



3D Modeling and Testing of Contact Problems in Resistance Welding

Nielsen, Chris Valentin; Zhang, Wenqi

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Nielsen, C. V., & Zhang, W. (2013). *3D Modeling and Testing of Contact Problems in Resistance Welding*. Technical University of Denmark.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

3D Modeling and Testing of Contact Problems in Resistance Welding

PhD Thesis

Chris Valentin Nielsen

Department of Mechanical Engineering
Technical University of Denmark

September 2012

PhD Student:

Chris Valentin Nielsen
Department of Mechanical Engineering, Manufacturing Engineering
Technical University of Denmark
E-mail: cvni@mek.dtu.dk

Supervisor:

Professor Niels Bay
Department of Mechanical Engineering, Manufacturing Engineering
Technical University of Denmark
E-mail: nbay@mek.dtu.dk

Co-supervisor:

President Wenqi Zhang
SWANTEC Software and Engineering ApS
E-mail: wz@swantec.com

Co-supervisor:

Professor Paulo A.F. Martins
Department of Mechanical Engineering, Instituto Superior Técnico
Technical University of Lisbon, Portugal
E-mail: pmartins@ist.utl.pt

Funding:

Department of Mechanical Engineering, Technical University of Denmark

The Danish Research Agency, FTP INNOJoint project at the Technical University of Denmark

SWANTEC Software and Engineering ApS

Preface

This thesis is submitted in partial fulfillment of the requirements for obtaining the degree of PhD at the Technical University of Denmark (DTU). The PhD project was carried out in the period October 1st, 2009 – September 30th, 2012. The supervision of the project was shared among Professor Niels Bay as main supervisor and co-supervisors PhD Wenqi Zhang, President of SWANTEC Software and Engineering ApS and Professor Paulo A.F. Martins. All three supervisors are deeply thanked for their engagement and the environment spanned among them. Niels Bay always kept the project on the right track by his supervision while leaving a large degree of freedom for me to explore several areas related to the topic. Wenqi Zhang provided a strong connection to the industry that has been a great motivation throughout the project, and at the same time, he has been largely involved in supervision of the development due to his experience with modeling of the resistance welding process. Paulo Martins supervised with outstanding engagement and provided a broad modeling view to the project.

The majority of the project was carried out in the Section of Manufacturing Engineering at the Department of Mechanical Engineering at DTU, where Professor Niels Bay has his research group. Another part of the project was carried out in SWANTEC by means of regular and close contact during the whole project period. Finally, two abroad stays of total 16 weeks were spent in the Department of Mechanical Engineering, Instituto Superior Técnico, Technical University of Lisbon, Portugal (IST), being hosted by Professor Paulo Martins and his group. The environment comprised by these three units has presented great opportunities for the development and research presented in this thesis.

Thanks to a lot of people would be necessary to be fair, but to keep it short special thanks go to the following from DTU: Office mates Pavel Müller, Ermanno Ceron, Alessandro Godi, Jais Andreas Breusch Angel and Peter Søe Nielsen are appreciated for contributing to the nice environment with many different discussions, PhD Kasper Storgaard Friis is thanked for valuable discussions during the finishing of his PhD within resistance welding that overlapped the beginning of my PhD studies. Help and assistance during laboratory tests and analyses by Lars Peter Holmbæk, Peter Sanderhoff and Steffen Sonne Munch are acknowledged. Søren Giversen is thanked for his introduction and advice in operating a GLEEBLE test machine. Finally, Associate Professor Chresten Træholt is thanked for his clarifying answers related to an electrical control unit.

Kim Richardt Pedersen (former SWANTEC) has taken part in many valuable discussions and is highly appreciated for his process knowledge and broad knowledge on materials and material properties. The periods of abroad stays at IST are among the best periods of the project for many reasons. First of all it was a pleasure to work intensively with Professor Paulo Martins, who offered a unique possibility of working closely together on a topic to which we share the interest. Besides working with Paulo Martins, I enjoyed my time in his group with special thanks to Associate Professors Beatriz Silva and Luis Alves.

A number of companies have provided examples of their industrial cases, which have enriched the present work, both in terms of presentation and in terms of motivation for further developments. Although the majority of companies cannot be thanked by their names, their willingness to share their cases is very much acknowledged. One company can be mentioned due to a collaborative publication; namely ThyssenKrupp AG, Germany, represented by Dr. Azeddine Chergui who shared his ideas in valuable discussions.

At last, I wish to express my thanks to my closest family and my girlfriend Jiayuan Wang for understanding and patience during the project.

