbs free energy, \( G \), for phases existing in analyzed system. Principle thermodynamic properties of \( \varphi \) phase which are used in calculations are free energy of pure \( i \) elements, \( G_i^\varphi \) and interaction parameters, \( L_{ij}^\varphi \) of \( i, j \) elements dissolved in \( \varphi \) phase. In Fe-Fe₃C system exist four phases: liquid solution, existing above liquidus temperature, austenite, \( \gamma \), ferrite, \( \alpha \), and carbide, Fe₃C. Functions G for austenite and ferrite in Fe-M-C system, containing alloy element, M, are described using Hillert and Staffansson (1971) two sublattice model, \((\text{Fe,M})_b(C,\text{Va})_c\), with substitution elements, Fe, M in one sublattice and carbon, C, with vacancies, Va, in second sublattice. The \( b \) and \( c \) parameters denote the number of sites in each sublattice. For austenite \( b=1 \) and for ferrite \( b=1, c=3 \).

The Gibbs free energy, \( G_m^\varphi \), for one mole of austenite, \( \gamma \), and ferrite, \( \alpha \), are given by formulas:

\[
G_m^\varphi = y_{Fe}y_C^\varphi G_{Fe,C}^\varphi + y_{Fe}y_{Va}^\varphi G_{Fe, Va}^\varphi + y_M y_C^\varphi G_{M, C}^\varphi + y_M y_{Va}^\varphi G_{M, Va}^\varphi + bRT(y_{Fe} \ln y_{Fe} + y_M \ln y_M ) + cRT(y_{Fe} \ln y_{Fe} + y_{Va} \ln y_{Va}) + E G_m^\varphi + m^g G_m^\varphi
\]

where excess free energies, \( E G_m^\varphi \) are:

\[
E G_m^\varphi = y_{Fe}y_M^\varphi \left( y_{Fe} L_{Fe,C}^\varphi + y_{Va} L_{Fe, Va}^\varphi \right) + y_M y_C^\varphi \left( y_{Fe} L_{Fe,M,C}^\varphi + y_{Va} L_{Fe, M, Va}^\varphi \right) + y_M y_{Va}^\varphi \left( y_{Fe} L_{Fe,C}^\varphi + y_{M} L_{Fe, M, C}^\varphi \right)
\]

The term \( G_{i, C}^\varphi \) is the Gibbs energy of component \( i \) with structure \( \varphi \) in a hypothetical nonmagnetic state and \( G_{i, C}^\varphi \) is the Gibbs energy of a hypothetical state, where all interstitial sites are filled with carbon atoms, C. The \( L \) parameters can be composition dependant according to a Redlich-Kister polynomial.

\( y_i \) are the site fractions in sublattices and related to ordinary mole fractions, \( x_i \) by formulae:

\[
y_C = \frac{b}{c} \frac{x_C}{1-x_C}
\]

\[
y_{Va} = 1 - y_C
\]

\[
y_{Fe} = \frac{x_{Fe}}{x_{Fe} + x_C}
\]

\[
y_M = \frac{x_M}{x_{Fe} + x_M}
\]

The magnetic term, \( m^g G_m^\varphi \) is magnetic contribution to the Gibbs free energy in \( \gamma \) and \( \alpha \) phases:

\[
m^g G_m^\varphi = RT \ln(\beta + 1)f(\tau)
\]

where \( \tau=\frac{T}{T_c} \) (\( T_c \) – temperature Curie) and \( \beta \) resembles magnetic moment. For ferrite the parameters \( \beta=2.22 y_{Fe}, T_c=1043 y_{Fe} \) and for austenite: \( \beta=-2.1 y_{Fe}, T_c=201 y_{Fe} \).

For \( \tau < 1 \):

\[
f(\tau) = 1 - \frac{79 \tau^{-1} + 474}{140} \left( \frac{1}{p^2} - 1 \right) \left( \frac{\tau^4 + \tau^6 + \tau^{15}}{600} \right)
\]

and for \( \tau > 1 \):

\[
f(\tau) = \left( \frac{\tau^4 + \tau^{-15} + \tau^{25}}{10 + \frac{315}{1500}} \right)
\]
where:

\[
A = \frac{518}{1125} + \frac{11692}{15975} \left[ \frac{1}{p} - 1 \right]
\]  

(12)

Parameter \( p \) depends on the phase structure, \( p=0.4 \) for ferrite and \( p=0.28 \) for austenite.

For liquid phase is assumed that liquid can be described as a mixture between Fe, M and C on a single sublattice. Therefore the Gibbs free energy is described by formula:

\[
G_m^{\text{liq}} = x_F e^{\text{liq}} C_F + x_C C_C^{\text{liq}} + x_M C_M^{\text{liq}} + \nabla T \left( x_F \ln x_F + x_M \ln x_M + x_C \ln x_C \right) + G_m^{\text{liq}}
\]  

(13)

where

\[
E G_m^{\text{liq}} = x_F e^{\text{liq}} C_F + x_C C_C^{\text{liq}} + x_M C_M^{\text{liq}} + x_C x_M I^{\text{liq}}_M, C
\]  

(14)

Cementite has orthorombic structure and is stoichiometric with respect to C, \((\text{Fe}, \text{M})_3\text{C}\). This phase is described by two-sublattice model. The Gibbs free energy for one mole of cementite is given by following expression:

\[
G_m^{\text{cem}} = y_F e^{\text{cem}} C_F + y_M C_M^{\text{cem}} + 3 \nabla T \left( y_F \ln y_F + y_M \ln y_M \right) + y_F y_M I^{\text{cem}}_{\text{Fe}, \text{M}, C}
\]  

(15)

Thermodynamic properties of Fe-M-C system are published elsewhere (e.g. Huang, 1987), a comprehensive data base of these parameters is presented by Adrian (2011).

3. METHOD OF Fe-Fe\(_3\)C DIAGRAM CALCULATION

In Fe-Fe\(_3\)C phase diagram there is 11 different areas, where exist one or two phases: 4 one-phase areas (liquid, \( L \), ferrite high temperature, \( \delta \) (above \( A_4 \)), ferrite low temperature, \( \alpha \) (below \( A_3 \)) and austenite, \( \gamma \) and 7 two-phase areas (\( L + \delta \), \( L + \gamma \), \( L + \text{Fe}_3\text{C} \), \( \delta + \gamma \), \( \gamma + \text{Fe}_3\text{C} \), \( \alpha + \gamma \), \( \alpha + \text{Fe}_3\text{C} \)). Characteristic temperatures are: \( T_t \) – melting point for Fe, \( T_{\text{cem}} \) – melting point for cementite, \( A_4 \) – temperature of allotropic transformation, \( \text{Fe}_\gamma \rightarrow \text{Fe}_\delta \), \( A_3 \) - temperature of allotropic transformation, \( \text{Fe}_\alpha \rightarrow \text{Fe}_\gamma \), \( T_{\text{peritectic}} \) - temperature of peritectic transformation, \( L + \delta \rightarrow \gamma \), \( T_{\text{eutectic}} \) - temperature of eutectic transformation, \( L \rightarrow \gamma + \text{Fe}_3\text{C} \), and \( A_1 \) – temperature of eutectoid transformation, \( \gamma \rightarrow \alpha + \text{Fe}_3\text{C} \). Temperatures of transformations (peritectic, eutectic, eutectoid) depends on the chemical composition of iron alloy and addition of M alloying element may change these temperatures. Calculation of Fe-Fe\(_3\)C phase diagram involves the calculation of two-phase areas boundaries as functions of temperature. In these areas in any temperature there is constant chemical composition of existing phases resulting from

![Fig. 1. Relationships between Gibbs free energy and carbon content for ferrite, austenite and liquid at 1420 K. Tangent line to \( G_m^{\gamma} \) and \( G_m^{\text{liq}} \) determines the chemical composition of austenite, \( C_\gamma \) and liquid, \( C_{\text{liq}} \) in two-phase (\( \gamma + \text{liquid} \)) area.](image-url)
energy functions of phases $\varphi$ and $\Phi$. Therefore in order to calculate the carbon contents in equilibrium phases, $c_\varphi$ and $c_\Phi$, it is necessary to solve following non-linear equations system:

$$G^\varphi_m(c_\varphi) - G^\varphi_m(c_\varphi) - \frac{\partial G^\varphi_m}{\partial c}(c_\varphi)(c_\varphi - c_\varphi) = 0 \quad (16)$$

$$G^\Phi_m(c_\varphi) - G^\Phi_m(c_\varphi) - \frac{\partial G^\Phi_m}{\partial c}(c_\varphi)(c_\varphi - c_\varphi) = 0 \quad (17)$$

Graphical illustration of the method is shown in figure 1, which presents the relationships between Gibbs free energy, $G$, and carbon content, $C$, for liquid, austenite and ferrite at 1420 K. At this temperature exist two phases: austenite with composition $C_\gamma$ and liquid with composition $C_{liq}$.

For solution the non-linear equations system the steepest descent method was used, described in Burden (1985). Advantage of this method, in comparison to others (e.g. Newton method) is, that it does not require an accurate initial approximation to the solution in order to ensure convergence.

4. **PD_Fe_Fe3C COMPUTER PROGRAM**

The computer program PD_Fe_Fe3C was developed. The code was written in Delphi programming system. For solution of non-linear equation systems in two-phase areas, the steepest descent method was used (Burden, 1985). General scheme of calculation is as follows:

1. Input chemical composition of alloy
2. Calculate $T_n$, $T_{\text{cen}}$, $A_3$ and $A_4$
3. Define temperature limits for areas 1 to 7,
4. For $i=1$ to 7 repeat steps 5 to 11
   5. $T_a=T_d(i)$
   6. $j=1$
   7. Repeat steps 8 to 11 until $T_a<T_d(i)$
   8. calculate $C_\varphi(T_a)$, $C_\Phi(T_a)$
   9. store calculated data in Fe_C table
   10. $j:=j+1$
   11. $T_a=T_{a}+\Delta T$
12. Calculate the transformation temperatures – $T_{\text{peritectic}}$, $T_{\text{eutectic}}$, $A_1$

where $T_d(i)$, $T_a(i)$ – assumed temperature range for $i$ line. Except of pure Fe-Fe$_3$C phase diagram calculation the program enables to calculate modified diagram taking into account different alloying elements, such as Mn, Si, Cr but in low contents, common for low alloy steel, where these elements are dissolved in Fe containing phases and do not form separated phases, because the formation of compounds formed by these elements is not include in the algorithm of program. Program enables also to calculate the Gibbs free energy, $G=f(T)$ for different phases at given temperature.

![Fig. 2. Input data window of PD_Fe_Fe3C program.](image_url)
range (for definite chemical composition of alloy) - upper chart in figure 2, or $G = f(C)$ functions (for given temperature, $T$, figure 3). Calculated functions $G = f(C)$ for liquid, austenite and ferrite were compared with results of calculations using Therm-Calc program and good agreement was observed.

One of the option of program is the calculation of difference Gibbs free energy for ferrite and austenite as a function of undercooling. This difference, $\Delta G$, is driving force for $\gamma \rightarrow \alpha$ transformation (figure 2 – lower chart).

The effect of addition of 1.5 %Mn on $A_3$, $A_{cm}$ and $A_1$ temperatures in Fe-Fe$_3$C diagram is presented in figure 5. This addition mainly decreases the $A_3$ and $A_1$ temperatures as well as decreases the carbon content in eutectoid point.

![Fig. 3. The relationships between Gibbs energy, $G$, and carbon content, $C$, at 1000 K.](image3)

![Fig. 4. Calculated Fe-Fe$_3$C phase diagram.](image4)

Calculated Fe-Fe$_3$C phase diagram is presented in figure 4.

5. CONCLUDING REMARKS

For selection the optimum parameters for heat treatment of steels the useful tool is the quasi-equilibrium phase diagram, Fe-Fe$_3$C. This diagram enables also to predict the microstructure of annealed steel on the knowledge of carbon content. Developed computer program PD_Fe_Fe$_3$C enables to calculate this diagram. Additionally it enable to calculate the effect of alloying elements, commonly used in quenched and tempered steels, on the critical temperatures: $A_3$, $A_{cm}$ and $A_1$. For calculation the microstructure image of heated treated steel different numerical models and methods are used (e.g. Adrian & Spiradek-Hahn, 2009; Marynowski et al., 2013). The input data used for such calculations are Gibbs free energies of phases existing in considered system and data obtained using PD_Fe_Fe$_3$C program can be applied.

ACKNOWLEDGEMENT

Work performed within the AGH project no 11.11.110.299.

REFERENCES

Adrian, H., Spiradek-Hahn, K., 2009, *The Simulation of Dendritic Growth in Ni-Cu Alloy Using the Phase Field
Model, Archives of Materials Science and Engineering, 40, 89-93.
Chipman, J., 1972, Thermodynamics and Phase Diagram of Fe-C System, Metall Trans, 3, 55-64.

OBLICZANIE WYKRESU RÓWNOWAGI FAZOWEJ Fe-Fe₃C

Streszczenie

W obróbce cieplnej stopów żelaza ważnym narzędziem pomocniczym jest wykres równowagi fazowej Fe-Fe₃C. Umożliwia dobór parametrów obróbki cieplnej stali węglowych, jak również pozwala określić skład strukturalny stali po obróbce cieplnej. Obliczanie wykresów równowagi fazowej metodą CALPHAD jest oparte na znajomości energii swobodnej Gibbsa jako funkcji temperatury i składu chemicznego, G=f(T,C) faz istniejących w układzie: austenit, ferryt, cementyt, ciecz. Obliczanie wykresu równowagi fazowej polega na obliczaniu minimum energii swobodnej układu metodami numerycznymi. W prezentowanej pracy przedstawiono program komputerowy PD Fe-Fe₃C do obliczania wykresu wykresu Fe-Fe₃C. Program umożliwia również obliczanie wpływu pierwiastków stopowych na temperatury krytyczne wykresu w stali niskostopowej. Pozwala również obliczać siłę pędną przemian fazowych zachodzących podczas obróbki cieplnej stali, co może być przydatne przy obliczaniu obrazu mikrostruktury metodami np. automatów komórkowych czy pola fazowego.

Received: November 3, 2014
Received in a revised form: November 15, 2014
Accepted: December 3, 2014
OPTIMIZATION AND APPLICATION OF GPU CALCULATIONS IN MATERIAL SCIENCE

GRZEGORZ KORPALA*, RUDOLF KAWALLA

Institute of Metal Forming, Technische Universität Bergakademie Freiberg, Bernhard-von-Cotta-Str. 4 D-09599 Freiberg, Germany
*Corresponding author: Grzegorz.Korpala@imf.tu-frieberg.de

Abstract

Modern Graphic Processing Units (GPU) provide in combination with a very fast Video Random Access Memory (VRAM) very high computational procedure, outrunning the conventional combination of a Central Processing Unit (CPU) and Random Access Memory (RAM) in terms of parallel computing and calculation. Within this work a concept for parallel application of the CPU/GPU is presented which combines the approach for processing and managing of large amounts of data. The computer algebra system (CAS) Wolfram Mathematica is used for numerical calculation of a large Finite Difference Model (FDM). The CUDA-link feature of Mathematica was used to achieve a parallel working environment with a parallelized computation on available CPUs with a parallelization of calculations of Nvidia GPUs at the same time. An advanced desktop computer system was setup to use a high-end desktop CPU in combination with four TITAN GK110 Kepler GPUs from Nvidia. It will be shown, that the calculation time can be reduced by using shared-memory and an optimization of the used block and/or register size to minimize data communication between GPU and VRAM. Results for diffusion, stress field and deformation field for a deformation sample will be shown, which is numerically calculated from crystal plasticity, with over four million of FDM elements being calculated by each of the four used graphic cards. It will be clearly shown, that the overall calculation time is strongly depending on the storage time for the amount of data, both for the final result and as for the intermediate results for the different numerical increments. Nevertheless, a promising application of parallel computing for research in the field of materials science is presented and investigated, showing the possibilities for new approaches and/or more detailed calculations in a reasonable time.

Key words: GPU computation, numerical simulation, memory management, GeForce GTX TITAN, crystal plasticity

1. INTRODUCTION

Nowadays, many phenomena are described numerically. The data generated assist when designing technology, when estimating the safety of components or enable improved understanding of physical effects such as diffusion or heat transfer. The methods that have been predominantly used up until now are the Finite Difference Method (FDM) and the Finite Element Method (FEM). Both are different in the mathematical description. The FEM gives good calculation accuracy and speed with a lower number of elements. As shown in figure 1 schematically FDM 1 and 2 (with a low and high number of nodes) demonstrate lower calculation accuracy than FEM. The only opportunity for improving accuracy is to further refine the node density on FDM. However, this means a longer calculation time, more calculation steps and a greater quantity of data to transfer. However, one advantage of FDM is the simple application of individual physical phenomena.

Simple parallelisation of calculations in FDM yields potential for acceleration. Use of GPU architecture allows for additional acceleration. A faster Video RAM as well as a shared memory installed in the graphics processor and corresponding storage management are not always fully utilised. Commercial programs use either combination 1 or 2 (figure
2) due to the original source code. All combinations differ from each other not just in construction, but also in terms of calculation speed. This is strongly dependent on the data transfer paths. Modification of the source code makes it possible to apply combinations 3, 4 or 5 into FDM, where combinations 4 and 5 are the fastest possible variants. Variants 3, 4, and 5 can, due to the special GPU architecture, calculate several sections of the parent lattice at once, otherwise known as blocks, in a special process (Tariq, 2011), see figure 3.

A comparison of speeds of different types of memory can be seen in table 1. The fastest L1 cache and shared memory should be used the most according to the table, as the transfer speed of these is larger by several factors. When a few of a block’s pro-

---

**Fig. 1.** Schematic illustration of results in accordance with different calculation methods.

**Fig. 2.** Schematic illustration of data transfer in the computer with the use of different processing units.

**Fig. 3.** Parent lattice and block (data structure).
cesses access the same memory address then it is better to first store these in the shared memory and then allow these to be called up from the processes at higher speed (Sanders & Kandrot, 2010). When calculating with classic CPU a large proportion of the calculation time is used for storing data to the data carrier (HDD or SSD). Despite saving on several hard drives (via Raid) this problem has still not been solved. The only method appears to be writing the data to the hard drive in the background parallel to the calculation. However, this requires a great deal of RAM for intermediate storage of the data. Complete storage of data from 4 Titan graphics processors (6GB video RAM) requires RAM capacity of 48 GB, which can be implemented in most systems.

Table 1. Reading and writing speeds of different memory types (Woolley, 2013).

<table>
<thead>
<tr>
<th>Used in Combination:</th>
<th>Memory Type</th>
<th>Mem. Speed</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CPU</strong></td>
<td>SSD (reference)</td>
<td>550 MB/s</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
</tr>
<tr>
<td></td>
<td>RAM</td>
<td>18 GB/s</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
</tr>
<tr>
<td><strong>GPU</strong></td>
<td>Video RAM (Titan GTX)</td>
<td>288 GB/s</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
</tr>
<tr>
<td></td>
<td>L1 - Cache</td>
<td>~2.5 TB/s</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
</tr>
<tr>
<td></td>
<td>Shared Memory</td>
<td>~2.5 TB/s</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
</tr>
</tbody>
</table>

In order to determine the influence of calculation concepts (figure 2) on calculation time, a simple equation (1 was applied with different methods and with different data quantities (Grid Size) in Mathematica®:

\[ y_i = \sqrt{\left(\frac{(a_i+b_i)+2}{2}\right) \cdot 10 + 3} + 10, \quad (1) \]

where \( y_i, a_i \) and \( b_i \) tables are with a specified size. For compiling Mathematica® uses a C compiler from Microsoft Visual Studio® for combinations 1 and 2 and a NVCC compiler for combinations 3 and 4. Calculation procedure 5 could not be verified in this example as the use of shared memory is not advantageous in this calculation. The measured, absolute times for the simple calculation can be found in figure 4. It can be clearly seen that a reduction in calculation time is expected due to the use of rapid data transfer. Combination 5 should be even faster, as the data transfer from video RAM to GPU is accelerated via shared memory. This combination is used in further calculations.

![Fig. 4. Calculation times for different combinations of data processing depending on data size.](image)

2. OPTIMISATION POTENTIAL WITH GPU - COMPLEX CALCULATIONS

An algorithm is used for calculating complex physical phenomena in which the Finite Difference Method is used in accordance with Wilkins (1999) and Patankar (1980). The whole concept was newly programmed using combination 5, the fastest available one. When working on the compilation of subroutines, various block sizes and registry sizes were tested to gain the fastest possible acceleration. To test the efficiency the function (Wilkins, 1999) used for calculating the acceleration vectors was selected as an example in this case. It was surprisingly established that the compilation with the largest possible number of blocks (the largest number of threads running simultaneously) did not yield the shortest calculation times. It goes without saying that attention was paid to the highest level of occupancy for the GPU (the CUDA Occupancy Calculator was used). It must be pointed out that the data transfer from VRAM and the number of threads are the main influential factors. Dependency of the calculation time on the theoretical block size can be seen in figure 5. The theoretical block size can be found from the third root of the number of processes in the block. Various lengths, widths and depths of the block are considered and the fastest variant is taken for a theoretical length. Compilation with small block sizes, as expected, resulted in longer calculation times. The best compilation in this case is a block size of 4x4x4 and a registry size of 160. The fastest variant requires just 0.058 seconds to calculate the new acceleration vectors for 1 000 000 ele-
ments. The slowest example needed 0.174 seconds, which was three times the calculation time.

Fig. 5. Calculation time for 1,000,000 elements for movement in 3D.

Further calculations were also carried out such as the stress state, diffusion or heat transport which were also transferred in their entirety to the GPU without large data volumes via PCIe 3.0 Bus. The CPU is responsible for issuing commands and storing data on the hard drive. Controlling of graphic cards (calculation) and saving data takes place simultaneously. 4 graphic Geforce cards Titan GTX were used in the test, each with a Kepler processor architecture and 6 GB of storage capacity. 8 processes were also carried out in the CPU simultaneously: 4 for controlling GPUs and 4 for saving the data. The schematic illustration of the process for two processes running simultaneously can be seen in figure 6.

After optimisation calculation times were established to be 15% shorter. The calculation times (figure 7) were established for the complete procedure within figure 6 with storage of data on HDD. The maximum size of the lattice is approx. 12,000,000 elements per graphics card. The times were noticeably shorter after optimisation in all lattice sizes.

Fig. 6. Calculation algorithm of the new program and resources used.

3. APPLICATION OF GPU CALCULATIONS IN MATERIAL SCIENCE

In the calculation system the mathematical description of diffusion depending on chemical analysis (Mujahid & Bhadeshia, 1992), of heat transfer (Patankar, 1980), of stresses, strains and flows (Wilkins, 1999) was also implemented. The algorithm is structured for universal use and changing the source codes is possible without problems and without help from third parties. The calculations should help make the complex micro-structure development in the material easier to understand. The interpretation of the results of the numerical simulation is an additional support for the experimental investigations and aids clarification of the results that cannot be directly explained. Local conditions such as the chemical analysis, stress state or temperature provide the necessary information for model-
ling the energy contributions when converting or recrystallizing.

2.1. Example 1

In a 1920x1080 grid with a cell size of 0.25 μm², defined as austenite with a carbon content of 0.0224 at. %, the new phase was denoted as ferrite (in the form of ellipses 960 μm long and 20 μm high: as shown in figure 8). Calculation of diffusion and the stress state then takes place in two dimensions. The initial condition is stated as being without tension with a homogeneous chemical composition (C 0.5%, Si 2%, Mn 3%) and a temperature of 350°C.

![Fig. 8. Distribution of plates of over-saturated ferrite in austenite.](image)

Carbon diffuses to the phase boundary due to lower solubility in the defined over-saturated ferrite area. Surplus carbon can then be absorbed by the austenite there. Suppressing the carbides build-up in the steel promotes the transport of carbon in ferrite and austenite. In the austenite areas located close to the ferrite plates, a pronounced enrichment of carbon is observed due to the lower diffusion speed in austenite and the carbon concentration reaches 0.0326 at. %, figure 9.

![Fig. 9. Carbon distribution of concentration after 350 min.](image)

The resulting differences in the density of ferrite/austenite cause residual stresses to occur and even after 0.5 s an inhomogeneous stress field is caused in the simulated material. This stress field can be seen in both austenite and ferrite. The elastic distortion of the lattice can be used for further modelling stress-induced conversion figure 10.

![Fig. 10. Stress field in accordance with von Mises after 0.5 s.](image)

3.2. Example 2

A random generator is often necessary for modelling phenomena which take place in material. With the aid of CUDA, these can run very efficiently. In principle a seed is generated for every block and then made available for every process via conversion. The efficiency of such a generator is introduced here. Each process in the calculation carries out 18 random draws of a natural number from 0 to 999999. No shared memory is used here so that no additional calculations are carried out. Small, yet relevant differences in calculation times are established depending on block size with this approach. The largest possible grid size is 700³, which is limited by the 6 GB of the video RAM. The measured times for the calculation can be found in figure 11.