September 30th, 2012



Chris Valentin Nielsen

Abstract

A generic, electro-thermo-mechanically coupled finite element program is developed for three-dimensional simulation of resistance welding. The developed computer program has reached a level of a complete standalone software that can be utilized as a tool in the analysis of resistance welding processes. Contact between deformable objects is modeled by the penalty method to handle multiple objects that appear in joining processes. Two algorithms are implemented for the identification of contact pairs. Simulations are presented for two metal forming processes involving contact and a number of resistance welding processes, which cover a wide range of spot welding and projection welding applications. Three-dimensional simulation of spot welding enables the analysis of critical effects like electrode misalignment and shunt effects between consecutive spots. A single-sided spot welding case involving three-dimensional contact is also presented. This case was suggested by and discussed with a German steel manufacturer.

When it comes to projection welding, a natural need for three-dimensional analysis arises in many cases because of the involved geometries. Cross-wire welding and welding of square nuts to sheets by projection welding are presented by means of experiments and simulations. These two cases are used to explore the capabilities of the developed simulation software by comparing experiments and simulations. A number of other projection welding cases are presented for further application of the software. These include joining of parallel sheets by circular projections and joining of perpendicular sheets by longitudinal projections. In the former case, the effects of unequal projection heights are analyzed, and in the latter case, the simulations are compared to the corresponding experiment by a Japanese company that proposed the case. Another industrial case, by a German company, is joining of micro components. The joining is based on mechanical locking, and the deformation is accommodated by resistance heating, which at the same time is used to melt a polymer coating locally for creating electrical contact necessary for the end-product.

All the above cases are modeled by meshing techniques included in the computer program. Structured, isoparametric meshing is utilized for setting up initial meshes of individual objects. Unstructured, all-hexahedral meshing is utilized for creating initial meshes of objects defined by CAD surfaces and is applied for remeshing of selected objects in order to carry on certain simulations. The all-hexahedral meshing procedures are enhanced by adaptive bounding boxes, facilities for handling multi-object simulations and overall improved by applying topology based criteria in the creation of hexahedral meshes. Simulation time is significantly reduced by a developed parallel skyline solver. The new solver is developed for shared memory and can be implemented in existing finite element codes by changing the call to the solver, as long as the system matrix is prepared in skyline format.

Finally, the above models and procedures are operated by a developed graphical interface including its own pre and post processing facilities. This combines the above into a new complete, standalone software: SORPAS 3D.

Resume (in Danish)

Et generelt, elektrotermomekanisk koblet finite element-program er udviklet med henblik på tredimensionel simulering af modstandssvejsning. Det udviklede computerprogram har nået et niveau hvor det som et komplet og selvstændigt software kan bruges som et værktøj ved analyse af modstandssvejsprocesser. Kontakt mellem deformerbare emner er modelleret ved hjælp af penalty-metoden sådan at flere objekter, som optræder i samleprocesser, kan håndteres. To algoritmer er implementeret til identificering af kontaktpar. Simulationer af to metalformgivningsprocesser, som inkluderer kontakt, er præsenteret sammen med et antal simulationer af modstandssvejsprocesser, herunder punktsvejsning og pressvejsning. Tredimensionel simulering af punktsvejsning muliggør analyse af kritiske effekter som skævt indstillede elektroder og parallelkobling mellem fortløbende punktsvejsninger. Anvendelse af punktsvejsning fra én side med tredimensionel kontakt er også inkluderet. Dette eksempel er foreslået af og diskuteret med en tysk stålproducent.

Når det gælder pressvejsning opstår der ofte naturligt behov for tredimensionel analyse på grund af de involverede geometrier. Eksperimenter og simulationer er sammenlignet for svejsning af krydsede tråde og svejsning af firkantmøtrikker til plader. Disse to eksempler er brugt til at undersøge mulighederne med det udviklede simulationsprogram. Et antal øvrige pressvejseksempler inkluderer samling af parallelle plader ved hjælp af cirkulære projektioner samt samling af plader vinkelret på hinanden gennem aflange projektioner. I førstnævnte eksempel er effekterne af uens højde af projektionerne undersøgt, og i sidstnævnte eksempel er der sammenlignet med et tilhørende eksperiment fra en japansk virksomhed, som også har foreslået analysen. En anden industriel anvendelse, i en tysk virksomhed, er samling af mikrokomponenter. Samlingen er baseret på mekanisk fastlåsning, hvor deformationen er hjulpet på vej af elektrisk genereret varme. Varmen er samtidigt brugt til at smelte en polymercoating lokalt sådan at der dannes elektrisk kontakt, som er nødvendig for det færdige produkt.

Alle ovenstående eksempler er modelleret med finite element-diskretisering baseret på teknikker inkluderet i computerprogrammet. Struktureret, isoparametrisk netgenerering er anvendt til at diskretisere startgeometrierne af enkeltobjekter. Ustruktureret, "all-hexahedral"-netgenerering er anvendt til diskretisering af startgeometrier af enkeltobjekter defineret af CAD-overflader og er ligeledes anvendt til regenerering af nettet for udvalgte objekter for at kunne fortsætte visse simulationer. Fremgangsmåden i "all-hexahedral"-netgenereringen er forbedret med adaptive valg af indre net, håndtering af simulationer med flere objekter og overordnet forbedret ved at introducere topologibaserede kriterier i opbygningen af net med sekssidede elementer. Simulerings-tiden er kraftigt nedsat via en udviklet parallel skyline-løser. Den nye ligningsløser er udviklet til delt hukommelse og kan direkte implementeres i eksisterende finite element-koder, bare ligningssystemet er forberedt i skyline-format.

Slutteligt, betjenes ovenstående modeller og fremgangsmåder via en nyudviklet grafisk brugerflade som indeholder egne for- og efterbehandlingsfaciliteter. Dette kombinerer ovenstående til et nyt, komplet og selvstændigt software: SORPAS 3D.

Contents

Preface	iii
Abstract	v
Resume.....	vii
Contents.....	ix
Nomenclature	xiii
1. Introduction	1
1.1 Resistance welding	1
1.2 Applications of resistance welding	3
1.3 Background and motivation.....	4
1.4 Outline of remaining chapters.....	6
References	6
2. Finite Element Formulations.....	9
2.1 Quasi-static formulations.....	11
2.2 Dynamic formulation.....	14
References	16
3. Coupled Finite Element Flow Formulation.....	17
3.1 State-of-the-art	17
3.2 Theoretical background	19
3.2.1 Plastic flow	19
3.2.2 Heat transfer	22
3.2.3 Electricity	24
3.3 Numerical implementation	25
3.3.1 Basic coupling procedures.....	25
3.3.2 Finite elements.....	26
3.3.3 Mechanical model	27
3.3.4 Thermal model.....	32
3.3.5 Electrical model.....	33

3.4	Incorporation of anisotropy	33
3.4.1	Rotation between global axes and material axes	37
3.5	Incorporation of elastic effects	38
	References	42
4.	Contact Modeling	45
4.1	Contact between workpiece and tooling	46
4.2	Contact between deformable objects	48
4.2.1	Frictionless contact	52
4.2.2	Sticking contact	54
4.2.3	Frictional contact	54
4.2.4	Electrical and thermal contact	55
4.3	Physical contact properties	56
4.3.1	Electrical contact resistance	56
4.3.2	Thermal contact resistance	59
	References	60
5.	Meshing and Remeshing	61
5.1	Description of tooling	62
5.2	Isoparametric structured meshing	64
5.3	All-hexahedral unstructured meshing	65
5.3.1	Identification of geometric features	65
5.3.2	Adaptive core mesh and reconstruction of surfaces	66
5.3.3	Reconstruction of vertices	68
5.3.4	Reconstruction of edges	70
5.3.5	Edge repairment	73
5.3.6	Smoothing	75
5.3.7	Application of all-hexahedral meshing	77
5.4	Addition of interface layers	77
5.5	Remeshing	79
5.5.1	Multi-object procedures and tool contact	80
5.5.2	Transfer of history dependent variables	81
	References	83
6.	Parallel Equation Solver	85
6.1	Strategies of solution techniques	85
6.2	Parallel skyline solver	86
6.2.1	Brief overview of equation solving by Gauss elimination in skyline storage format	87
6.2.2	Parallelization of skyline solver	90
6.3	Comparison of skyline solver with other solvers	97
6.4	Performance evaluation of parallel skyline solver	99
6.4.1	Evaluation by a resistance welding case	102
	References	107