![Fig. 11. Necessary time for generating random numbers for different lattice sizes.](image)

3.3. Example 3

The calculation used in example 1 can also map more complicated 3D systems. This includes the
forming simulation with texture-dependent yield stress, which is shown in equation 2 as a model. The yield stress is modelled via two equations and brought together via a parameter $\gamma$. The cosine of the smallest angle between the direction of the greatest shear stress and the sliding direction serves as this parameter $\gamma$ here.

$$Y = Y_T \cdot \gamma + Y_T \cdot (1 - \gamma);$$

$$Y_T = 0.35 \cdot (1 + 4 \cdot \varepsilon_{e})^{0.35} \cdot p;$$

$$Y_S = 0.31 \cdot (1 + 4 \cdot \varepsilon_{e})^{0.3} \cdot p;$$

$$p = 1 + 3 \cdot P \cdot \sqrt[3]{V};$$

$Y_T$ - flow stress model from tensile test, $Y_S$ - flow stress model from torsion test, $\varepsilon_{e}$ - effective strain, $P$ - pressure, $V$ - relative volume, $\gamma$ - equal to cosines of angle between direction maximum shear stress and nearest slip direction. The parallel length of the model probes was 18 mm, the width was 10 mm and the thickness was 1 mm. 4,000,000 volumes were used. The nodes at the beginning and end of the length were fixed in direction $y$ and $z$ and pulled in direction $x$ with a speed of 1 m/s.

Depending on the grain, texture-dependent lines were shown via the stress field described in accordance with von Mises, along which incremental deformation took place. Depending on the grain, which had a random orientation, the lines were developed differently. In the final step the directions of the

Fig. 12. Results of simulation, stress field in accordance with von Mises: in the elastic area at the beginning of the rupturing process (a), in the plastic area during the rupturing attempt (b), in the plastic area at the end of the rupturing attempt (c).
4. SUMMARY

With continued developments in graphics processors it is expected that the potential in the algorithms shown in this publication will continue to grow. The algorithms used are based on the Finite Differences Method and are only limited by the size of the available video RAM. The principles of FDM are adhered to and the calculation algorithm was newly developed for the graphics processors. Storage of the calculation in the graphics card only becomes advantageous if transportation of the data via PCIe BUS is minimised. In complex functions, the use of shared memory speeds up calculations. For this purpose, the compilation of sub-routines using shared memory with different block sizes should be tested. New models of yield stress give the option of taking dependency of shear stress and texture into account, which can be used to describe the damage mechanisms.

REFERENCES

Sanders, J., Kandrot, E., 2010, CUDA by Example: An Introduction to General-Purpose GPU Programming 1st, Addison-Wesley Professional.

OPTYMALIZACJA I ZASTOSOWANIE OBLICZEŃ NA PROCESORACH GPU W INŻYNIERII MATERIALOWEJ

Streszczenie

Nowoczesne procesory graficzne (GPU) w połączeniu z bardzo szybką pamięcią typu VRAM stanowią wysoko wydajne obliczeniowo narzędzie, które w aspekcie obliczeń równoległych wyprzedza znacznie konwencjonalną centralną jednostkę obliczeniową (CPU) z pamięcią RAM. W pracy przedstawiona została koncepcja aplikacji wykonującej obliczenia równoległe na procesorach CPU/GPU, która może przetwarzać i zarządzać dużą ilością danych. Wykorzystano środowisko obliczeniowe CAS Wolfram Mathematica do rozwiązywania dużych modeli metodą różnic skończonych (FDM) oraz funkcjonalność Matematyki CUDA-link do równoczesnego zrównoleglenia obliczeń na procesorach CPU i Nvidia GPU. Na tej podstawie opracowano zaawansowany system komputerowy pozwalający na obliczenia na procesorze GPU w połączeniu z czterema procesorami TITAN GK110 Kepler GPU firmy Nvidia. Pokazano, że czas obliczeń został zredukowany przy wykorzystaniu pamięci dzielonej i optymalizacji bloku lub rozmiaru rejestru, w celu minimalizacji przesyłu danych pomiędzy GPU i VRAM. Przedstawiono wyniki dla dyfuzji, pola naprężeń i pola odkształcenia dla odkształcenia, przykładowej próbki, otrzymane z modelu plastyczności kryształu z ponad czterema milionami elementów FDM, dla których obliczenia wykonywano na czterech kartach graficznych. Przeprowadzone obliczenia jasno pokazały, że całkowity czas obliczeń jest silnie zależny od czasu dostępu do pamięci dla danych, zarówno w aspekcie otrzymywania wyników końcowych, jak i wyników pośrednich dla różnych kroków czasowych. Niemniej jednak w pracy przedstawiono obieguce wyniki badań nad zastosowaniem obliczeń równoległych w dziedzinie inżynierii materiałowej, pokazując możliwości wykorzystania nowych metod i bardziej dokładnych obliczeń w akceptowalnym czasie.

Received in a revised form: November 21, 2014
Accepted: December 9, 2014

INFORMATYKA W TECHNOLOGII MATERIAŁÓW
MULTI-FRONTAL PARALLEL DIRECT SOLVER FOR ONE DIMENSIONAL ISOGEOMETRIC COLLOCATION METHOD

PAWEL LIPSKI*, MACIEJ PASZYŃSKI

AGH University of Science and Technology, al. Mickiewicza 30, Krakow, Poland
*Corresponding author: lipski@student.agh.edu.pl

Abstract

In this paper we present a new multi-frontal solver for the isogeometric collocation method (ISO-C) on GPU. The ISO-C method constitutes an alternative for the isogeometric finite element method (ISO-FEM). The key advantage of ISO-C over ISO-FEM is that it does not include the computationally intensive operation of integrating the variational formulation. The ISO-C method requires using only a single collocation point per one basis function, whereas in ISO-FEM, Gaussian quadrature is applied on many points at each finite element. The presented multi-frontal solver for collocation method results in logarithmic execution time assuming that large enough number of GPU processors is available.

In this article, the method is employed for an exemplary 1D nanolithography problem of Step-and-Flash Imprint Lithography (SFIL). The algorithm, however, may be applied to a wide class of 2D and 3D problems.

Key words: step-and-flash imprint lithography, collocation method, isogeometric analysis, GPU, CUDA

1. INTRODUCTION

Efficient solvers and fast-processing techniques for the problems of material science have always been of great importance. Currently, with the emergence of multi-core and graphic processor architectures, a tremendous speedup can be achieved. As a consequence, there is a need to develop new state-of-the-art algorithms that could leverage this new power. Several attempts have already been made to rewrite existing solutions to benefit from the new architectures (e.g. PLASMA and MAGMA – see Agullo et al., 2009). The shared-memory graphic cards have already proven a good environment for running multi-frontal direct solvers for isogeometric finite element method (Wozniak et al., 2014). In this paper we propose a preliminary version of the multi-frontal direct solver for isogeometric collocation method, an efficient alternative to classical isogeometric finite element method (Auricchio et al., 2010). The structure of the solver is based on the one developed for one dimensional finite difference method (Obrok et al., 2010). The numerical results presented in this paper concern the Step-and-Flash Imprint Lithography (SFIL), a modern patterning process (Bailey et al., 2001a, Bailey et al., 2001b). In the paper we also analyze the convergence of the isogeometric collocation method for a uniform ("naive") selection of the collocation points and for Demko algorithm (Demko et al., 1985) for selection of the collocation points. We present that the Demko point selection guarantees unconditional convergence.

2. THE MODEL FOR POLYMER NETWORK DAMAGE PROBLEM

We focus on modeling the shrinkage of the feature after the photopolymerization occurred, still before removal of the template. It is assumed that there is a damage in the polymer network and thus the interparticle forces are weaker in one part of the domain, as shown in figure 1(a). For simplicity, we
consider one-dimensional horizontal cross-section of the domain.

Fig. 1. Horizontal cross-section of the photopolymer (a), numerical result (b).

We need to solve the following linear elasticity equation:

\[- \frac{d}{dx} \left( EA \frac{du}{dx} \right) = f \tag{1}\]

where \( E \) stands for the Young modulus, \( A \) – for the cross-sectional area and \( f \) – for the body force. We seek \( u \), which is vertical displacement field of the feature. The local damage is expressed by the \( E \) function. The boundary conditions are:

\[ u(0) = u(160) = 0 \tag{2} \]

The material data are given by:

\[ A(x) = 1, x \in [0,160] \tag{3} \]

\[ E(x) = \begin{cases} 2.5 \cdot 10^{-3} x^2 - 0.2x + 4, & x \in (20, 60) \\ 1, & x \in [0,20] \cup [60,160] \end{cases} \tag{4} \]

\[ f(x) = 0.0002 \sin \left( \frac{2\pi x}{160} \right) \tag{5} \]

The spatial dimension is expressed in nanometers. The numerical result computed by using our GPU collocation method with fourth order B-splines is presented in figure 1(b).

3. ISOGEOOMETRIC COLLOCATION METHOD

It is not only FEM that can be improved with the use of isogeometric analysis – the above concepts can as well be applied to the collocation method. In general, numerical solving of differential equations (either ordinary or partial) with collocation method (CM) is performed by choosing a space of candidate solutions (typically polynomials of relatively low degree) and then finding a linear combination of these candidates. This combination must satisfy given differential equations at certain points, called collocation points. In the isogeometric collocation method we provide a candidate space consisting of B-splines of an arbitrary order, from now on denoted as \( p \). Each spline of order \( p \) spreads over \( p+1 \) basic intervals. The concept is illustrated in figure 2, where we have the following knot-vector for isogeometric computations: \( \{0,0,0,1,2,3,4,4,4\} \). There are 7 uniform cubic B-spline basis functions in this example, denoted by \( N_{i,p} \). In this case we need to utilize 7 collocation points (marked as \( c_i \)) to assure that the number of basis functions is equal to the number of equations. Then we approximate the solution with a linear combination of B-spline basis functions

\[ u(x) \approx \sum_{i=1}^{7} B_i N_{i,p}(x) \tag{6} \]

We substitute the approximation to the equation:

\[ \frac{d}{dx} \left( EA \frac{d}{dx} \left( \sum_{i=1}^{7} B_i N_{i,p}(x) \right) \right) = f(x) \tag{7} \]

and we get

\[ \sum_{i=1}^{7} B_i \frac{d}{dx} \left( EA \frac{d}{dx} \left( N_{i,p}(x) \right) \right) = f(x) \tag{8} \]

Next, we select the collocation points \( c_{i,1,\ldots,7} \) and we prescribe the equation (in the strong form) in the collocation points, to obtain

\[ \sum_{i=1}^{7} B_i \frac{d}{dx} \left( EA \frac{d}{dx} \left( N_{i,p}(c_j) \right) \right) = f(c_j) \text{ for } j = 1,\ldots,7 \tag{9} \]

Since at any given collocation point there are up to \((p+1)\) non-zero B-spline basis functions (compare figure 2 where we have four B-spline basis functions at a point), we get a multi-diagonal matrix.
4. EMERGING ARCHITECTURES IN SCIENTIFIC COMPUTATIONS

Graphic Processing Units (GPUs) constitute a promising architecture for scientific computations since they allow to process data much faster when compared to the traditional CPU approach. GPUs feature large memory bandwidth and a substantial number of processing cores, allowing to perform the calculations in parallel on hundreds of threads. Compute Unified Device Architecture (CUDA) forms a platform for programming both CPUs and NVIDIA GPUs using a consistent programming model based off the C language with some extensions for expressing the parallelism.

CUDA treats the GPU as a coprocessor for the CPU. The CPU is called a host, whereas the GPU is called a device. The host plays the superior role, orchestrating computations and initiating memory transfers. Parallel fragments of the application are executed on the device as kernels, which, with respect to the source code, resemble serial programs. However, the GPU executes each kernel on many cores in parallel.

5. SOLVER ALGORITHM

Let us focus on the isogeometric collocation method with quadratic B-splines over knot vector \(\{0,0,0,1,2,3,4,5,5,5\}\) and control points \(\{0,1,2,3,4,5\}\) which results in \(N+p=5+2=7\) B-spline basis functions. In this case we select 5 collocation points located at the centers of elements \(\{0.5,1.5,2.5,3.5,4.5\}\) and we get the following system of linear equations, completed with Dirichlet boundary conditions:

\[
B_1N_{1,2}(0) = 0 \\
B_1N_{1,2}'(c_1) + B_2N_{2,2}'(c_1) + B_3N_{3,2}'(c_1) = f(c_1) \\
B_2N_{2,2}'(c_2) + B_3N_{3,2}'(c_2) + B_4N_{4,2}'(c_2) = f(c_2) \\
B_3N_{3,2}'(c_3) + B_4N_{4,2}'(c_3) + B_5N_{5,2}'(c_3) = f(c_3) \\
B_4N_{4,2}'(c_4) + B_5N_{5,2}'(c_4) + B_6N_{6,2}'(c_4) = f(c_4) \\
B_5N_{5,2}'(c_5) + B_6N_{6,2}'(c_5) + B_7N_{7,2}'(c_5) = f(c_5) \\
B_7N_{7,2}(160) = 0
\]

By computing the derivatives of the B-spline basis functions at the collocation points, as well as the values at the boundary nodes, we get a tri-diagonal system of linear equations. The system is sparse and can be decomposed into the following set of subsystems:

\[
\begin{bmatrix}
1 & 0 \\
1 & -1
\end{bmatrix}
\begin{bmatrix}
B_1 \\
B_2
\end{bmatrix} =
\begin{bmatrix}
0 \\
0.5f(c_1)
\end{bmatrix}
\begin{bmatrix}
-1 & 1 \\
1 & -1
\end{bmatrix}
\begin{bmatrix}
B_2 \\
B_3
\end{bmatrix} =
\begin{bmatrix}
-0.5f(c_2) \\
0.5f(c_2)
\end{bmatrix}
\]

These systems are equivalent to the original after summing up. We can then merge (which is denoted by \(\oplus\) and \(\rightarrow_M\)) adjacent pairs of the systems, perform Gaussian elimination (denoted by \(\rightarrow_g\)) and eliminate one fully assembled node from each system (denoted by \(\rightarrow_B\)).

\[
\begin{bmatrix}
1 & 0 \\
1 & -1
\end{bmatrix}
\begin{bmatrix}
B_1 \\
B_2
\end{bmatrix} =
\begin{bmatrix}
0 \\
0.5f(c_1)
\end{bmatrix}
\rightarrow_M \begin{bmatrix}
-2 & 1 & 1 \\
0 & 1 & 0 \\
1 & 0 & -1
\end{bmatrix}
\begin{bmatrix}
B_2 \\
B_3 \\
B_4
\end{bmatrix} =
\begin{bmatrix}
0.5f(c_2) \\
0 \\
0.5f(c_3)
\end{bmatrix}
\rightarrow_g \begin{bmatrix}
1 & -1/2 & -1/2 \\
0 & 1 & 0 \\
0 & 1 & 1/2
\end{bmatrix}
\begin{bmatrix}
B_3 \\
B_4 \\
B_5
\end{bmatrix} =
\begin{bmatrix}
-0.25f(c_2) \\
0 \\
0.5f(c_3) + 0.25f(c_2)
\end{bmatrix}
\rightarrow_B
\]

Similarly, for the other adjacent pairs, we perform Gaussian elimination and eliminate fully assembled equations (the latter operation not shown below):

\[
\begin{bmatrix}
-2 & 1 & 1 \\
1 & -1 & 0 \\
1 & 0 & -1
\end{bmatrix}
\begin{bmatrix}
B_4 \\
B_5 \\
B_6
\end{bmatrix} =
\begin{bmatrix}
0.5f(c_4) \\
0.5f(c_5) \\
0.5f(c_5)
\end{bmatrix}
\]
We end up again with three 2x2 sub-systems – one for $B_1$ and $B_3$, one for $B_3$ and $B_5$, one for $B_5$ and $B_7$. First two of them, which we consider as adjacent now, can be likewise merged on the next level:

$$\begin{bmatrix} 1 & -1/2 & -1/2 \\ 0 & -1/2 & 1/2 \\ 0 & 1/2 & -1/2 \end{bmatrix} \begin{bmatrix} B_4 \\ B_6 \\ B_8 \end{bmatrix} = \begin{bmatrix} -0.25f(c_4) \\ 0.5f(c_5) + 0.25f(c_4) \end{bmatrix}$$

$$\begin{bmatrix} -2 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} B_6 \\ B_5 \\ B_7 \end{bmatrix} = \begin{bmatrix} 0.5f(c_6) \\ 0.5f(c_5) \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 1/2 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & -1/2 \end{bmatrix} \begin{bmatrix} B_4 \\ B_3 \\ B_1 \end{bmatrix} = \begin{bmatrix} f(c_5) + 0.25f(c_2) + 0.25f(c_4) \\ f(c_5) + 0.25f(c_4) + 0.25f(c_6) \\ 0 \\ f(c_5) + 0.25f(c_4) + 0.25f(6) \end{bmatrix}$$

Again, we can eliminate the fully assembled first row, which yields a system for $B_1$ and $B_5$. The process of partial merging and eliminating continues up to the root of the tree. The process is followed by recursive backward substitutions. For higher order B-splines, the process requires merging more frontal matrices in order to get one fully assembled row, compare figure 3 for fourth order B-splines.

6. IMPLEMENTATION CONCERNS

In the forward elimination phase, the number of nodes is gradually decreasing (about twice with each level) to reach one root node on the topmost level. This means the number of working cores falls down over the time. After each iteration, about half of the cores becomes idle (assuming that each core is assigned to a separate node).

On the contrary, in the backward substitution phase the algorithm starts from a single node (which is the root). This implies only one core is working and the rest remains idle. With each layer, twice as much cores as in the previous layer are assigned a node to process. This can be seen as a flow reverse to the flow of forward elimination.

The bottleneck of the algorithm is the topmost level of the forward elimination and the corresponding topmost level of the backward substitution. Regardless of the size of the input matrix, at this level only one thread can be active and any other cores must wait for the completion.

![Fig. 3. Solver elimination pattern for fourth order B-splines.](image)

![Fig. 4. Comparison of total execution time of ISO-FEM solver and ISO-C solver. Total ISO-FEM solver time = integration + elimination. Total ISO-C solver = elimination time only: quadratic B-splines (a), cubic B-splines (b).](image)
7. NUMERICAL RESULTS AND CONCLUSIONS

The solution presented in figure 1b has been computed by a solver implemented using NVIDIA CUDA SDK 5.5 and run on Fedora 14.

Figures 4(a) and 4(b) point out the amount of time spent just on integration (generation of local systems) in the isogeometric finite element method in comparison to the elimination phase. In other words, switching from ISO-FEM to ISO-C method allows to save a substantial amount of time on the integration phase.

Figures 4 (a) and (b) also show that the execution times of both solvers scale logarithmically when the number of cores is larger than the number of collocation points (ISO-C) or finite elements (ISO-FEM).

Figures 5 (a) and (b) show the common logarithm of relative error depending on the number of collocation points for fourth order B-splines.

In the figure 5(a), the points were chosen according to the Demko scheme. In the figure 5(b), the points on the right side were evenly distributed over the equation’s interval. This is clearly complies with the statement that Demko collocation points guarantee convergence of the method.

ACKNOWLEDGEMENTS

The work presented in this paper has been supported by the National Science Center grant no. 2011/03/B/ST6/01393.

REFERENCES


WIELOFRONTALNY, RÓWNOLEGŁY SOLWER BEZPOŚREDNI DLA JEDNOWYMIAROWEJ METODY KOLOKACJI

Streszczenie


Received: September 21, 2014
Received in a revised form: November 21, 2014
Accepted: November 27, 2014
CRACK ANALYSES IN CONDUCTING AND NON-CONDUCTING PIEZOELECTRIC SOLIDS

JAN SLADEK¹*, VLADIMIR SLADEK¹, MILAN ZMINDAK², SLAVOMIR HRCEK²

¹ Institute of Construction and Architecture, Slovak Academy of Sciences, 84503 Bratislava, Slovakia
² Faculty of Mechanical Engineering, University of Zilina, 01026 Zilina, Slovakia

*Corresponding author: sladek@savba.sk

Abstract

The stress intensity factor and electric displacement intensity factor for cracks in conducting and non-conducting piezoelectric materials is investigated. Transient dynamic crack problems are analyzed. The coupled governing partial differential equations (PDE) for stresses, electric displacement field and electric current are satisfied in a local weak-form on small fictitious subdomains. Local integral equations are derived for a unit function as the test function on circular subdomains. All field quantities are approximated by the moving least-squares (MLS) scheme. The influence of the electric conductivity on the stress intensity and electric intensity factors is shown in numerical examples for an edge crack in a finite strip under a pure mechanical impact load with Heaviside time variation.

Key words: meshless approximation, local integral equations, intensity factors, impermeable boundary conditions

1. INTRODUCTION

Piezoelectric materials (PE) can be either dielectrics or semiconductors. Up to date dielectric materials are more intensively investigated than semiconductors since dielectric materials are described by simpler governing equations. The solution of the boundary value problems for coupled electromechanical problems requires advanced numerical methods due to the high mathematical complexity. Piezoelectric materials are brittle and they have a tendency to develop cracks even in manufacture process. Therefore, it is important to understand fracture of piezoelectric ceramics. Pak (1990) obtained the closed-form solutions for an infinite PE medium under an anti-plane loading by using a complex variable method. Later, Park and Sun (1995) obtained closed-form solutions for all the three fracture modes associated with a crack in an infinite PE medium. They investigated the effects of the electric field on the fracture of PE ceramics.

General computational methods like the finite element method (FEM) (Gruebner et al., 2003; Govorukha & Kamlah, 2004; Enderlein et al., 2005; Kuna, 2006) and the boundary element method (BEM) (Pan, 1999; Davi & Milazzo, 2001; Gross et al., 2005; Garcia-Sanchez et al., 2005, 2007) need to be applied for general crack analyses in PE solids. In recent years, meshless formulations are becoming popular due to their high adaptability and low costs in preparation of input and output data for numerical analyses. Even continuously varying PE material properties are considered in some numerical analyses for non-conducting dielectric PE (Sladek et al., 2007).

In piezoelectric semiconductors (conducting PE) the induced electric field produces also the electric current. The interaction between mechanical fields and mobile charges in piezoelectric semiconductors is called the acoustoelectric effect (Hutson & White, 1962; White, 1962). An acoustic wave traveling in a PE semiconductor can be amplified by application...
of an initial or biasing dc electric field (Yang & Zhou, 2005). There are only few papers devoted to crack problems in piezoelectric semiconductor materials. These papers concerned only the anti-plane crack problem in unbounded domain with a semi-infinite crack (Yang, 2005) or a finite crack (Hu et al., 2007) under stationary conditions. The Fourier transform technique was applied to reduce the problem to a pair of dual integral equations.

In the present paper, we aim at analyzing the in-plane crack problem in bounded domains under a mechanical and electric load. Static and transient boundary conditions are considered here. The meshless Petrov-Galerkin (MLPG) method (Sladek et al., 2013) is developed for the solution of the initial-boundary value problems in conducting piezoelectric solids. Nodal points are introduced and spread on the analyzed domain and each node is surrounded by a small circle for simplicity, but without loss of generality. The spatial variations of the displacement, electric potential and electron density are approximated by the moving least-squares (MLS) scheme. After performing the spatial integrations, a system of algebraic equations is obtained. The essential boundary conditions on the global boundary are satisfied by the collocation.

2. LOCAL INTEGRAL EQUATIONS AND NUMERICAL SOLUTION

Consider a homogeneous piezoelectric semiconductor with electron density $M_0$ in the unloaded state and vanishing initial electric field $E_0$. The quasi-static approximation can be supposed for electromagnetic fields, since the frequency of external loadings is assumed to be significantly lower than the characteristic frequency of the electromagnetic fields. Then, the governing equations within the linear theory are given by the balance of momentum, Gauss’s law and conservation of the electric charge (Hutson & White, 1962)

$$\begin{align*}
\sigma_{ij}(x, \tau) &= \rho \ddot{u}_i(x, \tau), \\
D_{ij}(x, \tau) &= qM_{ij}(x, \tau), \\
qM(x, \tau) + J_{ij} &= 0,
\end{align*}
$$

where $\dot{u}_i$, $\sigma_{ij}$, $D_{ij}$, and $q$ are the acceleration of elastic displacements, stress tensor, electric displacement field, and electric charge of electron, respectively. The electron density and electric current are denoted by $M$ and $J_{ij}$, respectively. Symbol $\rho$ is used for the mass density.

The constitutive equations are given as (Hutson & White, 1962; White, 1962)

$$\begin{align*}
\sigma_{ij}(x, \tau) &= c_{ijkl}E_{kl}(x, \tau) - e_{ij}E_i(x, \tau), \\
D_{ij}(x, \tau) &= e_{ijkl}E_{kl}(x, \tau) + h_{ij}E_i(x, \tau), \\
J_{ij}(x, \tau) &= qM_{ij}(x)E_i(x, \tau) - qd_{ij}M_j(x, \tau),
\end{align*}
$$

where $c_{ijkl}$, $e_{ijkl}$, $h_{ij}$, $\mu_{ij}$ and $d_{ij}$ are the elastic, piezoelectric, dielectric, electron mobility and carrier diffusion material coefficients, respectively.