7. Material Characterization	109
7.1 Mechanical properties at room temperature	109
7.2 Friction characterization	111
7.3 Mechanical properties at elevated temperatures	113
7.4 Electrical contact properties.....	114
References	116
8. Verification and Applicability of Mechanical Contact Model.....	119
8.1 Verification by contact experiments	119
8.1.1 Experimental setup	119
8.1.2 Materials	120
8.1.3 Comparison between simulations and experiments	122
8.2 Forming of seamless reservoirs	128
8.3 Mechanical joining of tubes.....	130
References	133
9. Challenges in Modeling Projection Welding.....	135
9.1 Projection welding of cross-wire by natural projection.....	135
9.1.1 Experimental analysis.....	136
9.1.2 Numerical modeling	148
9.2 Projection welding of square nut to sheet.....	155
9.2.1 Experimental analysis.....	155
9.2.2 Numerical modeling	165
References	168
10. Applications.....	171
10.1 Resistance spot welding.....	171
10.1.1 Three sheet spot welding.....	172
10.1.2 Electrode misalignment.....	174
10.1.3 Shunt effect.....	176
10.2 Single-sided spot welding of sheet to tube	179
10.2.1 Numerical simulations of sheet-to-tube welding.....	182
10.3 Projection welding with unequal projection heights.....	188
10.4 Projection welding by longitudinal embossment.....	194
10.5 Welding of bellow to disc by natural projection.....	196
10.6 Micro joining of fork and wire	201
References	204
11. Conclusions and Future Work	207
Appendix A.....	213

Nomenclature

α	Angle (Chapter 5)	$[\circ]$
α	Ratio of real contact area to nominal contact area	$[-]$
α_{conv}	Convergence tolerance	$[-]$
α_D	Direct iteration deceleration coefficient	$[-]$
α_i	Area coordinates	$[-]$
α_{NR}	Newton-Raphson deceleration coefficient	$[-]$
β	Dimensionless parameter	$[-]$
δ_{ij}	Kronecker delta	$[-]$
ε_{emis}	Emission coefficient	$[-]$
ε_{ij}	Strain tensor	$[-]$
$\dot{\varepsilon}_{ij}$	Strain rate tensor	$[s^{-1}]$
$\dot{\varepsilon}$	Effective strain rate	$[s^{-1}]$
$\dot{\varepsilon}_0$	Cut-off strain rate	$[s^{-1}]$
$\dot{\mathbf{\varepsilon}}$	Strain rate vector	$[s^{-1}]$
$\dot{\mathbf{\varepsilon}}_{mat}$	Strain rate vector in material coordinate system	$[s^{-1}]$
ζ	Element natural coordinate	$[-]$
η	Efficiency (Chapter 6)	$[-]$
η	Element natural coordinate	$[-]$
θ	Angle (Chapter 5)	$[\circ]$
θ	Dimensionless parameter	$[-]$
κ	Thermal resistivity	$[K \cdot mm / W]$
κ_c	Thermal contact resistivity	$[K \cdot mm / W]$
λ_L	Lagrange multiplier	$[MPa]$
$\dot{\lambda}$	Proportionality factor	$[(MPa \cdot s)^{-1}]$
μ	Friction coefficient	$[-]$
ν	Poisson ratio	$[-]$

ξ	Element natural coordinate	$[-]$
Π	Energy rate functional	$[W]$
ρ	Electrical resistivity	$[\Omega mm]$
ρ_c	Electrical contact resistivity	$[\Omega mm]$
$\rho_{contaminant}$	Electrical resistivity due to contaminants	$[\Omega mm]$
ρ_m	Mass density	$[kg / mm^3]$
σ	Flow stress	$[MPa]$
σ	Speed-up (Chapter 6)	$[-]$
σ_{ij}	Stress tensor	$[MPa]$
σ'_{ij}	Deviatoric stress tensor	$[MPa]$
σ_m	Mean stress	$[MPa]$
σ_{SB}	Stefan-Boltzmann coefficient	$[W / (K^4 \cdot mm^2)]$
$\bar{\sigma}$	Effective stress	$[MPa]$
σ'_{mat}	Deviatoric stress vector in material coordinate system	$[MPa]$
τ_f	Friction shear stress	$[MPa]$
τ_r	Friction shear stress in real area of contact	$[MPa]$
φ	Contact algorithm identifier	$[-]$
Φ	Electric potential	$[V]$
Φ_d^c	Potential difference in contact pair c	$[V]$
Φ	Vector of nodal potential	$[V]$
$\dot{\omega}_{ij}$	Spin rate tensor	$[s^{-1}]$
A	Area	$[mm^2]$
A	Candidate node (Chapter 5)	$[-]$
A	Dimension of Rastegaev specimen (Chapters 7-8)	$[mm]$
A_c	Nominal contact area	$[mm^2]$
\mathbf{A}_φ	Matrix of area coordinates	$[-]$
B	Pre-strain	$[-]$
\mathbf{B}	Strain rate-velocity matrix	$[mm^{-1}]$
c	Contact pair identification	$[-]$
c_e	Velocity of propagation	$[mm / s]$
c_k	Center Gauss point field variable	$[/]$
c_m	Heat capacity	$[J / (kg \cdot K)]$
C	Flow stress constant	$[MPa]$