The strain tensor $e_{ij}$ and the electric field vector $E_i$ are related to the displacements $u_i$ and the electric potential $\phi$ by

$$e_{ij} = \frac{1}{2}(u_{ij} + u_{ji}), \quad E_i = -\phi_i. \quad (3)$$

According to the meshless local Petrov-Galerkin (MLPG) method, we construct a weak-form of (1) over the local subdomains $\Omega_{s}$ around each node $x^s$ inside the global domain $\Omega$ (Sladek et al., 2013). The subdomains are distributed in the analyzed domain. The local subdomains could be of any geometrical shape and size. For the sake of simplicity, the local subdomains are taken here to be of a circular shape. For this case, the evaluation of domain-integrals is quite easy. The local weak form of the governing equations (1) can be written as

$$\int_{\Omega_s} \left[ \sigma_{ij}(x, \tau) - \rho \ddot{u}_i(x, \tau) \right] u^*_i(x) d\Omega = 0, \quad (4)$$

where $u^*_i(x)$ is a test function and $\Omega_s \subset \Omega$.

Applying the Gauss divergence theorem to the first integral and choosing the Heaviside step function as the test function $u^*_i(x)$ in each subdomain

$$u^*_i(x) = \begin{cases} 
\delta_{is} & \text{at } x \in \Omega_s \\
0 & \text{at } x \notin \Omega_s 
\end{cases}$$

the local weak-form (4) is converted into the following local boundary-domain integral equations

$$\int_{L_s \cup \Gamma_{ni}} t_i(x, \tau) d\Gamma - \int_{\Omega_s} \rho \ddot{u}_i(x, \tau) d\Omega = \int_{\Gamma_{ns}} \bar{f}_i(x, \tau) d\Gamma, \quad (5)$$

where the boundary of the local subdomain $\partial \Omega_s$ consists of three parts $\partial \Omega_s = L_s \cup \Gamma_{ni} \cup \Gamma_{ns}$ (Sladek et al., 2013). Here, $L_s$ is the local boundary that is totally inside the global domain, $\Gamma_{ns}$ is the part of the local boundary which coincides with the global trac-
tion boundary, i.e., $\Gamma_{su} = \partial \Omega_s \cap \Gamma_v$, and $\Gamma_{su}$ is the part of the local boundary that coincides with the global displacement boundary, i.e., $\Gamma_{su} = \partial \Omega_s \cap \Gamma_u$.

Similar definitions are valid also for other fields and related integration parts.

The local integral equation (5) is valid for both the homogeneous and nonhomogeneous solids. Nonhomogeneous material properties are included in equation (5) through the elastic and piezoelectric coefficients involved in the traction components:

$$t_j(x, \tau) = \left[ e_{ijkl} u_{k,j} (x, \tau) + e_{ijkl} \phi_k (x, \tau) \right] n_j (x).$$

The local integral equation corresponding to the second governing equation in (1) has the following form:

$$\int_{\Gamma_a} Q(x, \tau) d\Gamma - \int_{\Omega} q M(x, \tau) d\Omega = -\int_{\Gamma} \tilde{Q}(x, \tau) d\Gamma$$

where

$$Q(x, \tau) = D_j(x, \tau) n_j(x) = \left[ e_{ijkl} u_{k,j} (x, \tau) - h_{ik} \phi_k (x, \tau) \right] n_j.$$

Finally, the local integral equation corresponding to the last governing equation in (1) has the form:

$$\int_{\Gamma_a} S(x, \tau) d\Gamma + \int_{\Omega} q M(x, \tau) d\Omega = -\int_{\Gamma_a} \tilde{S}(x, \tau) d\Gamma$$

where the electric current flux is given by

$$S(x, \tau) = J_J(x, \tau) n_j(x) =$$

$$\left[ -q M_{a,ik} \phi_k (x, \tau) - q a_{ik} M_{a,ik} (x, \tau) \right] n_j.$$  

In the present paper the trial functions are approximated by the moving least squares (MLS) method on a number of nodes spread over the influence domain. According to the MLS (Sladek et al., 2013) method, the approximation of physical fields $f(x, \tau)$ (mechanical displacements, the electric potential and electron density) over a number of randomly located nodes $\{x^a\}$, $a = 1, 2, \ldots, n$, is given by

$$f(x, \tau) = \pi^T(x) a(x, \tau),$$

where $\pi^T(x) = [\pi^1(x), \pi^2(x), \ldots, \pi^n(x)]$ is a complete monomial basis of order $m$; and $a(x, \tau)$ is a vector containing the coefficients $a^j(x, \tau)$, $j = 1, 2, \ldots, m$ and $x \equiv (x_1, x_2)$. The coefficient vector $a(x, \tau)$ is determined by minimizing a weighted discrete $L_2$-norm defined as

$$J(x) = \sum_{a=1}^n w^a(x) \left[ \pi^T(x) a(x, \tau) - \tilde{f}^a(\tau) \right]^2,$$

where $n$ is the number of nodes used for the approximation. It is determined by the weight function $w^a(x)$ associated with the node $a$. The stationarity of $J$ with respect to $a(x, \tau)$ leads to the following linear relation between $a(x, \tau)$ and $\tilde{f}(\tau)$:

$$A(x) a(x, \tau) - B(x) \tilde{f}(\tau) = 0,$$

where

$$A(x) = \sum_{a=1}^n w^a(x) \pi(x^a) \pi^T(x^a),$$

$$B(x) = \left[ w^1(x) \pi(x^1), w^2(x) \pi(x^2), \ldots, w^n(x) \pi(x^n) \right].$$

The solution of equation (9) for $\tilde{a}(x, s)$ and the subsequent substitution into equation (8) lead to the following expression:

$$u^h(x, \tau) = \sum_{a=1}^n N^a(x) \tilde{u}^a(\tau),$$

$$\phi^a(x, \tau) = \sum_{a=1}^n N^a(x) \phi^a(\tau),$$

$$M^a(x, \tau) = \sum_{a=1}^n N^a(x) \tilde{M}^a(\tau),$$

where the nodal values, $\tilde{u}^a(\tau) = \left( \tilde{u}^a_1(\tau), \tilde{u}^a_2(\tau) \right)^T$, $\phi^a(\tau)$, and $\tilde{M}^a(\tau)$, are fictitious parameters for the displacements, electric potential and electron density, respectively, and

$$N^T(x) = \pi^T(x) A^{-1}(x) B(x).$$

$N^a(x)$ is the shape function associated with the node $a$. A 4th-order spline-type weight function is applied in the present work

$$w^a(x) = \begin{cases} 1 - 6 \left( \frac{d^a}{r^a} \right)^2 + 8 \left( \frac{d^a}{r^a} \right)^3 - 3 \left( \frac{d^a}{r^a} \right)^4, & 0 \leq d^a \leq r^a \\ 0, & d^a \geq r^a \end{cases},$$

where $d^a = \|x - x^a\|$ and $r^a$ is the size of the support domain.
Then, the traction vector \( t^a(x, \tau) \) at a boundary point \( x \in \partial \Omega_s \) is approximated in terms of the nodal displacements \( \hat{u}^a(\tau) \) and electric potentials \( \hat{\phi}^a(\tau) \) as

\[
t^a(x, \tau) = N(x)C(x) \sum_{a=1}^{t} B^a(x) \hat{u}^a(\tau) + N(x)L(x) \sum_{a=1}^{n} P^a(x) \hat{\phi}^a(\tau),
\]

where the matrices \( C(x), L(x) \) represent elastic constants and piezoelectric coefficients, respectively, the matrix \( N(x) \) is related to the normal vector \( n(x) \) on \( \partial \Omega_s \) by

\[
N(x) = \begin{bmatrix} n_1 & 0 & n_3 \\ 0 & n_3 & n_1 \end{bmatrix},
\]

and finally, the matrices \( B^a \) and \( P^a \) represent the gradients of the shape functions as

\[
B^a(x) = \begin{bmatrix} N_{a1}^a & 0 \\ 0 & N_{a2}^a \\ N_{a3}^a & N_{a4}^a \end{bmatrix}, \quad P^a(x) = \begin{bmatrix} N_{a1}^a \\ N_{a2}^a \end{bmatrix}.
\]

Symbols \( n_1(x) \) and \( n_3(x) \) are components of the normal vector \( n(x) \) to the boundary \( \partial \Omega_s \) at the Gaussian point \( x \) in plane \( x_1 - x_3 \).

Similarly one can write

\[
Q^a(x, \tau) = N^a(x)G(x) \sum_{a=1}^{t} B^a(x) \hat{u}^a(\tau) - N^a(x)H(x) \sum_{a=1}^{n} P^a(x) \hat{\phi}^a(\tau),
\]

where the matrices \( G(x) \) and \( H(x) \) represent piezoelectric and dielectric coefficients and

\[
N^a(x) = \begin{bmatrix} n_1 \\ n_3 \end{bmatrix}.
\]

Finally, the electric current flux \( S(x, \tau) \) is approximated by

\[
S^a(x, \tau) = -N^a(x)qM_0A(x) \sum_{a=1}^{t} P^a(x) \hat{\phi}^a(\tau) - N^a(x)qF(x) \sum_{a=1}^{n} P^a(x) \dot{\hat{\phi}}^a(\tau) - N^a(x)qF(x) \sum_{a=1}^{n} P^a(x) \dot{\hat{\phi}}^a(\tau),
\]

with the matrices \( A(x), F(x) \) denoting electron mobility and carrier diffusion parameters.

Substituting the traction approximation (11), electric charge (12) and electric current (13) into the local integral equations (5)-(7) one obtains a system of ordinary differential equations for nodal quantities

\[
\sum_{a=1}^{n} \left[ \int_{\Omega_s} N(x)C(x)B^a(x)d\Gamma \right] \hat{u}^a(\tau) - \left[ \int_{\Omega_s} \rho(x)N^a d\Omega \right] \hat{u}^a(\tau) + \sum_{a=1}^{n} \left[ \int_{\Omega_s} N(x)L(x)P^a(x)d\Gamma \right] \dot{\hat{\phi}}^a(\tau) = -\int_{\Gamma_{s}} \tilde{i}(x, \tau)d\Gamma \tag{14}
\]

\[
\sum_{a=1}^{n} \left[ \int_{\Gamma_{s}} N^a(x)G(x)B^a(x)d\Gamma \right] \hat{u}^a(\tau) - \sum_{a=1}^{n} \left[ \int_{\Gamma_{s}} N^a(x)H(x)P^a(x)d\Gamma \right] \dot{\hat{\phi}}^a(\tau) - \sum_{a=1}^{n} \left[ \int_{\Omega_s} qN^a(x)d\Omega \right] \dot{\hat{\phi}}^a(\tau) - \int_{\Gamma_{s}} \tilde{Q}(x, \tau)d\Gamma = \int_{\Gamma_{s}} \tilde{M}^a(\tau) \tag{15}
\]

\[
\sum_{a=1}^{n} \left[ \int_{\Omega_s} N^a(x)qM_0A(x)P^a(x)d\Gamma \right] \dot{\hat{\phi}}^a(\tau) + \sum_{a=1}^{n} \left[ \int_{\Omega_s} qN^a(x)d\Omega \right] \ddot{\hat{\phi}}^a(\tau) = -\int_{\Gamma_{s}} \tilde{S}(x, \tau)d\Gamma \tag{16}
\]

Above given system of ordinary differential equations is solved numerically by the Houbolt method.
The stress intensity factors (SIF) and electric displacement intensity factor (EDIF) $K_N$ are computed numerically from the asymptotic expansion of displacements and electric potential at the crack tip vicinity

$$u_i(r, \theta) = \sqrt{\frac{2r}{\pi}} \sum_{N=1}^{4} K_N d_i^N(\theta),$$

$$\phi(r, \theta) = \sqrt{\frac{2r}{\pi}} \sum_{N=1}^{4} K_N \nu^N(\theta),$$

where $d_i^N(\theta)$ and $\nu^N(\theta)$ are dependent on material properties and they are given by Park and Sun (1995).

3. NUMERICAL EXAMPLES

An edge crack in a finite strip is analyzed in the first example. Due to the symmetry with respect to $x_1$ only a half of the specimen is modeled with the boundary conditions and discretization nodes being shown schematically in figure 1. The geometry parameters are specified as: $a = 0.5$ m, $a/w = 0.4$ and $h/w = 1.2$. The material properties correspond to aluminum nitride (AlN) (Auld, 1973).

Electrically impermeable boundary conditions are assumed on crack faces. The strip is subjected to a pure mechanical load with Heaviside time variation and the intensity $\sigma_0 = 1$ Pa.

![Fig. 1. Edge crack in a finite homogeneous strip.](image1)

We have used 930 nodes equidistantly distributed for the MLS approximation of physical fields. The static stress intensity factor for the considered load and geometry is equal to $K_{I_{stat}} = 2.642$ Pam$^{1/2}$. The time evolution of the normalized SIF for the cracked strip under an impact pure mechanical load is presented in figure 2 and the normalized electrical displacement factor $\Lambda K_{I_{stat}}/K_{I_{stat}}$ in figure 3, where $\Lambda = e_{33}/h_{33}$. While the electrical displacement intensity factor for a pure static mechanical load is zero, the EDIF is not in the dynamic case with a finite velocity of elastic wave propagation for a pure mechanical load.

![Fig. 2. Normalized stress intensity factor for the edge crack in a strip under a pure mechanical load.](image2)

One can observe that the initial electron density has a vanishing influence on the SIF. However, the EDIF is strongly dependent on the initial electron density. In the conducting PE the EDIF is enlarged with respect to that in non-conducting PE under a pure mechanical load. Accuracy of numerical results for non-conducting PE solids was tested in earlier published work (Sladek et al. 2007).

![Fig. 3. Normalized electrical displacement intensity factor for the edge crack in a strip under a pure mechanical load.](image3)

ACKNOWLEDGEMENT

The authors gratefully acknowledge the supports by the Slovak Science and Technology Assistance Agency registered under number APVV-0014-10.
REFERENCES


ANALIZA PROCESU PĘKANIA W PRZEWODZĄCYCH I NIEPRZEWODZĄCYCH MATERIAŁACH PIEZOELEKTRYCZNYCH

Streszczenie

W pracy badano współczynnik intensywności naprężeń i współczynnik intensywności przemieszczeń elektrycznych dla pękania w przewodzących i nieprzewodzących materiałach piezoelektrycznych. Analizowano problemy zmiennego, dynamicznego pękania. Sprzężone równania różniczkowe cząstkowe dla naprężeń, pola przemieszczeń elektrycznych i prądu elektrycznego zostały spełnione poprzez wprowadzenie słabiej formy w małych, urojonych podobszarach. Lokalne równania całkowe zostały definiowane dla funkcji jednostkowej będącej funkcją testową w podobszarach. Wszystkie wielkości opisujące pole były aproksymowane z wykorzystaniem ruchomej metody najmniejszych kwadratów. Wpływ przewodnoci elektrycznej na intensywność naprężenia i współczynnik intensywności elektrycznej został pokazany w przykładach numerycznych pękania krawędzi w paśmie, które zostało poddane czystemu obciążeniu mechanicznemu będącemu funkcją czasu Heaviside’a.

Received: September 23, 2014
Received in a revised form: October 9, 2014
Accepted: December 17, 2014
APPLICATION OF THE FULLY AUTOMATIC
HP-ADAPTIVE FEM TO ELASTIC-PLASTIC PROBLEMS

MARTA OLEKSY*, WITOLD CECOT

Institute for Computational Civil Engineering, Cracow University of Technology,
ul. Warszawska 24, 31-155 Kraków, Poland
*Corresponding author: moleksy@l5.pk.edu.pl

Abstract

The hp-adaptive mesh refinement is a technique that for linear problems proved to deliver a fast, predicted by theory, exponential convergence. We have applied this approach to elastic-plastic problems modeled by associative constitutive law with the Mises yield surface. A modification of the automatic refinements and various types of mesh adaptation were tested. Exponential convergence was observed when the FEM mesh complied with the yielding zone. In general the original automatic hp-mesh refinements led to the fastest convergence even though it was only algebraic, what is theoretically justified for elastic-plastic problems.

Key words: fully automatic hp-adaptivity; elastic-plastic problem; exponential convergence

1. INTRODUCTION

Since the eighties of the 20th century both h and p adaptive mesh refinements were used by many researchers to improve convergence of the FEM solutions for inelastic problems. Cheng (1988) as well as Zienkiewicz and coworkers (Zienkiewicz et al., 1990) developed this approach to metal forming processes, Peric, Dutko and Owen (Peric et al., 1996; Peric et al. 1998) presented adaptive FEM solutions for large strain plasticity, Cramer et al. (1999) applied partitioning of elements to associative and non-associative plasticity. Ladeveze and Moes (1999) proposed a posteriori error estimation, to control time-dependent nonlinear finite element analysis. They used the error and the three indicators in an adaptive strategy (to adapt the mesh, the time and to limit the number of iterations of the global iterative algorithm). Johnson and Hansbo (Johnson et al., 1992) have developed adaptive strategy, based on residual error estimator, for small strain elastoplasticity using the Hencky model. Adaptive mesh refinement procedure, based on a-posteriori estimate, that takes into account the local directional interpolation error and a recovering technique to compute second derivatives of the finite element solution was proposed by Borges and coworkers (Borges et al., 2001) for limit analysis. This approach is able to capture discontinuities arising from localized plastic deformations during plastic collapse. The p-adaptive FEM was reported to be efficient approximation to physically nonlinear problems by Düster and Rank (2002) since elastic-plastic solution may be highly regular in interior of both elastic and plastic subdomains. The rp-adaptive approach was used by Nübel and coworkers (Nübel et al., 2007) to adjust finite element mesh to elastic-plastic border in order to take into account the loss of regularity and consequently observed exponential convergence rate for deformation theory of plasticity.

In this work we examined convergence of the fully automatic hp strategy, developed by Demkowicz and coworkers (Demkowicz, 2006; Dem-
This method leads to the exponential convergence for linear problems (Demkowicz et al., 2002; Gui & Babuška, 1986). In order to obtain an efficient method for elastic-plastic computational homogenization of metal matrix composites (Serafin & Cecot, 2010). We have applied the self adaptive hp-FEM to inelastic problems. Even though the elastic-plastic solutions are, in general, not analytic almost everywhere (Szabo et al., 2004) and consequently exponential convergence for hp-FEM is not guaranteed, one may still expect a faster convergence of hp than exclusively h or p mesh refinements.

2. FORMULATION OF THE PROBLEM

We consider here quasi-static elastic-plastic processes under small displacements and strains for polycrystalline materials with linear, kinematic hardening. Such processes may be modeled by associative plasticity with certain yield function (denoted by \( \Phi \)), resulting in the following initial boundary value problem defined in an open bounded domain \( \Omega \subset \mathbb{R}^n \), \( n = 1, 2 \) or 3 and pseudo-time interval \([0, T]\), where \( \Omega \) and \( \partial \Omega \) are elastic and plastic subdomains (zones):

\[
\text{find sufficiently regular fields of displacements } u(x; t), \text{ strains } \varepsilon(x; t), \text{ inelastic strains } \varepsilon^p(x; t), \text{ plastic multiplier } \gamma(x; t) (\gamma(x; t) = 0 \text{ in } \Omega) \text{ and stresses } \sigma(x; t) \text{ such that:}
\]

\[
\begin{align*}
\text{div } \sigma &= 0 \quad \forall x \in \Omega_0 \cap \Omega_p, \forall \tau \in [0, T] \\
\dot{\varepsilon} &= \frac{1}{2} \left[ \nabla u + (\nabla u)^T \right] \quad \forall x \in \Omega_0 \cap \Omega_p, \forall \tau \in [0, T] \\
\sigma &= C(\dot{\varepsilon} - \dot{\varepsilon}^p) \quad \forall x \in \Omega_0 \cap \Omega_p, \forall \tau \in [0, T] \\
\dot{\varepsilon}^p &= \frac{\gamma}{\sigma} \frac{\partial \Phi}{\partial \sigma} \quad \forall x \in \Omega_0 \cap \Omega_p, \forall \tau \in [0, T] \\
\gamma &\geq 0, \Phi \leq 0, \phi = 0 \quad \forall x \in \Omega_0 \cap \Omega_p, \forall \tau \in [0, T] \\
\dot{u} &= \dot{u} \quad \forall x \in \Omega_0 \cap \Omega_p, \forall \tau \in [0, T] \\
\sigma \dot{n} &= \dot{t} \quad \forall x \in \Omega_0 \cap \Omega_p, \forall \tau \in [0, T] \\
\varepsilon^p &= \varepsilon_0^p, u = u_0 \quad \forall x \in \Omega_0 \cap \Omega_p, \forall \tau \in [0, T]
\end{align*}
\]

(1)

where, as usual, the Lipschitz boundary \( \partial \Omega = \partial \Omega_D \cup \partial \Omega_N \) and \( \partial \Omega = \partial \Omega_D \cap \partial \Omega_N = \emptyset. \) For \( n > 1 \) the measure of Dirichlet boundary is greater than zero (\( \text{meas} \partial \Omega_D > 0 \)). We also assume that the initial conditions are compatible with both loading and plastic strains. Furthermore, continuity of displacement and stress vectors holds on the elastic-plastic interface at every time instant. Integration in time of problem (1) is usually performed by the backward Euler scheme with return mapping corrector (see e.g. Simo & Hughes, 1998). After converting the rate formulation into the incremental form, the total load history is represented by a sequence of sub-loads. For each sub-load the solution increment is evaluated iteratively by requiring at each Gauss integration point fulfillment of the yield condition and overall momentum equations at the end of each load increment. The well known algorithm makes use of the decomposition of the stress increment into two parts: the elastic trial component (\( \sigma^{trial} \)) and the plastic corrector component (\( \Delta \sigma = -C \Delta \varepsilon^p \)) evaluated by the return mapping method (Simo & Taylor, 1986) (radial return algorithm for the Mises yield function). The trial part is computed, for fixed plastic strain increment \( \Delta \varepsilon^p \), by the momentum equation (1), combined with equations (1)_{2,3}. Then the plastic flow rule and the yield condition constitute, in the case of constant hardening, the following nonlinear system of equations that is used to define increments of the corrector parts of stress (\( \Delta \sigma \)) and the plastic multiplier (\( \Delta \gamma \))

\[
\begin{align*}
C^{-1} \Delta \sigma + \Delta \gamma \frac{\partial \Phi}{\partial \sigma} &= 0 \\
\Phi(\sigma^{trial} + \Delta \sigma) &= 0
\end{align*}
\]

(2)

The system of equations (2) is, in general, solved by the Newton-Raphson method. Usually, in order to assure quadratic convergence of the iterative process, the so-called algorithmic tangent modulus is suggested instead of tangent one.

3. AUTOMATIC MESH ADAPTATION

The automatic mesh adaptation proposed in (Demkowicz et al., 2002) was successfully used for various linear problems. Its key idea is an appropriate strategy of anisotropic \( h, p \) or \( hp \) mesh refinement. It is based on the interpolation error estimate, which is a good upper bound of the best approximation error that in turn, for coercive problems by the Cea’s lemma, is the upper bound for the actual approximation error. The aforementioned interpolation error is estimated by making use of a fine mesh solution (\( \hat{u}_{h/2,p+1} \), denoted here for the sake of brevity by \( u \)) that serves as a substitute for the exact solution. Such an “exact” solution is interpolated locally on possible new \( hp \)-refined meshes. The difference between \( u \) and its interpolant approximates the inter-
polation error and the optimal anisotropic mesh refinement is that one for which the reduction of the interpolation error per number of additional degrees of freedom is maximal. It means, that for the coarse mesh the optimal \((h, p \text{ or } hp)\) refinement is determined by maximizing the following expression

\[
r = \frac{|u - \Pi_{hp}u|_{H^1}^2}{N_0 - N_c} \rightarrow \max (3)
\]

with additional assumption, that the mesh is oneirregular, where \(hp, h_{pop}\) denote \(H^1\) projection-based interpolants on the current and optimal meshes, respectively; \(N_0, N_c\) are the numbers of degrees of freedom in optimal and current meshes. The maximization is performed by search over a suitable subset of all possible \(hp\) refinements. Thus, the algorithm of adaptation approach starts with the solution of the problem on the current (coarse) mesh \((u_{hp})\). Then, the refinement in both \(h\) and \(p\) is performed and the optimal mesh is selected by maximization of the function \(r\) defined by equation (3). For large problems computation of the fine mesh solution may be time consuming. However, only partially convergent solution obtained by e.g. a fast two-grid solver may be used to guide the optimal \(hp\)-refinement.

For elastic-plastic problems accuracy of integration in semi-time \(\tau\) depends primarily on accuracy of stress approximation in space since both plastic multiplier and plastic strain rate are determined by stress rate (equation (1)4, (1)5). Incremental application of loading reduces number of iterations and prevents numerical solution from coming to far away from the true one. Since solution \(u_{h/2p+1}\) uses significantly greater number of degrees of freedom therefore, the error of plastic strain approximation is also taken into account by the error estimate used in the automatic \(hp\)-adaptivity algorithm. Consequently, one may expect a fast exponential convergence also for elastic-plastic problems. In order to obtain appropriate stress approximation accuracy, inelastic deformations should be accounted for in a special way in a-posteriori error estimates (Barthold et al., 1998; Gallimard et al., 1996; Nübel et al., 2007; Peric et al., 1994). Therefore, we have studied performance of the original and modified automatic \(hp\)-mesh refinement for elasticplastic problems. The objective of the modification was an additional \(h\)-refinement of the mesh along the elastic-plastic interface, which is a place of lower solution regularity.

4. NUMERICAL TESTS – 1D EXAMPLES

The first (1D) numerical tests were performed for a bar subject to axial body force.

\[
q(x) = -\sin(2\pi x) 2\sigma_0 x 
\]

For the considered example the following material properties were assumed: Young modulus \(E = 200\) GPa, Poisson ratio \(v = 0.3\), yield stress limit \(\sigma_0 = 200\) MPa and hardening parameter \(H = 0.2E\). The second derivative of the exact solution is not continuous at the elastic-plastic interface thet for this data is located at points \(x = 0.25, x = 0.75\). First, \(p\)-stability for presented example was verified and the results presented in figure 1 indicate, that \(p\)-enrichment is reasonable for our problem since it reduces the error. The main objective of the tests was verification whether the fully automatic \(hp\)-refinements should be complemented with additional \(h\)-refinements (or \(p\)-enrichment) in vicinity of the elastic-plastic zone. The asymptotic behavior of the solution below the error level of \(10^{-6}\) is a result of the tolerance assumed for meeting the yield condition in the radial return algorithm. Therefore, in all the examples considered in this paper the analysis was performed at most as long as this error level was reached.

Since the rate of convergence is much better for the meshes that initially comply with the elastic-plastic interface, the numerical analysis with two initial meshes was performed in order to examine the adaptation process for inelastic problems. In the first case the initial finite element mesh complied with a-priori known elastic and plastic zones, while in the second case initial mesh was independent of...
yielding. In the second case we have observed, that after few automatic adaptation steps the mesh accommodated to elastic-plastic interfaces at points \( x = 0.25 \) and \( x = 0.75 \) (first 1D example), as may be observed in figure 2. Convergences of the error norm for various refinement strategies are compared in figures 3, 4. One may observe that in the first case, i.e. with initially detected elastic-plastic interface, the results are better.

To make an example more realistic we assumed the elastic-plastic interfaces at points \( x = 0.23456789 \) and \( x = 0.703703670 \) (second 1D example). In this case only meshes that did not comply with elastic plastic interface were used. The corresponding plastic strain distribution and the resulting mesh refinements are presented in figure 5. In this example the \( p \)-stability was also observed (figure 6) and convergence of error norm for different refinements is presented in figure 7.
In the 1D examples additional h-refinements maybe profitable for convergence rate. However, since 1D problems may exhibit super convergence properties the further tests were performed for 2D examples.

5. NUMERICAL TESTS – 2D EXAMPLES

The following 2D examples were examined:
1. thick-walled cylinder
2. L-shaped domain
3. perforated plate
4. non-perforated plate 1
5. non-perforated plate 2

All these examples may be formulated in the following weak form: find field of displacements \( u(x; t) \) such that for every \( t \in [0; T] \)
\[
\int_\Omega \varepsilon(v) : C \varepsilon(u) d\Omega = \int_\Omega \varepsilon(v) : \sigma d\Omega + \int_{\Gamma_x} \hat{t} v ds \forall v \in V_0
\]
where \( V_0 = \{ v \in [H^1(\Omega)]^n, v = 0 \text{ on } \partial \Omega_D \}, \partial \Omega_D \) and \( \partial \Omega_N \) are the Dirichlet and Neumann parts of the boundary, \( \partial \Omega_D \cup \partial \Omega_N = \partial \Omega \) and \( \partial \Omega_D \cap \partial \Omega_N = \emptyset \).
\( \hat{t}, \hat{u} \) are known tractions and displacements along the Neumann and Dirichlet parts of the boundary, \( C \) stands for elastic tensor of material parameters. For each example the following data were assumed:
Young modulus \( E = 200 \text{ GPa} \), Poisson ratio \( \nu = 0.3 \), yield limit \( \sigma_0 = 200 \text{ MPa} \), hardening parameter \( H = 0.1E \). In 2D examples additional h-refinements for elements with both elastic and plastic zones were performed in order to fit better to the zones. These additional refinements were performed by considering four closest to vertices Gauss integration points. If only at two of those points located near one common edge (marked by the red color in figure 8) material yields, then the element is partitioned in the direction perpendicular to that common edge. Otherwise, if yielding is observed at some, but not all of Gauss points, the uniform h-refinement is performed.

5.1. Thick-walled cylinder

We assumed plane strain state for a quarter of a thick-walled cylinder with radii 1 m and 2 m. It was loaded by internal pressure \( p = 120 \text{ MPa} \), that resulted in yielding at points in distance smaller then -1.226 m from the center. The finite element mesh initially did not comply with the elastic-plastic interface and consisted of 8 second order elements (2 in the radial and 4 in the hoop directions). Automatic hp-adaptation was performed in the standard way (FE mesh after few steps of adaptations is shown in figure 9) and it was compared with some other possible refinements (figure 12). We assumed the same as in figure 2 meaning of colors in all figures that present FEM meshes. In one of them hp-adaptation was augmented with additional h-refinements (figure 10). The proposed modification, based on additional refinements, resulted in better convergence and higher order approximation in the elastic zone. For comparison, finite element mesh, which complies with elastic-plastic interface was assumed (figure 11) and automatically refined. As one may expect in such a situation the best convergence was observed.

Fig. 8. Cases of additional anisotropic element refinement.

Fig. 9. Cylinder test. Mesh after 20 steps of hp-refinements (colors indicate order of approximation).

5.2. L-shaped domain

Another classical test, i.e. L-shaped domain with singular derivatives of solution at the reentrant corner was also considered. We assumed plane strain state, fixed boundary on the right-hand side and loading presented in figure 13. The finite element meshes obtained by various adaptation strategies are shown in figure 14. The elastic-plastic interface was successfully modeled by additional h-refinements of
the partially yielded elements. In this example such a modification of adaptation process did not improve error convergence and the best one was observed for the original algorithm (figure 15). However, the number of necessary adaptation steps, that guarantees results with assumed accuracy, was reduced whenever the additional h-refinements were used.

Fig. 10. Cylinder test. Mesh after 14 steps of modified hp-refinements.

Fig. 11. Cylinder test. Initial mesh and after 13 steps of hp-refinements.

Fig. 12. Cylinder test. Convergence of error norm. A – uniform refinement, B – h-adaptation, C – original automatic hp-adaptation, D – automatic hp-adaptation with additional h-refinements of elements with both elastic and plastic zones, E – automatic hp-adaptation for mesh, which initially complies with known elastic-plastic interface.

5.3. Perforated plate

A quarter of a perforated plate in plane strain state with constant loading was analyzed as a next test. Assumed boundary conditions and loading are shown in figure 16. The next figures present meshes obtained by fully automatic refinement of $hp$ and $h$-type (figure 17), as well as after additional $h$-refinements along elastic-plastic interface (figure 18). This time, as it can be observed in the plot of convergence history, the rate of convergence was not improved by additional enforcement of mesh accommodation to plastic zone, but the number of refinement steps was two times smaller.
5.4. Non-perforated plate 1

In this example elastic-plastic deformations resulted in a rectangular plate \((1m \times 2m)\) from body forces assumed in the form:

\[
f_x = -\frac{\sigma_0 \pi}{2} \sin \left(\frac{\pi x}{4}\right), \quad f_y = -\frac{\sigma_0 \pi}{2} \cos \left(\frac{\pi y}{4}\right)
\]  

Such a loading was assumed in order to be able to generate various shapes of elastic-plastic interfaces. Symmetry along left and bottom edges, as well as traction-free boundary conditions on the remaining edges were assumed. The elastic and plastic zones convergence history, as well as the meshes obtained by various strategies are shown in figures 19, 20. Convergence history for the modified automatic mesh adaptation was slightly worse than for the original version. This time the number of necessary refinements was also reduced.

5.5. Non-perforated plate 2

The plate, fixed on the left edge and loaded on the right-hand side, was the next test (figure 21). A plane strain state was assumed. The results are presented in figures 22, 23. We may observe, for this elastic-plastic problem with singularity of solution, that the convergence rate for the original automatic \(hp\)-adaptation is much better than for the version with additional \(h\)-refinements.

6. SOLUTION TRANSFER

In all the examples described in this paper loading was applied in one step. Therefore, there was no need for solution transfer. However, if more steps were used the solution transfer would have to be done and it might be performed comparatively easily since the mesh is refined by subdivision of the elements (Wunderlich et al., 1998). Therefore, each new element is a part of only one old element. After a mesh refinement or unrefinement the time integration is performed on a new set of the Gauss points and accuracy of the space approximation changes. We propose the strategy that assumes a reanalysis of the actual load increment. Whenever the mesh is refined at the end of the load increment...
the second run over this increment should serve the purpose of solving the problem with specified accuracy in order to avoid the situation when large error in one time instant, that results from too coarse space discretization, influences the accuracy in the later time instances.

Fig. 21. Non-perforated plate 2. Boundary conditions. Elastic and plastic subdomains.

Fig. 22. Non-perforated plate 2. Mesh after 15 steps of hp-refinements and 14 steps of modified hp-refinements.

Fig. 23. Non-perforated plate 2. Convergence history (A – automatic hp-adaptation with additional h-refinements, B – original automatic hp-adaptation).

7. CONCLUDING REMARKS

The fully automatic (self automatic) hp-adaptive mesh refinement strategy was applied to analysis of elastic-plastic problems. Since the solution is less regular at the elastic-plastic interface (Nübel et al., 2007) the finite element meshes should comply with elastic and plastic zones. However, the elastic and plastic zones are not known a-priori thus appropriate adaptive mesh refinements are the way to construct meshes that at least approximately correspond to the shapes of the zones. Generally, the self adaptive mesh refinement technique generates the aforementioned meshes and for the considered physically nonlinear problems and various types of mesh adaptation strategies delivers the fastest convergence of the error. We tested additional h-refinements and p-enrichments along the elastic-plastic interface, as well as exclusively h-adaptive or p-adaptive mesh refinements. Only in the case of cylinder additional h-refinements resulted in a speed up of the convergence. Presuma-
bly the reason for significant improvement only in this case was the shape of elastic-plastic interface, which could be relatively easily captured in the cylinder. In the other examples even anisotropic additional $h$-refinements did not result in meshes that exactly complied with elastic-plastic zones. Therefore, neither convergence rate nor time of computation (figures 24, 25) were improved and the original $hp$-FEM delivered the fastest theoretically justified algebraic convergence. In the future the fully automatic $hp$-FEM will be supplemented with the $r$-adaptation, since it was successfully used for $p$-refinements (Nübel et al., 2007). Also the algorithm of searching for the optimal new $hp$ meshes should undergo further testing. Currently, for the sake of efficiency, only certain selected from all possible refinements are considered. Such a strategy works correctly for linear problems but its validation for elasto-plasticity is intended.

ACKNOWLEDGMENT

This research was supported by the National Science Center under grant 2011/01/B/ST6/07312.

REFERENCES


AUTOMATYCZNA HP-ADAPTACJA DLA ZAGADNIEŃ SPRĘŻYSTO-PLASTYCYZNYCH

Streszczenie


Received: November 20, 2014
Received in a revised form: November 25, 2014
Accepted: November 27, 2014
DEALING WITH PERIODIC BOUNDARY CONDITIONS FOR 1D, 2D AND 3D ISOGEOGRAPHIC FINITE ELEMENT METHOD

MARCIN ŁOŚ1, MACIEJ PASZYŃSKI1*, LISANDRO DALCIN2, VICTOR CALO2

1AGH University of Science and Technology, al. Mickiewicza 30, 30-059 Krakow, Poland
2King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
*Corresponding author: paszynski@agh.edu.pl

Abstract

In this paper we analyze the problem of implementing periodic boundary conditions in the isogeometric finite element method (ISO-FEM). The ISO-FEM method uses the B-spline-based basis functions, which facilitates usage of the same basis functions for approximation of the geometry as well as for the numerical solution of the modeled physical phenomena. The usage of the B-spline based basis functions results in C^(p-1) global continuity of the solution. The drawback is a difficulty in implementing the periodic boundary conditions, and special dedicated methods are necessary. In this paper we present two algorithms implementing the periodic boundary conditions. The first one is an iterative algorithm that utilizes widely available block-diagonal LAPACK solver. The second one is a modification of the multi-frontal solver algorithm itself, and it requires a dedicated solver with its source code modified accordingly. The presented methods can be applied in one, two or three-dimensional isogeometric finite element method.

Key words: isogeometric analysis, finite element method, periodic boundary conditions, direct solvers

1. INTRODUCTION

In this paper we discuss some algorithms dealing with periodic boundary conditions for one, two or three-dimensional isogeometric finite element method (IGA-FEM) (Cottrel et al., 2009). The IGA-FEM is a modern method of solving partial differential equations, where the global C^k continuity is kept even on finite element interfaces. This is not the case in classical higher order finite element methods (p-FEM or hp-FEM) where we have only C^0 continuity on finite elements interfaces. The higher continuity of the IGA-FEM is possible to obtain thanks to the utilization of the B-spline basis functions.

However, the B-spline basis functions of order p delivering C^(p-1) continuity have support that spreads over p+1 finite elements. This requires some special treatment in the case of periodic boundary conditions. In this paper we present two algorithms dealing with the periodic boundary conditions for IGA-FEM.

The first algorithm is based on an iterative procedure. It can be summarized in the following way. Periodic boundary conditions induce non-zero terms in heat capacity matrix in the corners outside the main diagonal region. The basic idea is to split this matrix into diagonal and non-diagonal parts, start with solution of the system without non-diagonal part as an initial approximation, and use fixed-point iteration method to obtain successive approximations of the solution. Convergence is linear, each iteration can be reduced to a simple vector equation. Using more sophisticated scheme, faster convergence can be obtained.

The advantage of the first algorithm is that it can be applied using classical, widely available solvers, e.g. LAPACK 2014 solver for block-diagonal matrices for simple 1D geometries (LAPACK, 2014), or
using Petiga framework, a part of PETSc toolkit. The disadvantage of this algorithm is its iterative nature.

The second algorithm is based on the following simple observation: basically, the boundary conditions can be treated in 1D case as a solution over the 1D circular mesh. In analogous way, we can think of the periodic boundary condition in 2D as a solution over the 2D ball, and in 3D case as a solution over the hyperball. The advantage of this approach is that it can be solved directly, without any iterative procedure. The disadvantage of this approach is that we need to design and implement special solvers for these cases, since classical available solvers, like LAPACK, do not provide such functionality.

2. THE MODEL PROBLEM

Consider the following equation, describing non-stationary heat transport in body with constant density $\rho$ and heat capacity $c$:

$$\rho c \frac{\partial u}{\partial t} = \nabla (k \cdot \nabla u) + f$$  \hspace{1cm} (1)

defined on the unit cube $\Omega = (0,1)^3$, with some initial state $u(x,y,z,0) = u_0(x,y,z)$ and periodic boundary conditions, i.e.

$$u(0,y,z,t) = u(1,y,z,t)$$  \hspace{1cm} (2)

and similarly for other directions.

Natural way to approach this problem seems to seek solution in a subspace consisting of periodic functions, that is, to choose a periodic basis, for example the tensor product of one-dimensional periodic B-splines, as it is presented in figure 1.

![Fig. 1. The periodic B-spline basis functions.](image)

The problem one immediately encounters is the structure of the heat capacity matrix. While in the case of general B-spline basis supports of basic functions are connected and localized in one place, periodic B-splines (or any periodic basis with support not being the whole interval, for that matter) situated “on the edge” have support consisting of two components, at the opposite sides of the interval. This yields additional nonzero entries in the heat capacity matrix, localized in the top right and bottom left corners, while in the usual case the heat capacity matrix is a band matrix. While sufficiently sophisticated specialized solvers can handle the resulting linear equation directly (e.g. the second method described in the paper), lack of banded structure seems to preclude straightforward use of LAPACK.

3. ITERATIVE METHOD

Crucial observation is that usually it should be possible to choose the periodic basis to be similar to an ordinary (non-periodic) one, so that there are not many additional terms in the corner, and the solution to the simplified system, with corner terms omitted, is relatively close to the exact one. This suggests an iterative scheme might be utilized.

For simplicity, let us consider the case of $C^0$-continuous B-splines (degree 1). In this case, the heat capacity matrix has the banded structure, except for two single entries at the corners. We have

$$M = A + e_1 \otimes e_n + e_n \otimes e_1$$  \hspace{1cm} (3)

where $A$ is banded. We wish to solve the system

$$Mx = b$$  \hspace{1cm} (4)

We have thus

$$(A + e_1 \otimes e_n + e_n \otimes e_1)x = b$$  \hspace{1cm} (5)

$$x = A^{-1}b - A^{-1}(e_1 \otimes e_n + e_n \otimes e_1)x$$  \hspace{1cm} (6)

Let us put $x^{(0)} = A^{-1}b$ (the solution to the simplified system) and $D = A^{-1}(e_1 \otimes e_n + e_n \otimes e_1)$, and iterate

$$x^{(t+1)} = x^{(0)} - Dx^{(t)}$$  \hspace{1cm} (7)

$D$ has very simple structure – it is easy to see it has two nonzero columns, first and last, with values $u = A^{-1}e_1$ and $v = A^{-1}e_n$, so the iterative step actually consists of merely two vector-by-scalar multiplications:

$$x^{(t+1)} = x^{(0)} - x_1^{(0)}u - x_n^{(0)}v$$  \hspace{1cm} (8)

Generalization is straightforward: heat capacity matrix is split into banded part and the extra terms in the corners, (here $e_1 \otimes e_n + e_n \otimes e_1$), solution to the original system is calculated and used as the first
guess. In general, the number of terms in the above final formula will grow linearly with the number of additional nonzero entries in the corners. In the case of B-spline basis of order $p$, it means $p(p + 1)$ terms. The scheme converges linearly, as shown in figure 2.

3. CHANGE OF BASIS FUNCTIONS

Sometimes, instead of explicitly using periodic basis, it might be more convenient to compute the heat capacity matrix in simpler basis, and then convert it to the heat capacity matrix corresponding to periodic basis (for example, to utilize existing, efficient code computing the heat capacity matrix for non-periodic basis). This section describes a general way to do this.

Let $\{N_i\}$ be the initial basis, $i = 1, ..., n$ and $\{\tilde{N}_i\}$ denote the target basis, $i = 1, ..., m$. Let $M$ be the heat capacity matrix corresponding to the initial basis, that is $M_{ij} = \langle N_i, N_j \rangle$. Assuming the space spanned by $\{\tilde{N}_i\}$ is a subspace of the one spanned by $\{N_i\}$ (as will usually be the case – in practice, $\{\tilde{N}_i\}$ will span periodic subspace of the full span of $\{N_i\}$), there exists a change-of-basis matrix $T$, so that

$$\tilde{N}_j = \sum_{i=1}^{n} T_{ij}N_i \quad \text{(9)}$$

Thus, if $\tilde{M}$ denotes the heat capacity matrix for $\{\tilde{N}_i\}$, we have by linearity

$$\tilde{M}_{ij} = \langle \tilde{N}_i, \tilde{N}_j \rangle = \sum_{r,s=1}^{n} T_{ir}T_{js} \langle N_i, N_j \rangle = \sum_{s=1}^{n} \left\{ \sum_{r=1}^{n} T_{ir}M_{ij} \right\} T_{js} = [T^TMT]_{ij} \quad \text{(10)}$$

hence $\tilde{M} = T^TMT$. Let $b$ be the RHS, that is $b_i = \langle f, N_i \rangle$. Similarly, $\tilde{b}_j = \langle f, \tilde{N}_j \rangle = \sum_{i=1}^{n} T_{ij} \langle f, N_i \rangle = [Tb]_j \quad \text{(11)}$

so $\tilde{b} = Tb$. Therefore, if $T^TMT\tilde{x} = Tb$, then $x = T\tilde{x}$ is the solution to the projection problem posed in the “right” basis - $\{\tilde{N}_i\}$, expressed in the original basis.

As the bases are most likely fixed and known in advance, $T$ matrix may be precomputed, symbolically or by some numerical method, e.g. by solving system of linear equations arising from evaluating basis functions in suitably chosen set of points. The cost of computation seems negligible compared to the rest of the FEM algorithm, so the method is applicable even if the bases are chosen dynamically.

4. DIRECT METHOD

Let us focus on one dimensional isogeometric finite element method with linear B-splines to discuss the necessary modifications to the solver algorithm when dealing with periodic boundary conditions. In particular, consider knot vector $\{0,0,1,2,3,4,5,5\}$ and control points $\{0,1,2,3,4,5\}$ which results in $N+2=7$ B-spline basis functions. Let us compare the ordinary formulation (zero Dirichlet b.c.) and formulation with periodic b.c.

Ordinary formulation yields the following system of equation:

$$B_1 = 0$$

$$B_2b(N_{3,1}, N_{2,1}) + B_3b(N_{3,1}, N_{3,1}) + B_4b(N_{3,1}, N_{4,1}) = f(N_{3,1})$$

$$B_5b(N_{4,1}, N_{3,1}) + B_6b(N_{4,1}, N_{4,1}) + B_7b(N_{4,1}, N_{5,1}) = f(N_{4,1})$$

$$B_8b(N_{5,1}, N_{4,1}) + B_9b(N_{5,1}, N_{5,1}) + B_{10}b(N_{5,1}, N_{6,1}) = f(N_{5,1})$$

$$B_{11} = 0$$

Let us compute the above ordinary formulation with periodic boundary condition formulation:

$$B_{12}b(N_{1,1}, N_{7,1}) + B_{13}b(N_{1,1}, N_{4,1}) + B_{14}b(N_{1,1}, N_{2,1}) = f(N_{1,1})$$

$$B_{15}b(N_{2,1}, N_{1,1}) + B_{16}b(N_{2,1}, N_{2,1}) + B_{17}b(N_{2,1}, N_{3,1}) = f(N_{2,1})$$

$$B_{18}b(N_{3,1}, N_{2,1}) + B_{19}b(N_{3,1}, N_{3,1}) + B_{20}b(N_{3,1}, N_{4,1}) = f(N_{3,1})$$

$$B_{21}b(N_{4,1}, N_{3,1}) + B_{22}b(N_{4,1}, N_{4,1}) + B_{23}b(N_{4,1}, N_{5,1}) = f(N_{4,1})$$

$$B_{24}b(N_{5,1}, N_{4,1}) + B_{25}b(N_{5,1}, N_{5,1}) + B_{26}b(N_{5,1}, N_{6,1}) = f(N_{5,1})$$

$$B_{27}b(N_{6,1}, N_{5,1}) + B_{28}b(N_{6,1}, N_{6,1}) + B_{29}b(N_{6,1}, N_{7,1}) = f(N_{6,1})$$

$$B_{30} = 0$$

In the periodic case, we have $N_{7,1} = N_{1,1}$.
We can draw the following conclusions from the above example:
- The ordinary (with Dirichlet b.c.) system of linear equations is block-diagonal.
- The block diagonal LAPACK solver can be used in the case of Dirichlet b.c., delivering linear O(N) computational cost.
- The periodic b.c. system is no longer block-diagonal, it has non-zero entries at the top right and bottom left part of the matrix.
- The block diagonal LAPACK solver cannot be used in this case.
- The size of the off-diagonal parts in the periodic b.c. case grows when we increase global continuity of the solution, since the support of B-spline basis function of order \( p \) spreads over \( p+1 \) elements.

Let us recall now the multi-frontal solver approach for solution of the ordinary, non-periodic boundary conditions in the case of isogeometric finite element method computations. The first one has been verified experimentally. The second one requires a straightforward modification of the classical algorithm (Obrok et al., 2010; Wozniak et al., 2014), and it will not alter the computational cost of the multi-frontal solver algorithm (Collier et al., 2012). It is therefore possible to effectively handle such constraints using isogeometric methods. Iterative method, though does not require custom solver.

\[
\begin{pmatrix}
-2 & 1 & 0 \\
1 & 0 & 0 \\
0 & 0 & 1
\end{pmatrix} \begin{pmatrix}
B_2 \\
B_3 \\
B_7
\end{pmatrix} =
\begin{pmatrix}
0.5 \\
0 \\
0
\end{pmatrix}
\]

\[
\begin{pmatrix}
0.5 \\
0 \\
0
\end{pmatrix} \begin{pmatrix}
1 & -1/2 & -1/2 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix} \begin{pmatrix}
B_2 \\
B_4 \\
B_5
\end{pmatrix} =
\begin{pmatrix}
-0.25 \\
0.75 \\
0
\end{pmatrix}
\]

We end up again with three 2x2 sub-systems, two of them can be merged at this point in the higher level to obtain

\[
\begin{pmatrix}
-1 & 1/2 & 1/2 \\
1/2 & 0 & -1/2
\end{pmatrix} \begin{pmatrix}
B_4 \\
B_5
\end{pmatrix} =
\begin{pmatrix}
1.5 \\
0.5
\end{pmatrix}
\]

Again, we can eliminate the fully assembled first row, and continue this process of partial merging and eliminating up to the root of the tree. The process is followed by recursive backward substitutions.

Let us focus now on the case of the multi-frontal solver algorithm with periodic b.c. The local sub-systems looks like

\[
\begin{pmatrix}
-1 & 1 \\
1 & -1
\end{pmatrix} \begin{pmatrix}
B_2 \\
B_3
\end{pmatrix} =
\begin{pmatrix}
0.5 \\
0.5
\end{pmatrix} \begin{pmatrix}
-1 & 1 \\
1 & -1
\end{pmatrix} \begin{pmatrix}
B_2 \\
B_3
\end{pmatrix}
\]

\[
\begin{pmatrix}
(0.5) \\
0.5
\end{pmatrix} \begin{pmatrix}
B_2 \\
B_3
\end{pmatrix} =
\begin{pmatrix}
0.5 \\
0.5
\end{pmatrix}
\]

Notice that in the last sub-system, there is again \( B_1 \) variable. The multi-frontal solver for the case of zero Dirichlet b.c. works over the binary tree structure, but the multi-frontal solver for the case of periodic b.c. works over the binary tree stretched over the 2D circular graph, compare figure 3. The 2D and 3D versions of the multi-frontal solver work in the same way, with the circle replaced by the sphere and hypersphere, respectively.

5. CONCLUSIONS

In this paper we have shown two algorithms, iterative and direct one, for dealing with periodic boundary conditions in the case of isogeometric finite element method computations. The first one has been verified experimentally. The second one requires a straightforward modification of the classical algorithm (Obrok et al., 2010; Wozniak et al., 2014), and it will not alter the computational cost of the multi-frontal solver algorithm (Collier et al., 2012). It is therefore possible to effectively handle such constraints using isogeometric methods. Iterative method, though does not require custom solver,
seems to exhibit rather slow rate of convergence. Using more sophisticated scheme, faster method should be possible to obtain.

ACKNOWLEDGEMENTS

The work presented in this paper has been supported by the National Science Center, project 2012/07/B/ST6/01229.

REFERENCES


IMPLEMENTACJA PERIODYCZNYCH WARUNKÓW BRZEGOWYCH W JEDNO, DWU I TRÓJWYMIAROWEJIZOGEOMETRYCZNEJ METODZIE ELEMENTÓW SKOŃCZONYCH

Streszczenie


Received: September 16, 2014
Received in a revised form: November 26, 2014
Accepted: December 10, 2014
EVALUATION OF MACROSCOPIC STRESSES IN DISCRETE ELEMENT MODELS OF SINTERING PROCESSES

JERZY ROJEK1,*, SYZYMON NOSEWICZ1, KATARZYNA PIETRZAK1,2, MARCIN CHMIELEWSKI2

1 Institute of Fundamental Technological Research (IPPT PAN), Pawińskiego 5B, 02-106 Warszawa, Poland
2 Institute of Electronic Materials Technology, Wólczyńska 133, 01-919 Warszawa, Poland
*Corresponding author: Jerzy.Rojek@ippt.pan.pl

Abstract

This paper presents investigation of macroscopic stresses in powder metallurgy process modelled with the discrete element method. The discrete element model belongs to the class of micromechanical models. In the DEM model the material is represented by an assembly of particles interacting by contact forces and the method is formulated in terms of forces and displacements. In order to evaluate macroscopic stresses a special upscaling procedure is necessary.

The paper presents basic formulation of the discrete element method with special attention for the contact interaction models for powder compaction and sintering. A method to evaluate macroscopic stresses based on the two level averaging is presented. The discrete element model of sintering is verified using own experimental results. Macroscopic stresses are calculated for the whole process including loading, heating, sintering, cooling and unloading. It has been found out that the macroscopic stresses are consistent with changing process parameters. The procedure is suitable for multiscale modeling of sintering.

Key words: sintering, modeling, discrete element method, macroscopic stresses

1. INTRODUCTION

Sintering is an essential stage of powder metallurgy processes in which solid parts are manufactured from metal or ceramic powder mixtures. Sintering consists in consolidation of loose or weakly bonded powders at elevated temperatures, close to the melting temperature with or without additional pressure. Sintering is a complex process affected by many factors and sensitive to many defects. Some of the most frequent defects such as cracks and shape distortions are associated with stresses during and after the sintering process.

Numerical modelling can be employed to simulate a sintering process and analyse sintering stresses and shape distortion of sintered parts. There are different approaches in modelling of sintering processes, ranging from continuum phenomenological models to micromechanical and atomistic ones.

This work presents application of the discrete element method (DEM) to modelling of sintering. The discrete element model belongs to the class of micromechanical models. In the DEM model the material is represented by an assembly of particles interacting by contact forces. The discrete element model of sintering requires a special interaction model. This work employs an original viscoelastic model developed by Nosewicz et al. (2013).

Numerical simulation of sintering at the microscopic level allows us to study material phenomena occurring during sintering, such as interaction between the grains during sintering and their rearrangement, material shrinkage and gradual decrease of porosity. Use of an averaging procedure will al-
allow us to transfer force-type interactions to macroscopic stresses.

2. SINTERING MECHANISMS

During sintering, a particulate material (figure 1) is converted into a solid compact body (Olevsky, 1998). In the initial stage, cohesive bonds (necks) are formed between grains. Microstructure at an early stage of sintering is shown in figure 2a. When the sintering process is continued the necks between particle grow. Grain rearrangement and increase of grain compaction can be observed during sintering (figure 2b).

3. NUMERICAL MODEL OF SINTERING

Numerical model of sintering used in the present work has been developed within the framework of the discrete element method which assumes that a particulate material can be represented as a collection of spherical particles interacting among one another, thus the discrete element model takes explicitly into account the particulate nature of the sintered material (Nosewicz et al., 2013). The numerical model of sintering has been implemented in the finite/discrete element code DEMPack (Dempack, 2013).

3.1. Discrete element formulation

In the discrete element method, the motion of rigid spherical elements (particles) is governed by the standard equations of rigid body dynamics. In general, both the translational and rotational motion is considered in Rojek et al. (2005). Here, however, we have neglected the tangential interaction between particles, similarly as in Martin et al. (2006), therefore the rotational motion of the particles is not considered. This favors particle compaction. Thus, the translational motion of the $i$-th particle is described by the equation:

$$m_i \ddot{u}_i = F_i,$$  

(1)

where $u_i$ is the element centroid displacement in a fixed (inertial) coordinate frame $X$, $m_i$ – element (particle) mass, and $F_i$ – resultant force. The force vector $F_i$ includes all the forces applied to the $i$th element due to external load, $F^\text{ext}_i$ and contact interactions with neighbouring spheres and boundary surfaces $F^\text{cont}_j$, $j = 1, \ldots, n_i^c$, where $n_i^c$ are the number of elements being in contact with the $i$-th discrete element

$$F_i = F^\text{ext}_i + \sum_{j=1}^{n_i^c} F^\text{cont}_j.$$  

(2)

Contact forces $F^\text{cont}_j$ are obtained using a suitable constitutive model formulated for the interaction
of two particles. The present model employs two different interaction models for different stages of the process, one model for the compaction stage and the other for sintering.

3.2. Contact model for powder compaction

Powder compaction prior to sintering is modeled assuming cohesionless and frictionless contact conditions. The rheological scheme of this model is shown in figure 3.

Fig. 3. Rheological scheme of the contact interaction for powder compaction.

The contact interaction is represented by the Kelvin-Voigt element consisting of a spring and a dashpot connected in parallel. The total contact force \( F_n \) is a sum of the elastic force in the spring \( F_n^e \) and the viscous component \( F_n^d \)

\[
F_n = F_n^e + F_n^d.
\]  (3)

The elastic part of the normal contact force \( F_n^e \) can be evaluated assuming a linear force–displacement relationship

\[
F_n^e = k_n u_n,
\]  (4)

where \( k_n \) is the contact stiffness and \( u_n \) is the penetration of the two particles, calculated as

\[
u_n = d_{ij} - r_i - r_j,
\]  (5)

where \( d_{ij} \) is the distance of the particle centres, and \( r_i, r_j \) their radii.

No cohesion is allowed, so no tensile normal contact forces are allowed

\[
F_n^e \leq 0.
\]  (6)

The viscous component of the normal force is assumed to be a linear function of the normal relative velocity \( v_n \)

\[
F_n^d = c_n v_n
\]  (7)

where

\[
v_n = (\dot{u}_j - \dot{u}_i) \cdot \mathbf{n}.
\]  (8)

The value of the viscosity coefficient \( c_n \) can be taken as a fraction \( \xi \) of the critical damping \( C_{cr} \) for the system of two rigid bodies with masses \( m_i \) and \( m_j \), connected with a spring of the stiffness \( k_n \)

\[
c_n = \xi C_{cr}
\]  (9)

where the critical damping can be calculated as, cf. Taylor and Preece (1992):

\[
C_{cr} = 2 \frac{m_i m_j k_n}{m_i + m_j}.
\]  (10)

3.3. Contact model for sintering

Contact interaction during sintering is represented by an original viscoelastic model developed by Nosewicz et al. (2013). The rheological scheme of this model is shown in figure 4. The model consists of two parallel elements, one representing the sintering driving force \( F_{sint} \) and the other being the Maxwell element comprising the elastic component in series with the viscous one.

Fig. 4. Rheological scheme of the contact interaction during sintering.

For the Maxwell element, we have the following relationships for forces and velocities:

\[
F^e = F^v
\]  (11)

\[
v_n = v^e + v^v
\]  (12)

Equation (11) means that the forces transferred through the spring and viscous component, \( F^e \) and \( F^v \), respectively, are equal. Equation (12) expresses the additive decomposition of the relative normal velocity between particles \( v_n \) into the elastic and viscous parts, \( v^e \) and \( v^v \), respectively.

The elastic force is expressed by the linear relationship

\[
F^e = k_n u^e
\]  (13)

where \( k_n \) is the contact stiffness and \( u^e \) is the elastic part of the relative normal displacement at the contact point. The viscous force is written in the form:

\[
F^v = \eta v^v
\]  (14)
where $\eta$ is the viscosity coefficient. Substituting equations (14) and (13) into (12) we obtain the evolution equation for the force in the Maxwell branch

$$v_i = \frac{F_i^e}{k_n} + F_i^e \frac{1}{\eta}$$  \hspace{1cm} (15)

The sintering driving force and viscosity are evaluated according the the classical models developed for two-particle sintering (Coble, 1961; Johnson, 1969; De Jonghe & Rahaman, 1988) and used in previous implementations in the discrete element method, cf. Parhami and McMeeking (1998), Martin et al. (2006), Olmos et al. (2009). The sintering driving force $F_{\text{sint}}$ and viscosity coefficient $\eta$ are given by the following formulae:

$$F_{\text{sint}} = \pi \gamma_s \left[ 4r \left( 1 - \cos \frac{\Psi}{2} \right) + a \sin \frac{\Psi}{2} \right]$$  \hspace{1cm} (16)

$$\eta = \frac{\pi a^4}{8D_b}$$  \hspace{1cm} (17)

where $r$ – the particle radius, $a$ – the radius of the interparticle grain boundary, $\Psi$ – the dihedral angle, $\gamma_s$ – the surface energy and $D_b$ – the effective grain boundary diffusion coefficient. The geometrical parameters of the model are defined in figure 5.

4. EVALUATION OF MACROSCOPIC STRESSES

The discrete element model presented in this work can be used in the framework of multiscale modelling of sintering as the model at microscopic level (Pan, 2003). Multiscale analysis requires transfer of the parameters between scales. Transition between scales can be performed two ways, from the lower scale to the upper one (this is called upscaling) or from the upper scale to the lower one (downscaling).

The transfer can be applied to different types of parameters. In this work special attention is paid to the micro-macro transition (upscaling) involving macroscopic stresses and microscopic forces.

Theoretical bases of upscaling have been developed within the theory of homogenization which was also applied to granular materials modelled with the discrete element method (Kruyt & Rothenburg, 2004; Kruyt & Rothenburg, 1998). In this work, The macroscopic stress tensor will be determined using the concept of two-level averaging procedure presented in Chang et al. (1995), Luding (2004).

Fig. 5. Two-particle model of sintering.

The effective diffusion coefficient is given by the following equation, cf. Parhami and McMeeking (1998):

$$D_b = \frac{D_g \delta_g \Omega}{kT}$$  \hspace{1cm} (18)

where $D_g$ – diffusion coefficient, $\delta_g$ – thickness of the grain boundary, $\Omega$ – atomic volume, $k$ – Boltzmann constant, $T$ – sintering temperature.

3. EVALUATION OF MACROSCOPIC STRESSES

The discrete element model presented in this work can be used in the framework of multiscale modelling of sintering as the model at microscopic level (Pan, 2003). Multiscale analysis requires transfer of the parameters between scales. Transition between scales can be performed two ways, from the lower scale to the upper one (this is called upscaling) or from the upper scale to the lower one (downscaling).

The transfer can be applied to different types of parameters. In this work special attention is paid to the micro-macro transition (upscaling) involving macroscopic stresses and microscopic forces.

Theoretical bases of upscaling have been developed within the theory of homogenization which was also applied to granular materials modelled with the discrete element method (Kruyt & Rothenburg, 2004; Kruyt & Rothenburg, 1998). In this work, The macroscopic stress tensor will be determined using the concept of two-level averaging procedure presented in Chang et al. (1995), Luding (2004).

Fig. 6. Definition of inter-particle interaction.

We assume that the discrete elements (figure 6) interact among themselves with contact forces determined according to the model described earlier, being the material model at the microscopic scale. In the first stage we perform averaging over the representative volume elements coinciding with discrete elements volumes $V_p$. Thus we obtain the quantity $Q$ represented by the constant value $Q_p$ over the volume of the $p$-th discrete element. In the second stage we perform averaging over representative volumes containing a certain number of discrete elements.
The stress tensor $\sigma_p$ for a single discrete element will be calculated by averaging over the element volume $V_p$:

$$\sigma_p = \frac{1}{V_p} \int \sigma \, d\Omega$$  \hspace{1cm} (19)

Using the equilibrium condition and divergence theorem the volume integral \((19)\) can be transformed into the surface one (Luding, 2004):

$$\sigma_p = \frac{1}{V_p} \int \mathbf{q} \, dS$$  \hspace{1cm} (20)

where $\mathbf{q} = \sigma \, \mathbf{n}$ is the stress vector, with $\mathbf{n}$ being the unit vector normal to the element surface $S_p$ (cf. figure 6). Taking into account that the surface loading in our case is produced by concentrated forces $\mathbf{F}_c$ the surface integral in equation \((20)\) can be written as the following sum (Chang et al., 1995; Luding, 2004):

$$\sigma_p = \frac{1}{V_p} \sum_{c=1}^{n_c} \mathbf{x}_c \mathbf{F}^c$$  \hspace{1cm} (21)

where $n_c$ is the number of elements being in contact with the $p$-th element. Writing the position vector of the contact point $\mathbf{x}_c$ as the sum of the position vector of the element center, $\mathbf{x}_p$, and the vector, $\mathbf{s}^c$, connecting the element center with the contact point

$$\mathbf{x}_c = \mathbf{x}_p + \mathbf{s}^c$$  \hspace{1cm} (22)

for the static equilibrium we obtain the stress tensor for a single discrete element in the following form:

$$\sigma_p = \frac{1}{V_p} \sum_{c=1}^{n_c} \mathbf{s}^c \mathbf{F}^c$$  \hspace{1cm} (23)

After calculation of stresses for single elements $\sigma_p$, we can perform averaging over representative volume elements, defined for any point $\mathbf{x} \in \Omega$. Given constant stresses over elements, the average stress in the representative volume element can be calculated as, cf. Chang et al. (1995), Luding (2004):

$$\mathbf{\bar{\sigma}} = \left\langle \sigma \right\rangle = \frac{1}{V} \sum_{p \in V} \mathbf{V}_p \sigma_p = \frac{1}{V} \sum_{p \in V} \sum_{c=1}^{n_c} \mathbf{s}^c \mathbf{F}^c.$$  \hspace{1cm} (24)

The expression \((24)\) for the mean stress of a representative volume can be written in an alternative equivalent form, cf. Chang et al. (1995):

$$\mathbf{\bar{\sigma}} = \frac{1}{V} \sum_{c=1}^{n_c} \mathbf{L}^c \mathbf{F}^c,$$  \hspace{1cm} (25)

in which summation is over all $N_c$ contacts in the representative volume element and $\mathbf{L}^c$ is the so called branch vector connecting the centroids of two particles, for two particles $i$ and $j$ defined as follows

$$\mathbf{L}^c_{(ij)} = \mathbf{x}_p^{(i)} - \mathbf{x}_p^{(j)}$$  \hspace{1cm} (26)

where $\mathbf{x}_p^{(i)}$ and $\mathbf{x}_p^{(j)}$ are the position vectors of the two particle centroids.

5. EXPERIMENTAL STUDIES OF SINTERING

Experimental studies of sintering have been carried out for NiAl powder. Morphology of the NiAl powder used for sintering is presented in figure 2. Sintering has been performed under pressure of 30 MPa and at temperature of 1400°C. Temperature and pressure variation during the process are plotted in figure 7. The samples of the sintered material are shown in figure 8. The process was interrupted at different time instants in order to study the evolution of microstructure and density during sintering. The density evolution, strictly related to sintering kinetics, was used in the validation of the developed numerical models of sintering.

Fig. 7. Pressure and temperature profiles.

Fig. 8. Sintered specimens.
6. NUMERICAL RESULTS

Maintaining the original grain size and grain size distribution, a cylindrical container of diameter 200 \( \mu m \) has been filled with 1751 particles (figure 9a). It has been assumed that such a reduced geometric model represents correctly sintering process in a real specimen with diameter of 120 mm. This assumption is justified provided the parameters characterizing sintering are uniformly distributed in a real specimen volume. The model parameters used in the analysis are given in table 1. The whole process consisting of loading, heating, sintering, cooling and unloading has been simulated using the pressure and temperature profiles shown in figure 7. Figure 9b shows the final geometry of the sintered specimen. A significant height reduction can be observed. Density evolution during sintering is plotted in figure 10. It can be observed that numerical results are in quite a good agreement with experimental data.

Table 1. Material data for NiAl sintering.

<table>
<thead>
<tr>
<th>Material constant</th>
<th>Parameter value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diffusion coefficient, ( D )</td>
<td>1.85 ( \times ) 10^{-20} m^3/s</td>
</tr>
<tr>
<td>Atomic volume, ( \Omega )</td>
<td>1.20 ( \times ) 10^{-29} m^3</td>
</tr>
<tr>
<td>Surface energy, ( \gamma_s )</td>
<td>1.57 J/m^2</td>
</tr>
<tr>
<td>Dihedral angle, ( \Psi )</td>
<td>129°</td>
</tr>
<tr>
<td>Density, ( \rho )</td>
<td>5910 kg/m^3</td>
</tr>
</tbody>
</table>

Macroscopic average stresses have been calculated using the procedure presented in chapter 4. The whole specimen has been treated as the RVE. Evolution of the macroscopic stresses during the whole process is plotted in figures 11 and 12. It can be seen in figure 11 that the axial stresses are in equilibrium with externally applied pressure during nearly all the process except for the initial stage of sintering. This is reasonable since at this stage the system undergoes the most intensive compaction. Evolution of the three principal stresses is shown in figure 12. We can see that under external loading before sintering all the three stress components are nearly equal which is expected since we should have the state of hydrostatic compression. With the progress of sintering the radial stresses gradually decrease to zero due to radial shrinkage of the specimen and finally the state of uniaxial compression is obtained.
7. CONCLUDING REMARKS

The results presented in this paper show that the micromechanical model of sintering developed within the framework of the discrete element method reproduces correctly macroscopic behaviour of the material during sintering. The macroscopic stresses and other macroscopic quantities can be obtained by upscaling of the microscopic parameters. This will allow to use the discrete element model for multiscale modelling of sintering.

ACKNOWLEDGEMENTS

The results presented in this paper have been obtained within the projects funded by the National Science Centre awarded by decision numbers DEC-2013/11/B/ST8/03287, DEC-2012/05/N/ST8/03376 and DEC-2014/12/T/ST8/00681, as well as Operational Programme Human Capital 8.2.1.

REFERENCES


WYZNACZANIE NAPRĘżeŃ MAKROSKOPOWYCH W DISKRETNYM MODELU SPIEKANIA

Streszczenie

Artykuł przedstawia analizę naprężeń makroskopowych w procesie metalurgii proszków modelowanym metodą elementów dyskretnych. Metoda elementów dyskretnych należy do metod modelowania mikromechanicznego. W tej metodzie materiał jest reprezentowany przez liczny zbiór cząstek oddzielających się między sobą poprzez siły kontaktu. Sformułowanie metody wykorzystuje związki pomiędzy siłami i przemieszczeniami.


Received: November 27, 2014
Received in a revised form: December 17, 2014
Accepted: December 19, 2014
METHOD TO IDENTIFY RHEOLOGICAL CONSTITUTIVE MODEL ADAPTED FOR POWDER INJECTION MOULDING PROCESS USING INVERSE METHOD

DIMITRI CLAUDEL*, JEAN-CLAUDE GELIN, MOHAMED SAHLI, THIERRY BARRIÈRE

FEMTO-ST Institute / Applied Mechanics Department, 24 rue de l’Epitaphe, 25000 Besançon, France
*Corresponding author: dimitri.claudel@edu.univ-fcomte.fr

Abstract

The aim of this paper is to show a method to identify rheological parameters of constitutive models for powder injection moulding process. The constitutive rheological model used is generalized model mixing several rheological laws such as Carreau-Yasuda, Maron-Pierce and Arrhenius. Thus, the constitutive model takes accounts shear rate, powder volume loading, temperature and particle size.

The material used for this study is Inconel 718, nickel-chromium-based superalloy is typically used in high-temperature and high-performance applications, particularly in the aeronautic industry. To elaborate the feedstock, powder was mix with a formulation composed of three different binder ingredients: polypropylene (PP), polyethylene glycol (PEG) and a stearic acid (SA).

Then, a rheological characterization on the powder and feedstock was carried out. The rheological properties of the resulting binder formulations and feedstocks were characterized using a capillary rheometer. First all the binder granules were filled into rheometer barrel heated to 170, 180 or 190°C. More than, the powder particle size distribution was measured by laser scattering particle analyser.

Then, the data collected from the first characterization were used to identify parameters of the model. Then, this identification of parameters could be used to carry out numerical injection simulations. Also, sensitivity of parameters analysis was carrying out to determine influence of each of rheological law.

Key words: rheological model, metal injection moulding, identification parameter

1. INTRODUCTION

The Metal Injection Moulding (MIM) process is an economically attractive method of producing large amounts of small and complex metallic parts. This process is expanding regarding the numerous scientist papers published these 20 last years (Shivashankar et al., 2013; Özgün et al., 2013; Quinard et al., 2011). The dimensions and mechanical properties of MIM components are influenced by the feedstock characteristics, the process parameters of the injection moulding, as well as the atmosphere and kinetics of debinding and the sintering.

Numerical simulations are a very important feature of the beginning of any product or technology development. It requires also accurate constitutive models describing material behaviour at large shear rates up to $10^5$ s$^{-1}$. The choice of a rheological model and the determination of its parameters should be made from tests generating such conditions.

To identify rheological constitutive models of loaded polymer feedstocks, one generalized constitutive law is applied. The binder was composed of three polymers: polypropylene (PP), Polyethylene glycol (PEG) and stearic acid (SA) (Uterkin et al., 2011). It was prepared by twin screw mixing and mixed with powders. After mixing, the binder has
been granule and use for rheometer analyses. The super alloy powders used in the elaborated feedstock are Inconel 718 for airplane or automotive applications. The formulation is based on thermal and solvent debinding (Omar et al., 2003; Onbattuvelli et al., 2014).

The rheological properties of the resulting binder formulations and feedstocks were characterized by a capillary rheometer. In the rheometer, shear rates were varied from 100 to 105 s⁻¹.

Then, comparisons are carried out among the experimental and analytical results. Thus, several models were used such as power law, Williams-Landel-Ferry, Arrhenius, Maron-Pierce and Carreau-Yasuda. In the wake of this comparison, identifications of parameters used in models were performed. This parameter identification carried out with nonlinear least squares method and a trust-region algorithm.

Finally, identified parameters were used to complete identification of a more complex law. This equation takes account of several parameters such as particle size, temperature, shear rate and powder volume loading and take back identifications of parameters from previous laws.

Then, was performed sensitivity analysis on parameters of the rheological model, in order to determine the main parameters that need to be precisely estimated.

2. MATERIAL CONSTITUTIVE BEHAVIOUR LAWS

In the present work, comparisons are made between experimental results and those obtained from analytical models. Thus, several models more or less complicated were used such as power law, Arrhenius, Chong and Carreau-Yasuda. In the wake of this comparison, identification of parameters used in models and sensitivity analysis were performed, see table 1. In this paper, Ratkovich et al. (2013) show several rheological models suitable for non-Newtonian material among the numerous models predicting the apparent viscosity with different degrees of complexity. Table 2 shows several models suitable for viscosity of load polymers. The symbols in the tables mean: \( \tau_0 \) the yield stress (Pa), \( n \) the flow behaviour index (−), \( k \) the flow consistency index (Pa·s⁻¹), \( \dot{\gamma} \) the shear rate (/s), \( \mu \) the apparent viscosity (Pa·s), \( \mu_\infty \) the infinite rate apparent viscosity (Pa·s), \( \mu_0 \) the zero shear apparent viscosity (Pa·s), \( \lambda \) the (Cross) time constant (s) and \( m \) the Cross rate constant (−) in table 1, and \( \eta \) the viscosity (Pa·s), \( \eta_0 \) the apparent viscosity (Pa·s), \( \phi \) the powder load rate and \( \phi_m \) the maximal powder load rate in table 2, respectively.

The Carreau-Yasuda model is a generalized model of Carreau for which \( n = 2 \) (Yasuda et al., 1981). As shown in figure 1, different power law models may be appropriate for different shear rate regimes. The Carreau-Yasuda model allows getting a best link between the power-law model and the Newtonian plateau at low shear rates. It is why this model has been chosen for the identification. However, the identification of the power-law model has been done before in the aim to determine the flow behavior index \( n \).

<table>
<thead>
<tr>
<th>Model</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power-law (Ostwald de Waele)</td>
<td>( \tau = k\dot{\gamma}^n )</td>
</tr>
<tr>
<td>Bingham</td>
<td>( \tau = \tau_0 + k\dot{\gamma} )</td>
</tr>
<tr>
<td>Herschel and Bulkley</td>
<td>( \tau = \tau_0 + k\dot{\gamma}^n )</td>
</tr>
<tr>
<td>Casson</td>
<td>( \tau = \tau_0 + k\dot{\gamma}^n + \mu_\infty \dot{\gamma}^n )</td>
</tr>
<tr>
<td>Sisko</td>
<td>( \mu = \mu_\infty + K\dot{\gamma}^{n-1} )</td>
</tr>
<tr>
<td>Cross</td>
<td>( \frac{\mu - \mu_\infty}{\mu_0 - \mu_\infty} = \frac{1}{1 + (\lambda \dot{\gamma})^m} )</td>
</tr>
<tr>
<td>Carreau</td>
<td>( \frac{\mu - \mu_\infty}{\mu_0 - \mu_\infty} = (1 + (\lambda \dot{\gamma})^m)^{\frac{1}{m}} )</td>
</tr>
</tbody>
</table>

In the same way of Senapati et al. (2010) and Hidalgo Garcia (2013), a generalized model (see equation (1)) is elaborated from Arrhenius, Carreau-Yasuda and Maron-Pierce laws:

\[
\eta/\eta_0 = \left( 1 + \frac{1.25\phi}{1 - \phi/\phi_m} \right)^2
\]

\[
\eta/\eta_0 = \exp\left( \frac{2.5\phi}{1 - \phi/\phi_m} \right)
\]

\[
\eta/\eta_0 = \left( 1 + \frac{\phi}{\phi_m} \right)^{-2}\phi_m
\]

\[
\eta/\eta_0 = \left( 1 + 0.75 \frac{\phi}{\phi_m} \right)^2
\]

\[
\eta/\eta_0 = \left( 1 - \frac{\phi}{\phi_m} \right)^{-2}
\]

\[
\eta/\eta_0 = \frac{1 - \phi/\phi_m}{(1 - \phi/\phi_m)^2}
\]

\[
\eta/\eta_0 = \left( 1 + 0.5 k \phi - \phi \right) / (1 - k \phi)^2 (1 - \phi)
\]
\[ \eta = \frac{C_u}{D_{50}} \exp \left( \frac{E_a}{R \cdot T} \right) \eta_0 (1 + (\lambda \gamma)^2)^{\frac{n-1}{m}} \left( \frac{\varphi_m}{\varphi_m - \varphi} \right)^m \] 

(1)

where \( E_a \) is the Arrhenius activation energy (kJ·mol\(^{-1}\)), \( R \) is perfect gas constant (J·K\(^{-1}\)·mol\(^{-1}\)), \( C_u \) the D\(_{60}/D_{10} \), \( D_{60} \) the particle size at 60%, \( D_{10} \) the particle size at 10% and \( D_{50} \) the particle size at 50%.

3. RESULTS AND DISCUSSIONS

3.1. Powder particle size

Figure 2 illustrates the particle size distribution for Inconel super alloy powders. The particle size distribution was shown in terms of the number of particles. The graph shows the \( d_{10} \), \( d_{50} \) and \( d_{90} \) diameters equivalent to 3.35 \( \mu \)m, 6.24 \( \mu \)m and 10.97 \( \mu \)m, respectively. The standard deviation in this case, was 0.86 \( \mu \)m. The measure of SSA indicates a specific area of 0.095 m\(^2\)/g.

3.2. Shear viscosity

The results of the shear rate viscosity measurements vs shear rate and temperature of the Inconel superalloy elaborated feedstocks are shown in figure 3. It exhibits pseudo-plastic flow behaviour and the viscosity decreased as the shear rate increased at all test temperatures. The values of shear viscosity are less 100 Pa·s in injection range, the developed and elaborated feedstock is very low and very easily injectable for injection and micro-injection process. This viscosity is plotted for 170, 180 and 190°C. Shear viscosity decreases with shear rate for all temperatures. This behaviour matches with pseudoplastic behaviour flow. So, the Carreau-Yasuda model is suitable to model this viscosity behaviour curves.

3.4. Identification of the rheological parameters

The first identification by inverse method (Szela et al., 2006) was carried out with the power-law model and is presented in figure 4. This identification allows to define the flow behaviour index equivalent to a value of 0.36. For this model, only the experimental measurements with shear rates bigger than \( 10^3 \) s\(^{-1}\) were taken into account for identification.
The second identification was performed with the Carreau-Yasuda model and is given on the figure 5. With this model, the experimental measurements and analytical predictions are almost similar. The parameters allowing this good correlation between analytical and experimental curves are given in table 3.

### 3.5. Sensitivity analysis

To determine the influence of each parameter of the Carreau-Yasuda model, a sensitivity analysis was carried out (figure 6 and figure 7). For each curve of this analysis, only one parameter is increased.

![Fig. 5. Comparison of generalized law modeling after parameter identification of the evolution of shear viscosity vs shear rate and temperature for Inconel loaded polymer.](image)

**Table 3.** Identified parameters using a generalized full model (see equation (1)).

<table>
<thead>
<tr>
<th>T (°C)</th>
<th>Cu/d50</th>
<th>E_a (kJ·mol⁻¹)</th>
<th>η₀ (Pa·s)</th>
<th>λ</th>
<th>φ</th>
<th>φ_m</th>
<th>m</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>170</td>
<td>0.498</td>
<td>20.22</td>
<td>683.7</td>
<td>0.005031</td>
<td>0.7</td>
<td>0.72</td>
<td>0.009</td>
<td>1</td>
</tr>
<tr>
<td>180</td>
<td>0.498</td>
<td>20.26</td>
<td>521.4</td>
<td>0.005597</td>
<td>0.7</td>
<td>0.72</td>
<td>0.009</td>
<td>1</td>
</tr>
<tr>
<td>190</td>
<td>0.498</td>
<td>20.06</td>
<td>376.8</td>
<td>0.01005</td>
<td>0.7</td>
<td>0.72</td>
<td>0.009</td>
<td>1</td>
</tr>
</tbody>
</table>

![Fig. 6. Generalized law coefficients variation vs shear rate.](image)

![Fig. 7. Zoom in generalized law coefficients variation curve: on poor shear rate area (a), on high shear rate area (b).](image)
of 10% with regards to the initial value. This application allows to see the influence of each parameter with regards to the baseline.

The $a$ and $n$ parameter influences on the viscosity curve are depending of the shear rate. These parameters are the most important when shear rate is superior to 1000/s. Inversely, they have no influence with low shear rates and $\eta_0$ becomes the most influential in this situation.

4. CONCLUSIONS

The determination of rheological parameters using data obtained through standard viscometric flows opens the possibility of building reliable analytical models, which can be used in the injection simulations. The present work aims at the identification of rheological parameters associated to a non-linear constitutive equation. Firstly, a robust analysis has been used, in order to determine the constitutive coefficient values, starting from experimental data obtained by rheological tests. Then, an extensive sensitivity analysis has been performed. The aim of this work is to examine a proper constitutive model to more accurately describe the influence of the shear rates and temperature on the feedstock behavior in large shear rates ranges.

To improve the rheological sensitivity in the injection simulations, in a future work we propose to analyze the influence of a more complex law. It is then necessary to take into account the solid loading, shapes and particle size of powder grains and then interaction between powders and binders.

REFERENCES


METODO IDENTYFIKACJI KONSTYTUTYWNEGO MODELU REOLOGICZNEGO ZASTOSOWANEGO W PROCESIE ODWROTKOWA METODY WTRYSKIWANIA PROSZKU PRZY UŻYCIU METODY ODWROTNEJ

Streszczenie

Celem pracy jest przedstawienie metody identyfikacji parametrów reologicznych modelu konstytutywnego dla procesu wtryskiwania proszku. Reologiczny model konstytutywny jest uogólniony model będący połączeniem wielu praw reologicznych, takich jak Carreau-Yasuda, Maron-Pierce’a i Arrheniusa. Zatem model konstytutywny bierze pod uwagę prędkość ścianania, objętość proszku, temperaturę i wielkość cząstek.

Analizowanymi materiałami były Inconel 718 i nadstop na bazie niklu i chromu, stosowane w procesach wysokotemperaturowych i o wysokiej wydajności, szczególnie w przemyśle lotniczym. Material wadlowy wytworzono z proszku składającego się z trzech składników: polipropylenu (PP), glikolu polietylenowego (PEG) i kwasu stearynowego (SA).

Następnie opracowano charakterystykę reologiczną dla proszku i materiału wadlowego. Reologiczne właściwości spoiwa i materiału wadlowego zostały wyznaczone reometrą kapilarnego. Na początku spojowo umieszczono w reometrze, który podgrzany do temperatury 170, 180 lub 190°C, następnie został zmierzony rozkład wielkości cząstek proszku z użyciem laserowego analiza- tora rozmiarów cząstek.

W kolejnym kroku badała zebrańe dane zostały wykorzystane do identyfikacji parametru modelu. Oszacowane parametry modelu mogą zostać zastosowane w symulacjach numerycznych proce-
su wtryskiwania. Ponadto w pracy przeprowadzono analizę wrażliwości, określając wpływ parametrów poszczególnych modeli na wartości wyjściowe z modeli.

Received: September 30, 2014
Received in a revised form: December 18, 2014
Accepted: December 23, 2014
CHARACTERIZATION BY INFRARED SPECTROSCOPY OF BINDER BASED ON POLYETHYLENE GLYCOL AND INCONEL 718 FEEDSTOCK FOR POWDER INJECTION MOLDING

ALEXANDRE ROYER, JEAN-CLAUDE GELIN, THIERRY BARRIERE*

Femto-ST Institute, Applied Mechanics Department, UMR 6174 CNRS, ENSMM, 25030 Besançon cedex, France
*Corresponding author: thierry.barriere@univ-fcomte.fr

Abstract

Metal injection moulding (MIM) has over the past decade established itself as a competitive manufacturing process to produce in large quantities small precision components with complex shape which would be costly to produce by alternative methods. MIM is a process which combines the versatility of plastic injection moulding with the strength and integrity of machined, pressed or otherwise manufactured small, complex metal parts. MIM consists in shaping powder particles and sintering them. During the injection phase, segregation appears in the feedstock and defects will appear in the component during the sintering. To limit this effect, during decades a vast variety of binder systems have been developed. Binder systems are formulated as a mixture of different organic or inorganic substances with several functions. Binder system has the main commitments of giving the necessary rheological behavior to the feedstocks for injection moulding to transport the powder particles into the mould cavity and the cohesion of the green part.

In this paper a study of chemical interactions between polymers in binder and in Inconel feedstocks was investigated by Fourier Transform InfraRed spectroscopy (FTIR) and by differential scanning calorimetry (DSC). These methodologies were also investigated to study the thermal behavior of the binder at a temperature close to the temperature of injection. Analyzes shows relationship between chemical interactions and miscibility of polymer and different rheological and mechanical behavior. All methodologies revealed no interactions between the different component of the binder and the powder. This result shows the necessity to develop a better formulation of binder to improve the homogeneity of the feedstock and reduce the segregation during injection phase.

Key words: metal injection moulding, feedstock, binder, FTIR analysis

1. INTRODUCTION

Metal injection moulding (MIM) is a process for competitive manufacturing to produce in mass (more than 10,000 units per year) small parts (less than 50g) with complex geometry from varied materials such as stainless steel, superalloys, carbides or ceramics. The MIM process provides good dimensional accuracy (tolerances <5 microns), a good surface finish (Ra <1.5 microns) and high mechanical strength similar to those of the powder material. The PIM is a process that has great potential and involves many industrial sectors such as automotive, aerospace, information technology or medical. This method is based on the injection of a fluid material composed of powder of the desired material for the final part, and of a binder generally consisting of several polymers. The piece is then subjected to debinding to remove the binder and then the piece is sintered to obtain a dense part (Enneti, 2012).

The binder is the most important part in the MIM process as it contributes fully to multitask like to be able to support a load rate by importing powder, typically 60%, and to direct the powder in the mold (Enneti, 2012). For this it must give strength and cohesion to the molded part and be easily re-
moved from the molded part and be recyclable, environmentally friendly and economical. It thus requires a low viscosity, good adhesion to the powder, no chemical interactions with the powder (Enneti, 2012; Tam et al., 1997; Scott Weil et al., 2006) and a low coefficient of thermal expansion. Binders are usually classified into three categories:

- Based on thermoplastic: the most currently used binders formulations are generally based on thermoplastic and compound of polyethylene, propylene and natural or synthetic wax and/or stearic acid (Scott Weil et al., 2006; Enneti, 2012).
- Based on thermoset: binders based on thermoset are used to achieve greater resistance of the work piece after injection but are not recyclable. They are generally composed of polycarboxilane or epoxy resin.
- Based on gel: the gel binders reduce the use of polymers and may be formed of an aqueous gel or silica gel (Hidalgo et al., 2013). Some polysaccharides are used as a component due to their gelling properties in water.

In this study, a powder of Inconel 718 was investigated to an application in the MIM process. Inconel superalloy is used in aviation, aerospace and nuclear power for its high resistance to corrosion and oxidation but also for its excellent mechanical strength at high temperature (Özgün et al., 2013a, 2013b). Superalloy use most often is currently Inconel 718 and 625. Here we are using an inconel 718 whose composition is given in table 1. The binders used in these superalloys are generally composed of PP, CW, PW and SA (Özgün et al., 2013a, 2013b). However, here we will work with a conventional formulation of MIM (Urterkin et al., 2011; Enneti, 2012) and already used in the laboratory.

The objective of this study is to determine the behavior of binders used in MIM on an Inconel 718. For this, a study of the temperature behavior and chemical interactions between polymers was performed. The morphological characteristic of the powder used has also been studied.

2. MATERIALS AND METHODS

The polymers used for this study are polypropylene, polyethylene glycol and stearic acid (Urterkin at al., 2011). The PEG used in this study has a molar weight of 20000 g/mol. The polypropylene is a PP670Kh supplied by Sabic. The polymer blended are made in a Brabender twin screw mixer which have a volume capacity of 50 cm³ at a speed of 30 rpm and at a temperature of 180°C. The behavior of polymer and binder are further characterized by FTIR and DSC.

The powder used is an Inconel 718 atomized by argon. The powder was supplied by Sandvik Ospreys Ltd. The chemical composition of the powder is given by table 1. The morphological characteristic of the powder was characterized by laser granulometer and by krypton physical adsorption and by infrared spectroscopy. FTIR studies were performed on a Rheonant Resultec. This device couples a ThermoScientific Smart OMNI-Transmission Nicolet iS10 infrared spectrometer. This device allows for testing the temperature to 400°C, which is used to study the chemical and thermal behavior of polymers, binders and feedstocks. Temperatures used in tests are choose to be the nearest of the injection temperature conditions.

Fig. 1. Spectrum FTIR of PEG 20K at 40°C.
3. RESULTS AND DISCUSSION

3.1. Thermal behavior of PEG20K

The FTIR absorption spectrum of PEG at 40°C is shown in figure 1. PEG spectrum shows the characteristics peak of pure PEG (Finocchio et al., 2014). A broad band centered at 2890 cm⁻¹ corresponds to the CH₂ band. The peaks of the CH₂ are detected at 963 cm⁻¹ and 842 cm⁻¹ (rocking), 1467 cm⁻¹ (asymmetric deformation) 1360 cm⁻¹ and 1343 cm⁻¹ (wagging vibration mode), 1280 and 1242 (twisting), 1280 cm⁻¹ and 1242 cm⁻¹ (twisting) and 1148 (symmetric deformation). The other peaks correspond to the C-O-C stretching mode (1116 cm⁻¹) and to the C-OH stretching mode (1060 cm⁻¹).

The thermal decomposition of PEG investigated by FT-IR (figure 2) at increasing temperature shows the formation of decomposition products in the condensed phase characterized by a C=O stretching band at 1728 cm⁻¹, with a shoulder at 1754 cm⁻¹ (figure 2). These bands appear at 175°C and reach their maximum intensity between 200°C and 250°C. This effect indicates the formation of a complex mixture of decomposition products, likely low-molecular-weight fractions formed as a result of chain scission processes at the weak carbon-oxygen bonds in the PEG backbone.

3.2. Thermal behavior of polypropylene

Figure 3 shows the thermal behavior of the PP analyzed by FTIR. At 40°C, the spectra show the characteristic peaks of the polypropylene. The peaks at 1370 cm⁻¹ and 2964 cm⁻¹ correspond to the CH₃ deformation and stretching respectively. The peaks of the CH₂ are detected at 1465 cm⁻¹ (asymmetric deformation), 963 cm⁻¹ and 842 cm⁻¹ (rocking). A broad band centred at 2890 cm⁻¹ corresponds to the CH₂ band with the peak of the CH₃ at 2964 cm⁻¹. No differences are visible between the spectra at 40°C and the spectra at 200°C. The peaks are not shifted and no distorted. The result shows a stable behavior of the polypropylene at this temperature.

3.3. Mixing behavior

Figure 4 shows the temperature behavior of the mixture. The FTIR analysis shows that at 40°C the spectra obtained for the PEG60% + PP40% mixture is the superposition of the spectra of pure PEG and pure PP. The peaks
of the PEG and of the PP are not shifted and not distorted. This shows no chemical change and therefore no interactions between the two polymers. At 200°C the appearance of C = O peak of degradation of the PEG is visible, it means that the behavior of the PEG is not modified by the addition of PP. This result was confirmed by DSC, see figure 5. The DSC curve shows a melting temperature at 69°C for the PEG and at 151°C for the PP. These results are the same that the pure PEG and PP and show a non-interaction between the two polymers because the melting temperature of the polymer doesn’t change in the mix (Belhaneche-Bensemra & Bedda, 2001). These results also show an immiscibility of PP with the PEG (Belhaneche-Bensemra & Bedda, 2001).

Figure 6 shows the spectra of the mixture PP40%+PEG60% after the mixing process and after the injection process. This result shows the decomposition of the PEG with the presence of the carbonyl’s peak. Moreover the presence of the band of the C=C at 1640 cm$^{-1}$, which correspond to an intermediate component of the degradation’s reaction, shows the decomposition is not over after the mixing process. The spectrum of the PEG after the injection process, with the growth of the carbonyl’s peak, shows the decomposition going on during the injection phase.

Fig. 4. Spectra FTIR of PP40%+PEG60% at 140°C (a), 180°C (b) and 200°C (c).

Fig. 5. DSC curve of PP40%+PEG60%.
3.4. Powder analysis

The composition of the powder is given in table 1. The Inconel 718 powder was characterized by laser granulometer and nitrogen physical absorption to determine particle size and possible porosities. Table 2 shows the results. The SEM analysis figure 7 shows the shape of the powder which is confirmed spherical. Considering this shape and with the result of the specific area the size of the powder can be calculated. The result is given in the table 2. The result of the size of the powder is confirmed by the granulometer analysis, figure 8, which shows the size distribution. The fact that the result is near shows that the powder have not porosities and a proper surface quality.

Table 1. Chemical composition of Inconel 718 powder in weight percentage.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Ni</th>
<th>Cr</th>
<th>Fe</th>
<th>Nb</th>
<th>Mo</th>
<th>Ti</th>
<th>Co</th>
<th>Al, Cu, C…</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage</td>
<td>50-55</td>
<td>17-21</td>
<td>Bal.</td>
<td>4.75-5.25</td>
<td>2.8-3.3</td>
<td>0.65-1.15</td>
<td>1</td>
<td>&lt;1</td>
</tr>
</tbody>
</table>

Table 2. Particle size of inconel 718 calculated by BET method and specific area.

<table>
<thead>
<tr>
<th></th>
<th>Specific area, m²/g</th>
<th>Size BET, µm</th>
<th>d10, µm</th>
<th>d50, µm</th>
<th>d90, µm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0,1128</td>
<td>6,7</td>
<td>3,5</td>
<td>5,4</td>
<td>8,2</td>
</tr>
</tbody>
</table>

3.5. Feedstocks behavior

One formulation has been studied in this analysis. The mix was composed with the binder formulation (55% PEG + 40% PP + 5% SA) and the Inconel powder at a rate load of 60%. The result of the FTIR analysis is given in figure 9. The spectrum shows no interactions between the powder and the binder. In the spectra of the powder no peaks are detected. In the spectra at 40°C, the result of the FTIR analysis is
the same that the binder spectra (figure 4). The peaks are not shift and no distort sign of no interactions. The thermal behavior of the component is also unchanged. In the spectra in figure 4b the peak at 1728 cm⁻¹ of the decomposition of the PEG at 180°C is always present.

4. CONCLUSIONS

The present study has shown that the PEG is decomposed at the injection temperature in the classic MIM process conditions. The use of PP or Inconel 718 does not influence the thermal behavior of the PEG, the decomposition is always present in binder and feedstocks. These results show no chemical interactions and no influences between each component of the formulation. This fact may prove that the homogeneity of a MIM feedstock is not optimal. The PEG decomposition going on in the MIM process. The conditions of the process depend on the PP. This result shows the necessity to change the polypropylene on the formulation to conserve the characteristics of the PEG during the process. Further tests will be lead on the miscibility of the PEG with polymers and powder to improve the homogeneity and the characteristics of binders and feedstocks.

ACKNOWLEDGMENTS

The authors wish to thank the FUI ProPIM project for the financial support.

REFERENCES


WYKORZYSTANIE SPEKTROSKOPII W PODCZERWIENI DO CHARAKTERYSTYKI LEPISZCZA NA BAZIE GLIKOLU POLIETYLENOWEGO I WSADU Z INCONELU 718 DO FORMOWANIA WTRYSKOWEGO PROSZKU

Streszczenie

Formowanie wtryskowe, na przestrzeni ostatniej dekady, stało się konkurencyjnym procesem stosowanym do produkcji, w dużych ilościach, małych, precyzyjnych komponentów o złożonych kształtach, których wytworzenie byłoby kosztowne przy zastosowaniu innych procesów. Formowanie wtryskowe metali (ang. MIM) łączy wszechstronność procesu wtryskiwania z wytrzymałością i integralnością innych procesów wytwarzania, na
przykład tłoczenia, stosowanych do produkcji małych, złożonych elementów. Proces MIM składa się z nadania cząstkom proszku odpowiedniego kształtu a następnie spiekania ich. Podczas fazy wtryskiwania w proszku zachodzi segregacja i w materiale, w trakcie procesu spiekania, mogą pojawić się wady. Stąd opracowywane są systemy mieszania, pozwalające na ograniczenie tego zjawiska. Systemy te opisywane są jako mieszanina różnych substancji organicznych lub nieorganicznych o wielu funkcjach. System mieszania powinien zapewniać odpowiednie własności reologiczne materiału wsadowego, potrzebne do transportu cząstek proszku do formy i zachowania spójności tzw. części „zielonej”.

W pracy przeprowadzono badania nad oddziaływaniom chemicznymi między polimerem w lepiszczu i w materiale wsadowym Inconel z wykorzystaniem spektroskopii w podczerwieni z transformatą Fouriera (ang. FTIR) oraz różnicowej kalorymetrii skaningowej (DSC). Metody te zastosowano również do badania zachowania termicznego lepiszca w temperaturze zbliżonej do temperatury wtryskiwania. Analizy pokazały zależność między oddziaływaniom chemicznymi i mieszalnością polimeru a różnymi zachowaniami reologicznymi i mechanicznymi. Wszystkie metody nie wykazały żadnych oddziaływań pomiędzy komponentami lepiszca a proszkiem. Otrzymane wyniki pokazały konieczność opracowania nowej, lepszej formuły dla lepiszca w celu poprawy homogeniczności materiału wsadowego i zmniejszenia zjawiska segregacji podczas fazy wtryskiwania.

Received: September 30, 2014
Received in a revised form: November 3, 2014
Accepted: December 17, 2014
MODELLING OF CLINCHING JOINT PULL-OUT TEST

ZBIGNIEW GRONOSTAJSKI*, SŁAWOMIR POLAK, BARTOSZ BARTCZAK

Wrocław University of Technology, Łukasiewicza 27, Wrocław
*Corresponding author: zbigniew.gronostajski@pwr.wroc.pl

Abstract

The paper presents the model of the tests on joints obtained by clinching (mechanical joining) high-strength TRIP690 steels. A numerical finite element analysis of the joint with a bottom thickness of 0.6, 0.7 and 0.8 mm was carried out using the MSC.MARC&MENTAT implicit software. Similarly as in the physical tests, numerical analysis predicted that the joint can carry a greater normal force when the thickness of the bottom is reduced. The model can be used in further studies aimed at determining the optimum shape and strength of such joints. In the literature, two basic modes of failure are distinguished. The first is associated with insufficient material deformation, whereas the second, with the lack of material in the joint’s neck due to, for example, excessive displacement of the tools. For TRIP690 steel the only first mode of failure was observed.

Key words: high strength steel, FEM, clinching

1. INTRODUCTION

Mechanical press joining, also called clinching, is a method of joining in which parts of metal sheets are locally deformed without using any additional elements (Hahn & Horstmann, 2007; Varis, 2003). Press joining consists in the local pressing of one metal sheet into another in order to block the bottom of the pressed in sheet against the pressed out sheet. A protrusion forms in the pressed out sheet while a cavity forms in the pressed in sheet. The quality of such joints depends on the plasticity of the components being joined and on process parameters, including dimensions and shapes of the tooling and the magnitude of the applied pressure dependent on the materials being joined (Lee et al., 2010).

No heat is needed to produce this kind of joint. As opposed to conventional welding techniques, press joining proceeds without any heat effects on the layers of the materials being joined. A typical joint is formed in about 1 second. No preparation is needed, as opposed to many other joining methods, e.g., riveting requires the drilling of holes, gluing requires a bonding agent and a clean and rough surface and welding requires the preparation of the edges to be joined. A joint produced by press forming can be immediately subjected to load. The process is simple and without any unproductive time. Press joining can be used to join materials of different thickness and covered with different coatings (Oudjene & Ben-Ayed, 2008; Varis, 2006).

Studies have shown that the principal parameters determining the extent of failure of the joint are neck thickness and undercut width (figure 1) (Varis & Lepisto, 2003).

Clinched joints are increasingly often used in industry. The largest manufacturer of clinching tools is the Eckold company. It offers a wide range of press joining methods for different applications. Using Eckold tools one can join materials characterized by different properties.

Figure 2 schematically shows, how clinched joint R-PJ (round press-joining without cutting) is formed according to Eckold (Eckold). In order to make an R-PJ joint the metal sheets should be placed between the punch and the die and then a force
should be applied to the punch, causing the local deformation of the sheets in the pressing operation until the force increases as a result of the contact between the sheets and the die which thanks to its segmental structure widens perpendicularly to the movement of the punch. Consequently, the material flows in the radial direction while the bottom of the joint is being compressed (press-through operation). Then the punch is released and the segment die returns to the initial position (Eckold).

- an element joined by press joining is gas-tight,
- neat appearance; there is only a slight elevation of the metal sheet on the die side,
- this press joining variant is particularly suitable for press joining sandwich plates.

The strength parameters of the joints made by clinching are determined by the pull-out and tensile tests performed according to the standards (figure 3).

Fig. 1. Cross section through clinched joint R-PJ (a) and two types of joint failure (b),(c).

Fig. 2. Successive operations in formation of joint R-PJ (Eckold).

The special features and advantages of round press joining are:
- press joining without cutting,
- easy joining of materials characterized by different plastic properties,
- an element joined by press joining is gas-tight,
- neat appearance; there is only a slight elevation of the metal sheet on the die side,
- this press joining variant is particularly suitable for press joining sandwich plates.

The strength parameters of the joints made by clinching are determined by the pull-out and tensile tests performed according to the standards (figure 3).

Fig. 3. Strength testing: pull-out test (a) and tensile test (b).

Fig. 4. Dimensions of clinching.

The aim of this research was to model the pull-out test of clinching joint with dimensions presented in figure 4 made from sheet of TRIP690 steel.
2. NUMERICAL MODELLING

A numerical analysis was carried out using the finite element method (FEM) MSC.MARC&MENTAT employing the implicit method for solving the system of equations. The formation of the clinched joint and its subsequent deformation during the pull-out test were analyzed.

Stage 1 – modelling of clinched joint

The clinched joint is not ideally but axisymmetric. However, since it would be difficult to take into account the actual small circumferential variations a simplified axisymmetric computing model, consistent with the dimensions of the tools used in the real process (figure 5), was adopted. The thickness of the TRIP690 sheet was 1.5 mm. The stress-strain curves determined in tensile test are presented in figure 6. The coefficients of friction between tools and sheet \( \mu = 0.1 \) and between sheets \( \mu = 0.15 \) were taken from previous research.

![Fig. 5. Axially symmetrical process scheme.](image)

The shape of the joint obtained from mathematical modelling for 1.5 mm thick metal sheet TRIP690 is very similar to the actual joint (figure 7). The largest deformation occurred in the undercut region.

![Fig. 6. Stress-strain curve of TRIP690 steel for different strain rates.](image)

Fig. 7. Cross section of: a) mathematically modelled joint- strain distribution b) real joint.

Stage 2 – modelling joint deformation during the pull-out test

The clinched joint model developed in Stage 1 was then used to model the pull-out test in 2D and 3D (figure 8). In the literature, two basic modes of failure are distinguished (Varis, 2003). The first is associated with insufficient material deformation (figure 1b), whereas the second, with the lack of material in the joint’s neck due to, for example, excessive displacement of the tools (figure 1c). For TRIP690 steel the first mode of failure was observed.

The results of the simulation of joint tension are to a large extent consistent with the experimental results (figure 9). Similarly as in modelling, also in reality the joint undergoes local deformation and the top sheet slips off the bottom sheet as a result of plastic deformation in the undercut region.
greater normal force when the thickness of the bottom is reduced. The simulation results show slightly higher strength values than those obtained experimentally. The largest differences occur for the joints with a thicker bottom. For a 0.8 mm thick bottom the ultimate force amounts to 4 kN and 3.4 kN according to the simulation and the experiment respectively. For a 0.6 mm bottom the ultimate force according to the simulation amounts 4.2 kN, which is in full agreement with the experiment.

3. CONCLUSION

The strength of materials influences the failure mode of clinching joints. In the case of soft material
two modes are known. The first is associated with insufficient material deformation (figure 1b), whereas the second, with the low thickness of joint’s neck (figure 1c). For the high strength material such as the TRIP690 steel the only first mode of failure was observed.

A numerical finite element analysis of the clinching joint pull-out test was divided into two steps: the formation of the clinched joint and its subsequent deformation during the pull-out test. The joints were carried out with a bottom thickness of 0.6, 0.7 and 0.8 mm. Similarly as in the physical tests, the joint can carry a greater normal force when the thickness of the bottom is reduced. The model can be used in further studies aimed at determining the optimum shape and strength of such joints.

REFERENCES


NUMERICAL RESEARCH ON BRAKE CALLIPER PISTON’S WEIGHT REDUCTION

MIKHAIL PETROV, YULIAN PHILIPPOV, PAVEL PETROV*

University of Mechanical Engineering, Department of Machines and Metal Forming Technologies, B. Semyonovskaya 38, 107023 Moscow, Russian Federation
*Corresponding author: p.petrov@mami.ru

Abstract

Under understanding the necessity of the weight reduction without decrease in strength properties of a construction part the authors have tried to declare the methodology of the parts’ construction on the example of the brake calliper piston. Taking commonly used high strength aluminium wrought alloy of the 6xxx-serie the behaviour of the piston under plain compression load was investigated separately for elastic and plastic deformation. After that it was compared with the results for carbon steel C10. Under definition of the maximal allowed parameter’s value for elastic and minimal allowed parameter’s value for plastic deformations, represented by factor of safety (FoS) and yield stress respectively, the sensitivity maps of the proposed piston constructions were calculated. Opposite to the total topology optimization the outer geometry of the original piston was unchanged. The results of the study have shown that the changing of the inner geometry can be done in a certain range without any significant strength decrease of the construction part but not arbitrary.

Key words: aluminium wrought alloys 6061, carbon steel C10, brake calliper piston, construction optimisation, numerical simulation, FEM, QFORM 3D, T-Flex Analysis

1. INTRODUCTION

Braking systems is permanent developing system, which is responsible for the vehicle stop after human reaction and pushing the brake pedal or button. For different vehicle types it represented by different constructions. But only two main types of the breaks are widely exist, namely drum and disc brakes. Due to low weight, better heat dissipation and more durability under severe usage the disc brakes superior to drum brakes, that generally the state of art and described in many engineering books on the problem, e.g. by Heissing and Ersoy (2011), Gilles (2005). To protect the pressure fluid against overheating the calliper pistons have to be done regarding the material and geometry. Used in disc brake calliper pistons driven by fluid push the friction pads to clamp the rotor or brake disc during the movement of the brake pedal (figure 1). The generated heat will be dissipated inside the rotor and partly will be transmitted through the piston into the fluid. It causes the increase of the temperature inside the fluid volume and building the gas bubbles that prevent the movement of the brake pedal.

Today’s market is overflow with the pistons of different constructions, verified experimentally. There is no unique technique for construction of the optimal pistons due to distinguished driving mode of each car driver. And commonly there are only three main construction material groups with just a few representative materials per group, which are widely used through the numerous geometries of the pistons. These groups are grey cast iron, steel and aluminium. For flexible production more information about material properties regarding the application case is required. For example, the correlation between the geometry and density or weight of the piston is needed to calculate the generated and transmitted heat through the piston.
Fig. 1. Fixed caliper: 1 – brake disc, 2 – brake pistons, 3 – caliper (hydraulic connections, bleed screws and friction pads are not presented). 3D view (a), cross-section (b).

1.1. Fundamental ways of weight reduction of technical constructions

There are several methods in the optimization practice for making new construction parts. Among them the methods like topology optimization and lightweight materials’ application. Topology optimisation means that non-relevant material volume will be selected and neglected during the production stage. It means that the pistons with thin walls, smaller radiuses and complex surfaces can be created. Lighter materials can be also used but without decrease in strength of the end product. The most popular materials stay carbon steels and aluminium wrought alloys of the 5xxx, 6xxx and 7xxx-series thanks their good till excellent strength properties and high corrosion resistance against aggressive brake fluids.

The application of new material needs to be technically feasible that can be achieved by development of new technological processes, for instance processes based on the cold bulk metal forming. Philippov et al. (2011, 2013), Gnevashev et al. (2003), Philippov and Molodov (2012) have shown the practicable realisation of the numerical simulation of the developed cold bulk metal forming operations and good agreement of the simulation results with the experiments for many part produced by cold bulk forming operations.

1.2. Construction for brake calliper pistons under study

In the ideal case the new optimized piston should have the reduced weight and the same or higher strength values. The authors tried to couple in one theoretical function two main influence parameters (arguments), i.e. geometry and material. Each argument was considered commonly and designated as a criterion in the following equation:

\[ F = F(G, M), \]

where \( G \) – geometry criterion (G-criterion), \( M \) – material criterion (M-criterion).

Taking into account this fact and based on the two mentioned above methods the geometry and consequently the weight of the original mass produced pistons (figure 2, case A-1 and B-1) was changed. For the cases A-2 and B-2 the thick wall of the initial pistons was replaced by two thin walls. For the cases A-3 and B-3 the volume of the inner free space was increased. In both investigated modifications the replaced material volume was approximately the same (table 1). From the physical point of view two thin walls represented one connected shell and the induced during the brake operation stresses are closed inside it (compare A-1 vs. A-2). Stress distribution in one thin wall is comparable to one thick wall but with the correction on the ultimate material properties (compare A-1 and A-3). In principle the case A-2 can be converted into case A-3 then there is no gap between two thin walls of the piston in case A-2.

The number of investigated materials was restricted to two (one per main material group), namely aluminium alloy (EN AW-6061) and steel (carbon steel C10) with the chemical composition and mechanical properties presented in tables 2 and 3, respectively.
Table 1. Comparison of the weight characteristics of the pistons.

<table>
<thead>
<tr>
<th>Variant</th>
<th>Volume, mm³</th>
<th>Weight, kg</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>C10</td>
</tr>
<tr>
<td>A-1</td>
<td>42.050,8</td>
<td>0.328</td>
</tr>
<tr>
<td>A-2</td>
<td>26.800,2</td>
<td>0.209</td>
</tr>
<tr>
<td>A-3</td>
<td>27.411,2</td>
<td>0.214</td>
</tr>
<tr>
<td>B-1</td>
<td>35.003,9</td>
<td>0.273</td>
</tr>
<tr>
<td>B-2</td>
<td>21.412,3</td>
<td>0.167</td>
</tr>
<tr>
<td>B-3</td>
<td>21.321,1</td>
<td>0.166</td>
</tr>
</tbody>
</table>

1.3. Loading schemes

Two loading schemes were applied separately to the top surface: one to perform the static simulation in T-Flex Analysis (the problem is solved only in scope of Hooke’s law by the system of linear algebraic equations) and one to perform the simulation of the plastic deformation in QForm 3D. To check the construction reliability of each case, presented in figure 2 the maximal (in elastic area) and the minimal (in plastic area) values of deformation stresses or forces have to be determined.

Firstly the material of the piston has to withstand the middle force value of 10,000 N under uniaxial compression stresses, which are normally exists on the working surface rotor-friction pad with piston (figure 3) and are necessary boundary condition for elastic problem formulation.

The stress-strain state of the piston’s material during the brake pedal activation is quit the same as for the classical compression test except the temperature fields, which are induces by severe friction between the rotor represented by brake disk and pads, and small friction on the contact surfaces, which can be neglected. Secondly, if the material of the investigated piston obtains plastic deformation or stays close to that transient point, i.e. near to the yield stress point, the stresses increase inside the pistons.

Table 2. Chemical composition of the materials (GOST 4784-97, DIN EN 573-3).

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>Si</th>
<th>Mn</th>
<th>Ni</th>
<th>S</th>
<th>P</th>
<th>Cr</th>
<th>Cu</th>
<th>Al</th>
<th>As</th>
</tr>
</thead>
<tbody>
<tr>
<td>carbon steel C10</td>
<td>0.07-0.14</td>
<td>0.17-0.37</td>
<td>0.35-0.65</td>
<td>max. 0.3</td>
<td>max. 0.04</td>
<td>max. 0.035</td>
<td>max. 0.15</td>
<td>max. 0.3</td>
<td>max. 0.08</td>
<td></td>
</tr>
<tr>
<td>aluminium wrought alloy EN AW-6061</td>
<td>max. 0.7</td>
<td>0.4-0.8</td>
<td>max. 0.15</td>
<td>0.04-0.35</td>
<td>max. 0.15</td>
<td>0.15-0.4</td>
<td>0.8-1.2</td>
<td>max. 0.25</td>
<td>rest</td>
<td></td>
</tr>
</tbody>
</table>
material rapidly. As a result the factor of safety (FoS) against equivalent stresses tends to zero, follows from the equation (2) and (3). Additionally the authors specified a construction coefficient through equation (4), which should obviously correctly describe the dependency between critical deformation force and FoS.

\[ K_{fs} = \frac{\sigma}{\sigma_{eq}} \]  
\[ \sigma_{eq} = \sqrt{\frac{1}{2} \left[ (\sigma_x - \sigma_y)^2 + (\sigma_y - \sigma_z)^2 + (\sigma_z - \sigma_x)^2 + 6(\tau_{xy}^2 + \tau_{yx}^2 + \tau_{zx}^2) \right]} \]  
\[ K_c = K_{fs} \cdot F_d \]

where \( K_{fs} \) – factor of safety against equivalent stresses; \( \sigma_{eq} \) – equivalent stress value, which is calculated from the stress tensor components; \( [\sigma] \) – allowed stress value, which was equal to the flow stress of the material during the calculation, \( K_c \) – construction coefficient, \( F_d \) – deformation force.

If the same factor is equal to minimal (recommended) allowed value of 1–1.5 and is more higher the material works properly in elastic area and the construction is robust. For ultimate deformation, i.e. the situation, which can occur during the staff clamping of the friction pads, each piston was loaded according to the material’s flow law and the maximal obtained force, which resolves stresses to perform plastic deformation, was determined on the first calculation steps.

### Table 3. Mechanical properties of the materials (databases of the QForm and T-Flex).

| Parameter                        | EN AW-6061 |               | C10           |               |
|---------------------------------|-------------|---------------|---------------|
|                                 | T-Flex | QForm 3D   | T-Flex | QForm 3D   |
| Young module, GPa              | 69      | -            | 210             | -            |
| Poisson ratio                   | 0,33    | -            | 0,28             | -            |
| Thermal conductivity, W/(m K)   | 170     | 250           | 43              | 28           |
| Linear thermal expansion coeff. | 2,4     | -            | 1,3             | -            |
| Density, kg/m³                  | 2.700   | 2.800        | 7.800           | 7.550        |
| Flow stress, MPa                | 55,15   | Table function | 220,6           | Table function |
| Ultimate stress, MPa            | 124,1   | -            | 399,8           | -            |
| Heat capacity, J/(kg K)         | 1.000   | 1.230        | 440             | 649          |

### 2. NUMERICAL SIMULATION

Firstly, the elastic problem was numerically solved in finite element (FE) code T-Flex Analysis. It deals with the isotropic material with a specified in program’s database mechanical properties. Under isotropic material the material with an invariant with respect to orientation of the body in the space, i.e. identical properties in all directions, has to be understood. According to the loading schemes showed in the figure 3a full restrains were assigned to the bottom and side surfaces of the piston.

Following the stress-strain diagram the simulation runs were continued by the solving of the plastic problem, which was done with the help of the FE commercial code QForm v.7. It is based on flow formulation, with independent variables, represented by velocity vector and mean stress. In rigid-viscoplastic model the material is considered as incompressible, isotropic continua, whereby elastic deformations are neglected. The restrains were applied only to the top and bottom surfaces and represent the piston contacts with both tools (figure 3b).
2.1. Process setup

The simulation runs were done for normal environmental conditions, presented in the table 4. For investigation the material behaviour near yield stress point the hydraulic press was chosen. To eliminate the influence of the friction during the simulation the friction factor was set to zero.

Table 4. Process parameters (for QForm 3D and T-Flex Analysis).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tools’ temperature, °C</td>
<td>20</td>
</tr>
<tr>
<td>Environmental temperature, °C</td>
<td>20</td>
</tr>
<tr>
<td>Workpiece temperature, °C</td>
<td>20</td>
</tr>
<tr>
<td>Friction factor</td>
<td>0</td>
</tr>
<tr>
<td>Press nominal capacity, MN</td>
<td>1</td>
</tr>
<tr>
<td>Nominal ram velocity, m/s</td>
<td>0,001</td>
</tr>
</tbody>
</table>

2.2. Mesh preparation

All models were discretised by finite-elements generated automatically, whereby meshes based on the 4-node tetrahedral elements in QForm and on the 10-node tetrahedral elements with curved sides in T-Flex Analysis were prepared. Moreover finite-elements in T-Flex were generated automatically with global size propagation factor, which refers to the speed control of the mesh variation from reduced-size mesh cells to large cells of the general size. Its default value of 1 was accepted. This number means that the size of the element will double with each next calculation step until its size reaches the largest allowed mesh size. The generated mesh is not homogeneous and has clearly determined stacking places. During simulation this problem was corrected automatically just in QForm by a self-consistency algorithm of the program, that allows to increase mesh density locally for small materials volumes. Thanks to the internal algorithm of the program the number of the finite elements generated in T-Flex Analysis has not increased the calculation time.

3. RESULTS

As it was expected, the obtained simulation results have shown the dependency between the level of deformation force and geometry-material criteria. From the figure 4 it could be noticed that the pistons made from aluminium alloy did not reach the FoS of 1 – 1.5 and due to the construction recommendations such pistons could not be applied for real loaded brake systems. But also two representatives made from steel (A-2 and A-3) do not correspond this criterion.

Cases with the designations B-1 and B-3 are the most robust and trusted. They show that the piston with one thin wall (B-3) is much better than the piston with hollow area created by two thin walls (B-2), whereby the summarized wall thickness stays the same. Now analysing the maximal deformation forces, that can follow to the plastic deformation of the piston it is obviously that cases A-2 and B-2 cannot be applied due to high possibility of its plastic deformation during clamping stage and the A-3 case with low FoS can be applied conditionally because the construction is more rigid (figure 5). All investigated pistons’ constructions made from aluminium alloy could not be recommended for application due to too low total elastic ability.

Fig. 4. Minimal allowed value of the FoS (dashed horizontal line) against equivalent stresses show the fracture possibility of the piston (calculated in T-Flex Analysis).

There were two sensitivity maps calculated (figure 6), that had shown two separate areas for corre-
spondent material, whereby the first map (a) include only single function based on the elastic solution, since the second map (b) deals with both elastic and plastic solutions. It was found out that the same point of the diagram (e.g. B-2 for C10) changes its position and can move within the geometrical figure (circle or ellipse) of the certain radius.

**Fig. 5.** Maximal allowed values of the deformation force ($F_{de}$) before the plastic deformation take place (dashed horizontal line); reference value of 0.1 MN corresponds the loading force value of 10,000 N applied during the elastic problem simulation (calculated in QForm v.7).

4. SUMMARY

In the carried out investigation the authors tried to complete the fundamental task on the coupling of the deformation force with geometry and material. Numerical simulation was performed in two programs to solve different tasks from one hand – on searching of the maximal allowed elastic stresses represented by the FoS against equivalent stresses and from the other hand – the minimal resolved stresses follows to the undesirable plastic deformation of the piston. Both results had shown quit similar results and strongly cut the aluminium alloy as not desirable at least for the investigated geometry. This tendency can be changed if the wall thicknesses of the aluminium piston will be increased. Moreover the small gaps between the correspondent points from the figures 4 and 5 for the A-2 and B-2 cases confirm such assumption although the investigated cases are outside of the allowed ranges. The proposed construction coefficient $K_{c}$ is included into the M-criterion in equation (1) and can be the interface parameter between the elastic and plastic problem solution and also be used as a criterion for real constructions of the pistons for disc brakes in sense of the used materials.

**Fig. 6.** Sensitivity maps of the piston’s construction: one criterion (a), two criteria (b).

**REFERENCES**


WYKORZYSTANIE MODELOWANIA NUMERYCZNEGO DO BADAN NAD REDUKCJĄ WAGI TŁOKA W ZACISKU HAMULCA

Streszczenie

Rozumiając konieczność redukowania masy elementów części samochodu przy jednoczesnym zachowaniu ich własności wytrzymałościowych, autorzy pracy podjęli próbę opracowania metodologii konstrukcji tego typu elementów na przykładzie tłoku w zacisku hamulca. Do produkcji tłoka powszechnie stosowane są stopy aluminium z serii 6xxx, charakteryzujące się wysoką wytrzymałością. Stąd dla tego materiału przeprowadzono badania w płaskim stanie odkształcenia w zakresie odkształceń sprężystych oraz plastycznych. Następnie wyniki doświadczeń porównano z wynikami otrzymanymi dla stali węglowej C10. Na podstawie współczynnika bezpieczeństwa (FoS), definiowanego z wykorzystaniem maksymalnej dopuszczalnej wartości dla odkształcenia sprężystego i minimalnej dozwolonej wartości dla odkształcenia plastycznego, oraz granicy plastyczności, obliczono mapy wrażliwości dla zaproponowanych wariantów konstrukcji tłoków. W przeciwieństwie do zastosowania globalnej optymalizacji, wewnętrzny kształt oryginalnego tłoka pozostał niezmieniony. Wyniki przeprowadzonych badań pokazały, że zmiana wewnętrznegokształtu tłoka jest możliwa w pewnym zakresie bez znaczącego obniżenia własności wytrzymałościowych konstrukcji.
MEASUREMENT OF RESIDUAL STRESSES IN HOT-ROLLED STEEL SHEETS FOR LASER CUTTING

WOJCIECH SZYMAŃSKI1*, MARZENA LECH-GREGA1, MACIEJ GAWLIK1, ADAM KOKOSZA2, ADAM CHOCHOROWSKI3

1 Institute of Non-Ferrous Metals Light Metals Division, Piłsudskiego 19, 32-050 Skawina, Poland
2 AGH University of Science and Technology, al. Adama Mickiewicza 30, 30-059 Krakow, Poland
3 ArcelorMittal Kraków, ul. Ujastek 1, 30-969 Kraków, Poland
*Corresponding author: wszymanski@imn.skawina.pl

Abstract

Due to the specific nature of this process, the sheet metal designated for laser cutting must have appropriate chemical composition, structure, and low level of residual stresses. A key role in understanding the causes of the formation of residual stresses and in their control play various computer programs, targeted at this particular problem. One of the most important input data to such a programme is the distribution of residual stresses occurring in the material at different stages of the production process.

The most recognized method of measuring residual stress is the X-ray method $\sin^2\psi$, expensive and selective due to the measurement time. This method can be replaced with cheaper and faster, but less accurate and requires calibration method magnetic Barkhausen noise technique.

The purpose of this study was to compare the results of stress measurements taken by the X-ray $\sin^2\psi$ method with the results obtained by magnetic Barkhausen noise technique.

Key words: residual stress measurement, hot-rolled steel sheets, laser cutting

1. INTRODUCTION

High efficiency of the laser cutting process has rapidly increased the use of this method in recent years (Manohar, 2006). Due to the specific nature of this process, the sheet metal designated for laser cutting must have appropriate chemical composition, structure, and low level of residual stresses.

Previous experience gained in the development of laser cutting methods proves that in this process the state of the material is of primary importance (Totten et al., 2002). Studies show that the sheet metal behaviour during cutting process depends on factors such as the parameters of the steelmaking process, chemical composition of steel, and the type of applied thermo-plastic treatment. Properly adapted to the method of laser cutting, these factors help us obtain the material with special properties dedicated to this method. Their effect is noticeable especially when high cutting speeds are used during cutting of thick sections, in the punching operation, or when the quality of laser beam drops (clogging of lens and nozzles). Even if the sheet intended for laser cutting has a high uniformity of chemical composition, structure and mechanical properties, it still must meet two basic conditions for the proper run of the cutting process, namely, it should have (Andersen, 2000):

- very low level of residual stresses (even slight twisting of the sheet changes the focal point and angle of incidence, and this, in turn, changes the...
beam power at the point of contact with the processed material),
- adequate surface quality and uniformity of physicochemical properties (phenomena occurring on the surface are very important for the edge quality).

Achieving a low and stable level of residual stresses in the sheets is possible through the use of appropriate cooling conditions, checking the strip dimensions and using appropriate temper rolling at the end of the manufacturing process.

A key role in understanding the causes of the formation of residual stresses and in their control play various computer programs, targeted at this particular problem. A disadvantage of these programmes, however, is the difficulty in their adaptation to the technological process. Therefore, the aim of the scientist is to develop a dedicated software which would enable simulation of the development of microstructure and residual stresses to predict the final structure and mechanical properties in advanced hot rolling process of sheet metal. This will help control the level of stress in the sheets, significantly increasing their quality parameters. One of the most important input data to such a programme is the distribution of residual stresses occurring in the material at different stages of the production process.

The measurement of residual stresses is not a typical test method due to some difficulties in the interpretation of results and the specific type of apparatus used for these studies. Additionally, previous studies carried out on samples cut out from larger items using typical diffractometers would not reflect the correct state of stress in the material. Among numerous methods used for the measurement of residual stress, the X-ray \( \sin^2\psi \) method is the most recognised (Kocaña et al., 1990; Lech-Grega & Kłyszewski, 1991; Lech-Grega, 2003; Krasun et al., 2005; Jurcius et al., 2010). Therefore, the development of test apparatus is moving towards the possibility of measurements taken on large items. An apparatus of this type is the industrial, mobile X-ray diffractometer made by PROTO Manufacturing Ltd. Windsor, Ontario, Canada. Owing to this apparatus it will be possible to determine by the same method the state of stress in both laboratory samples and in the entire sheet. The results obtained by this method are, however, still expensive and selective. Therefore, this study compares the results of stress measurements by the X-ray technique with the results obtained using magnetic Barkhausen noise technique. This is faster and less expensive method to use because of the price of equipment, but it requires a complicated calibration and interpretation of the results obtained. Another key limitation is the fact that it can be applied only to ferromagnetic materials.

The purpose of this study was to compare the results of stress measurements taken by the X-ray \( \sin^2\psi \) method with the results obtained by magnetic Barkhausen noise technique.

2. APPARATUS

For the X-ray stress measurements, an X-ray diffractometer made by PROTO Manufacturing Ltd. (figure 1) was used. Both the measurement and the calculation method were in accordance with the SAE J84a and ASTM E915 standards. The apparatus is a portable X-ray diffractometer, fully computerised, used only for the measurement of residual stress and retained austenite content. All the measurement parameters of x-ray apparatus and tested material are shown in table 1.

Stress measurements using magnetic Barkhausen noise were taken with the STRESSCAN 500C apparatus, whose operation is based on the Barkhausen effect (figure 2). Tests included measurement of the value of the MP parameter. The higher is the value of this parameter, the higher (more positive) are the stresses. The data on the stress level were derived from the examination of layers with different thicknesses of 0.02 mm, 0.07 mm, and approx. 0.2 mm.

Table 1. Measurement parameters of x-ray apparatus and tested material.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value applied</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voltage</td>
<td>20 kV</td>
</tr>
<tr>
<td>Tube current</td>
<td>4 mA</td>
</tr>
<tr>
<td>X-ray tube applied</td>
<td>Cr</td>
</tr>
<tr>
<td>Length of radiation ( \lambda_{Cr} )</td>
<td>2.103, Å</td>
</tr>
<tr>
<td>Mirror planes</td>
<td>Fe 211</td>
</tr>
<tr>
<td>Measured angular range 2θ</td>
<td>145° - 165°</td>
</tr>
<tr>
<td>Size of the X-ray beam incident</td>
<td>5x1 mm</td>
</tr>
<tr>
<td>Tested material</td>
<td>ferritic steel</td>
</tr>
<tr>
<td>Bragg angle in unstressed material 2θ₀</td>
<td>156.41°</td>
</tr>
<tr>
<td>Young’s modulus</td>
<td>21·10³, kG/mm²</td>
</tr>
<tr>
<td>Poisson’s ratio</td>
<td>0.28</td>
</tr>
<tr>
<td>X-ray penetration depth</td>
<td>~10 μm</td>
</tr>
</tbody>
</table>
Fig. 1. PROTO iXRD X-ray diffractometer during the measurement of stress in sheet metal.

Fig. 2. STRESSCAN 500C apparatus for the stress measurement by magnetic Barkhausen noise method.

Fig. 3. The results of stress measurements taken in the rolling direction in sample of the N375900 strip, side A, prior to (a) and after (b) straightening. 0.2 K measurements by X-ray method, 0.2 and 0.02 measurements by magnetic method in layers of 0.2 and 0.02 mm thickness.
3. TEST MATERIALS AND RESULTS

Stress measurements were carried out on sample sheets of SA31 steel cut out from the strips prior to and after straightening. The study involved samples taken from two strips designated as N375900 and N375904 with the thickness of 4mm and 3mm, respectively. Each of the strips was 1500 mm wide. Samples were collected from the strips at a distance of approximately 15m from the beginning of the strip prior to and after straightening. Stresses were measured on the entire width of the strip on both sides in the direction of rolling.

The test results are shown in figures 3 – 6.

4. SUMMARY

Studies have shown that straightening reduced the level of residual stresses in the examined material in the rolling direction and caused homogenisation of their distribution across the width of the strip.

The results of stress measurements by magnetic Barkhausen noise technique expressed as an MP parameter and by the X-ray sin^2ψ method show a very similar nature of changes across the width of the strip (figures 3a-b, 4a, 5a-b and 6a), which indicates the correctness of the used test methods.

Fig. 4. The results of stress measurements taken in the rolling direction in sample of the N375900 strip, side B, prior to (a) and after (b) straightening. δKW-measurements by X-ray method, 0.2 and 0.02 measurements by magnetic method in layers of 0.2 and 0.02 mm thickness.
This means that when the magnetic method cannot be calibrated for the tested material, it is possible to calibrate it through the X-ray method. Then, the magnetic method can be used for routine monitoring under industrial conditions, as a cheaper and faster means of control than the X-ray method.

**Fig. 5.** The results of stress measurements taken in the rolling direction in sample of the N375904 strip, side A, prior to (a) and after (b) straightening. αKW-measurements by X-ray method, 0.2 and 0.02- measurements by magnetic method in layers of 0.2 and 0.02 mm thickness.
REFERENCES


Fig. 6. The results of stress measurements taken in the rolling direction in sample of the N375904 strip, side B, prior to (a) and after (b) straightening, δKW-measurements by X-ray method, 0.2 and 0.02- measurements by magnetic method in layers of 0.2 and 0.02 mm thickness.

POMIAR NAPRĘZEŃ WŁASNYCH W STALOWYCH BLACHACH GORĄCOWALCOWANYCH PRZEZNACZONYCH DO CIĘCIA LASEROWEGO

Streszczenie

Z uwagi na specyfikę cięcia laserowego, blachy do cięcia tą metodą muszą charakteryzować się odpowiednim składem chemicznym, strukturą oraz niskim poziomem naprężeń własnych. Kluczową rolę w zrozumieniu przyczyn powstawania i kontroli...
naprężenia własnych spełniły opracowane do tego celu programy komputerowe.

Jedną z najistotniejszych danych wejściowych do takiego programu jest rozkład naprężeń własnych występujących w materiale na różnych etapach procesu produkcji. Najbardziej uznaną metodą pomiaru naprężeń własnych jest rentgenowska metoda \(\sin^2\psi\), kosztowna, i ze względu na czas pomiaru, selektywna. Można próbować ją zastąpić tańszą i szybszą, ale mniej dokładną i wymagającą kalibracji, metodą magnetyczną szumów Barghausena.

Celem niniejszej pracy było porównanie wyników pomiaru naprężeń metodą rentgenowską \(\sin^2\psi\) z wynikami uzyskanymi metodą magnetyczną szumów Barkhausena.

Received: November 5, 2014
Received in a revised form: November 22, 2014
Accepted: December 2, 2014
DEVICE SIMULATION AND MULTI-OBJECTIVE GENETIC ALGORITHM-BASED OPTIMIZATION OF GERMANIUM METAL-OXIDE-SEMICONDUCTOR STRUCTURE

CHIEH-YANG CHEN, YIMING LI*

Parallel and Scientific Computing Laboratory, Department of Electrical and Computer Engineering, National Chiao Tung University, 1001 Ta-Hsueh Road, Hsinchu 300, Taiwan
*Corresponding author: ymli@faculty.nctu.edu.tw

Abstract

Germanium (Ge) and high-κ dielectric materials draw many attentions due to their fascinating electrical characteristics comparing with silicon (Si) material. However, in physical and electrical simulation, the physical model may have deviation to reality case due to the process condition and manufacturing technology. To computationally study the device with Ge material, it is necessary to optimize the theoretical result with experimental data. This paper originally provides a new method to examine the static characteristic of Ge metal-oxide-semiconductor field effect transistors (MOSFETs) with aluminum oxide (Al₂O₃) by integrating device simulation, multi-objective evolutionary algorithm (MOEA), and unified optimization framework (UOF). To deal with the realistic problem, especially for the steep change of capacitance, we consider not only residual sum of squares (RSS) (i.e. the sum of squares of residuals) function but also physically crucial points in the optimization problem. Comparing to single-objective genetic algorithm (GA) with a weighted fitness, the preliminary result of this study shows the method has great improvement to optimize the suitable parameters which not only minimize the RSS of capacitance but also agree the key capacitance values from physical view.

Key words: germanium MOSFET, aluminum oxide, fitting, capacitance-voltage curve, residual sum of squares, device simulation, genetic algorithm, multi-objective evolutionary algorithm, unified optimization framework, non-dominating sorting genetic algorithm (NSGA-II)

1. INTRODUCTION

While the semiconductor industry is evolving, germanium (Ge) is a raised and crucial material due to its higher carrier mobility comparing with silicon (Si) material. High carrier mobility can provide superior on-state characteristic of metal-oxide-semiconductor field effect transistors (MOSFETs). Also, the process of Ge MOSFETs has high compatibility with process of silicon. Various high-κ insulating films are studied to boost the device’s performance; among them, gate stack structures of Al₂O₃/GeO₂/Ge could form low-defect high-κ oxide layers, where the GeO₂ layer is a suitable interfacial layer for realizing superior high-κ Ge interfaces (Zhang et al., 2012). Based on the experimentally measured data, we can estimate and achieve even better device design by using semiconductor device simulation with calibrated physical and structural parameters. However, the simulation result with the same experimental parameter input may differ from experimental results. The deviations depend on the process environment and the inaccuracy of physical model, thus the parameter values in simulator should vary in certain range. To accurately extract the crucial parameters, cost-effective ways used to optimize the results between simulation and targeted goal are necessary. Genetic algorithm (GA), a kind of evolutionary algorithm, is a globally-searching optimization method in a large population and its evolution
process is based on the mechanism of natural selection. Recent applications of GA in semiconductor devices (Li et al., 2011) and electronic circuits (Li et al., 2013) were reported and have shown remarkable achievements on design optimization. For real-world engineering optimization problems, there are many goals to be optimized at the same time. Therefore, we have to introduce a weighting factor for all concerned goals so that a single-objective function can be modeled empirically. However, to determine weighting factors for each goal may lead the original optimization problem to additional uncertainty. Intuitively, multi-objective problem (MOP) can find the optimal population for each objective on its own dimension without using any weighting factors (Bansal et al., 2013; Hariharan et al., 2014). Therefore, in contrast to diverse try-and-error methods, it will be an interesting study for us to optimize capacitance-voltage (C-V) curve of Ge MOS devices by using a device simulated based MOEA methods.

In this work, we implement a new method to automatically optimize the capacitance of Ge MOSFET with Al₂O₃/GeO₂ film by using a semiconductor-device-simulation-based MOEA method running in the unified optimization framework (UOF) (Li et al., 2008). The capacitance optimization problem with material-, structure- and process-related parameters can be formed as a multi-objective optimization problem. By minimizing each objective, the presented method can simultaneously optimize seven designing parameters successfully. This paper is organized as follows. In Section II, we introduce the studied Ge MOS. In Section III, we brief the problem modeling and the semiconductor-device-simulation-based MOEA method in the platform of UOF. Then, we report and discuss the optimization results obtained from GA and MOEA methods. Finally, we conclude this study and suggest future work.

2. THE EVOLUTIONARY METHODOLOGY

2.1. The Structure of Germanium Metal-Oxide-Semiconductor

The primary high-κ material is Al₂O₃ in the experiment sample. The GeO₂ layer is naturally generated from germanium substrate and is a satisfactory interfacial layer for realizing high quality interface between high-κ Ge substrate. Figure 1 shows the simulation structure to calculate the electrical characteristic of the Ge MOSFET. The device’s channel length is fixed at 10 μm. The measured thickness of GeO₂ and Al₂O₃ layers are 3 nm and 5 nm, and the workfunction of metal gate is around 5 eV. We use two-dimensional (2D) device simulation with a small drain voltage and apply the gate voltage (V₉) ranging from -2 to 2 V. According to the physical observations, the entire gate capacitance – gate voltage (Cᵥ-V₉) curve can be partitioned into three regions: the accumulation, the depletion, and the inversion regions. Some of the parameters are flexible to be varied including the dielectric constant and thickness of Al₂O₃ and GeO₂, the parameter to adjust doping profile in germanium substrate, the source/drain doping level, and the workfunction of metal gate induced by nanosized grain orientation which may vary due to different treatments of fabrication process.

![Fig. 1. A 2D plot of the studied Ge MOS structure with the high-κ oxide layers. There are seven parameters to be optimized.](image)

Parameters to be optimized of the studied Ge MOS Structure

- Thickness of GeO₂ layer: \( \text{To}_{-}\text{GeO₂} \) (nm)
- Thickness of Al₂O₃ layer: \( \text{To}_{-}\text{Al₂O₃} \) (nm)
- The ratio parameter of Gaussian doping: \( \text{ratio}_g \)
- Doping concentration of the source/drain doping: \( \text{SD}_\text{dop} \) (cm⁻³)
- The dielectric constant of GeO₂: \( \varepsilon_{\text{GeO₂}} \) (CV⁻¹m⁻¹)
- The dielectric constant of Al₂O₃: \( \varepsilon_{\text{Al₂O₃}} \) (CV⁻¹m⁻¹)
- Workfunction of metal gate: \( \text{WK}_\text{Metal} \) (eV)

With the purpose of exploring the physically transport operation in Ge MOSFET, three governing equations which include Poisson equation and current-continuity equations for electron and hole transportation in semiconductor are solved self-consistently (Li, et al., 2002). The equations indicate that the number of carriers is conserved and the electrostatic potential due to the carrier charges obeys Poisson’s equation.
2.2. The Integrated Device Simulation, MOEA, and UOF Methodology

Based on our earlier developed unified optimization framework (UOF), we utilize integrated MOEA and numerical semiconductor device simulator to approach the experimental data of Ge MOSFET. This framework provides flexible interfaces to deal with different optimization method and numerical simulation solver, enabling these two different research fields to be bridged. The components of the UOF consist of several classes for different purpose; two primary classes are adapted for diverse problem and solver components, as shown in figure 2. The corresponding numerical simulation tool is defined in the problem class and the optimization algorithm is in the solver components. These two parts can be adjusted independently allowing high-level code to be reused, and rapidly adapted to other problems and algorithms.

Because the UOF is based on C++ language and consists of several classes, it has high coding flexibility. The class to call simulator and the class to call optimization algorithm are separated, thus we can achieve different simulators and algorithm to solve various problems. The program “main” governs the process step of optimization flow, as listed below.

```
void main
{
1. Set max iteration count, min error, population size, cross rate, mutation rate
2. Readfiles( fileList ) // file list contains
   // the name of intermediate file, input
   // and output file of simulator
3. ReadSetting( configure file) // configure
   // file contain cross/mutation rate,
   // command for simulator, and the key
   // word in output files.
4. Problem GeMOSCV()
5. Algorithm NSGAII()
6. NSGAII→iteration
}
```

To achieve the device characteristic of the individuals, as listed below, we replace the values in simulator file by the values of population. For each population, the simulation parameter values are set and then the device characteristic is solved by simulation tool. While the simulation is finished, the value of characteristic are obtained by simulator and recorded the results in its own output files. The sub
function defined in the problem class can receive the values in certain output file. Then, it calls the function defined in algorithm to assign fitness of the individuals.

```cpp
void Problem::Run_Simulator
{
1. Remove old results
2. Real values of individuals
3. Replace these values into intermediate file
4. Call TCAD and solve Poisson and electron/hole current continuity equations
5. Collect Result
}
```

In this work, we implement the non-dominating sorting genetic algorithm (NSGA-II) in algorithm class. The computational complexity of NSGA-II is \(O(MN^2)\), where \(M\) is objective number and \(N\) is the population size. In addition to technically sorting by several fronts, NSGA-II defines crowd distance to ensure that the Pareto front can keep a good diversity. The components of algorithm class are listed below.

```cpp
Class Algorithm
{
void Selection, Crossover, Mutation
void Evaluation
bool Dominate
void CrowdDistance
}
```

In this work, we consider several objectives, including the six crucial values of the C-V curve. The \(C_{L_{\text{max}}}\) is the capacitance value for the lowest measured gate voltage, the \(C_{\text{slope}}\) is the most negative slope for capacitance decrease while the gate voltage increase, the \(V_{C_{\text{slope}}}\) is the gate voltage where the capacitance is \(C_{\text{slope}}\), the \(C_{\text{min}}\) is the minimum capacitance value of entire curve and the \(V_{C_{\text{min}}}\) is its corresponding gate voltage, and the \(C_{R_{\text{max}}}\) is the capacitance value for the highest measured gate voltage. The deviations of these values between the simulation result and given specs are considered as objectives directly. Our ultimate goal is achieving the goal capacitance perfectly by adjusting the most crucial material/structure parameters, so we define RSS as the sum of square to summarize the capacitance-value deviation between the simulation result and the given specification. The seven objective functions are all minimization problem since the results can fit the target well if all deviations are zero. The objectives are given by:

\[ \text{Objective } _1 \equiv C_{L_{\text{max}}}, -C_{L_{\text{max}, \text{spec}}} \]  
\[ \text{Objective } _2 \equiv V_{C_{\text{min}}}, -V_{C_{\text{min}, \text{spec}}} \]  

and

\[ \text{Objective } _7 \equiv \min \sum_{v=2}^{7} r_v(C_v)^2 \]  

where \(r_v(C_v) = C_v - C_{v, \text{spec}}\). In this study, two strategies are applied to include these objectives; one is combining these objectives by weighting coefficients while another is directly using multi-objective evolutionary algorithm with all objectives at the same time.

3. RESULTS AND DISCUSSION

The optimization problem in this work is non-linear and is not exactly formulated, that’s the reason we use the simulation-based technique. If we use GA and consider only RSS, it’s hard to promise an acceptable deviation between simulation result and specification. To lead the evolution process to match the spec curve eventually, we consider physical key parts of C-V curve as well as RSS and use empirical weighted sum to model the problem into a single-objective fitting problem. The achieved results from GA and MOEA are shown in figure 4, we select the result with the minimum RSS in the population of the last generation. This diagram shows that adding those key consideration points to adjust the fitness can achieve desirable approach with small RSS by both two methods. Because these objectives have mutual interference and relate to different set of parameters, the fitting curve has inevitable difference to the specified curve. Once we have some objectives optimized by this way, the results are based on this set of weighting parameters. If we try to optimize other objectives, it is a difficult process to obtain another set of weighting coefficients. Although the RSS can be minimized by more evolutionary generations, the shape of C-V curve cannot be satisfied from device engineering’s view point for some gate bias region. Although the objective can be scaled and emphasized, the results still cannot highlight those important points simultaneously. The MOEA which use all objectives directly can achieve even better optimized result comparing with GA
method. The optimized curve fits the specified curve well everywhere. It is worth mentioning that the generation of MOEA method still keeps great diversity, meaning that the population will be continuously improved for more generations. In particular, results of the neighborhood of $C_{\min}$ which is the most physically complex point in the curve can be improved. We can find strong coupling between the capacitance and designing parameters around $C_{\min}$ region. As shown in figure 4, the comparison of results between GA and MOEA indicates the semiconductor-device-simulation-based MOEA method is computationally effective to handle the complex curve shape with suitable objectives. The reason is that the MOEA method can approach several characteristic with respect to each objective function; consequently, the optimization process governed not only by RSS but also the physical objectives.

Figures 5a-d show the optimized $|C_{\min}-C_{\min,\text{spec}}|$ versus $|C_{\text{slop}}-C_{\text{slop,spec}}|$ from the MOEA. We observe two important objectives, $|C_{\min}-C_{\min,\text{spec}}|$ and $|C_{\text{slop}}-C_{\text{slop,spec}}|$ which determine the behavior near the point $C_{\min}$ and are mutual interference. These four plots show the results of MOEA method under different crossover rate and mutation rate. The population moves toward smaller objective value with more iteration. The achieved Pareto fronts under four optimization settings are very similar. If we...
increase mutation rate, as shown in figures 5a and b, individuals may move away from the Pareto front to find other potential optimizers. On the other hand, high crossover rate makes the individuals evolve near the Pareto front and achieve smooth Pareto front surface, as shown in figures 5a and c.

4. CONCLUSIONS

In this study, based upon a unified optimization framework, an alternative way to optimize high-κ Ge MOSFET has been reported by using a semiconductor-device-simulation-based MOEA method. The MOSFET’s $C_G$-$V_G$ curve optimization problem has been modeled as a MOP with simultaneously considering capacitance’s physical constraints. Different from conventional try-and-error tuning method in device engineering, the main findings of this study indicate that the semiconductor-device-simulation-based MOEA method not only enables us to approach the capacitance in accumulation, depletion, and inversion regions of Ge MOSFET, but also achieves better results than the GA’s results. Notably, the MOEA method can effectively capture the complicated curve’s behavior compared with the GA method. The GA with a single-objective function strongly depends on empirically estimated weighted sum of multi- objectives which is difficult for real-world applications. We are currently using this method to optimize and design new experiments.

ACKNOWLEDGEMENTS

This work was supported in part by Ministry of Science and Technology, Taiwan under Contract No. NSC 102-2221-E-009-161, No. MOST 103-2221-E-009-180, and by tsme, Hsinchu, Taiwan under a 2012-2013 grant.

REFERENCES


SYMULACJA I WIELOKRYTERIALNA OPTYMALIZACJA ALGORYTMEM GENETYCZNYM PÓŁPRZEWODNIKOWYCH STRUKTUR TENKU GERMANU

Streszczenie

German (Ge) i materiały o wysokiej stałej dielektrycznej są interesujące ze względu na swoje niezwykle ciekawe charakterystyki elektryczne w porównaniu do krzemu (Si). Jednakże w symulacjach fizycznych i elektrycznych, model fizyczny może odbiegać od przypadku rzeczywistego ze względu na warunki procesu i technologię produkcji. Badania z wykorzystaniem metod obliczeniowych dla urządzeń wykonanych z germanu wymagają optymalizacji wyników teoretycznych z danymi doświadczalnymi. W pracy zaproponowano metodę badania statycznych charakterystyk tranzystorów polowych z tlenku germanu – półprzewodnika (MOSFETs) z tlenkiem aluminium (Al2O3), z wykorzystaniem zintegrowanego systemu składającego się z urządzenia do symulacji charakterystyk elektrycznych półprzewodników, wielokryterialnego algorytmu ewolucyjnego (MOEA) oraz zunifikowanej platformy do optymalizacji (UOF). Dla rozwiązania rzeczywistego problemu, zwłaszcza przy gwałtownej zmianie pojemności, w zadaniu optymalizacji rozwiązano nie tylko sumę kwadratów reszt (RSS), ale również kluczowe, z punktu widzenia fizyki, aspekty. W porównaniu z jednokryterialnym algorytmem genetycznym (GA) z ważoną funkcją dopasowania, wyniki przeprowadzonych badań pokazały, że opracowana metoda, która minimalizuje nie tylko błąd RSS dla pojemności, ale także bierze pod uwagę kluczowe wartości pojemności z fizycznych obserwacji, znacznie poprawiła zadanie optymalizacji wybranych parametrów zagadnienia.

Received: October 10, 2014
Received in a revised form: November 3, 2014
Accepted: December 12, 2014