C_0	Characteristic element side length of core mesh	$[mm]$
C	Matrix form of Kronecker delta	$[-]$
C_c	Heat capacity matrix	$[J / K]$
C_d	Damping matrix	$[N \cdot s / m]$
d	Distance	$[mm]$
d_0	Initial diameter	$[mm]$
D	Diagonal matrix defining effective strain rate	$[-]$
D^e	Elastic stress-strain matrix	$[MPa]$
D^{ep}	Combined elastic-elastoplastic stress-strain matrix	$[MPa]$
D^p	Elastoplastic stress-strain matrix	$[MPa]$
E	Elasticity modulus (unit in Eq. (2.13): $[kPa]$)	$[MPa]$
f	Friction factor (Chapter 4 and Chapter 7)	$[-]$
f	Generic function (Chapter 2)	$[-]$
f	Yield function (Chapter 3)	$[-]$
f^a	Anisotropic yield criterion	$[MPa^2]$
f_i	Nodal field variable	$[/]$
f_l	Factor defining lower limit of yielding range	$[-]$
f_u	Factor defining upper limit of yielding range	$[-]$
F	Anisotropic parameter	$[1 / MPa^2]$
F_i	Surface traction	$[MPa]$
\mathbf{f}	Load vector	$[N]$
\mathbf{f}_c	Load vector due to contact	$[N]$
\mathbf{F}	Matrix form of surface tractions	$[MPa]$
g	Weighting function for candidate nodes	$[-]$
g_n^c	Normal gap velocity in contact pair c	$[mm / s]$
g_t^c	Tangential gap velocity in contact pair c	$[mm / s]$
G	Anisotropic parameter (Section 3.4)	$[1 / MPa^2]$
G	Shear modulus	$[MPa]$
h	Heat transfer coefficient	$[W / (K \cdot mm^2)]$
h_0	Initial height	$[mm]$
H	Anisotropic parameter	$[1 / MPa^2]$
H'	Slope of stress-strain curve	$[MPa]$
I	Functional for least square minimization	$[/]$

i	Index vector	[–]
I	Identity matrix	[–]
<i>j</i>	Column number (Chapter 6)	[–]
<i>J</i>	Current density	[A/mm ²]
<i>k</i>	Node number (Chapter 5)	[–]
<i>k</i>	Shear flow stress	[MPa]
<i>k</i>	Thermal conductivity (Chapter 3)	[W/(K · mm)]
<i>K</i>	Volumetric penalty factor	[MPa · s]
K	Stiffness matrix (unit in Section 2.2: [N/m])	[N · s / m]
K_c	Conduction matrix	[W / K]
K_c	Stiffness matrix due to contact (Chapter 4)	[N · s / m]
K_e	Electrical conductance matrix	[mm]
<i>l</i>	Distance to edge segment	[mm]
<i>l_c</i>	Thickness of interface layer	[mm]
<i>L</i>	Anisotropic parameter (Section 3.4)	[1/MPa ²]
<i>L</i>	Latent heat	[J/kg]
<i>L_e</i>	Representative element size	[mm]
<i>L_k</i>	Distance along edge	[mm]
L	Matrix defining strain rate-velocity matrix	[mm ⁻¹]
<i>m</i>	Friction factor	[–]
<i>M</i>	Anisotropic parameter (Section 3.4)	[1/MPa ²]
<i>M</i>	Number of elements	[–]
<i>M_{ijkl}</i>	Tensor relating strain rate to deviatoric stress in anisotropy	[MPa ²]
M	Mass matrix	[kg]
<i>n</i>	Iteration number	[–]
<i>n</i>	Number of rows and columns (Chapter 6)	[–]
<i>n</i>	Strain-hardening exponent (Chapter 7)	[–]
<i>n_i</i>	Surface normal unit vector	[–]
<i>N</i>	Anisotropic parameter (Section 3.4)	[1/MPa ²]
<i>N</i>	Number of threads (Chapter 6)	[–]
<i>N_c</i>	Number of contact pairs	[–]
<i>N_i</i>	Element shape function	[–]
<i>N_p</i>	Contacting node	[–]

N_{rot}	Number of rotation steps	$[-]$
\mathbf{n}	Normal vector	$[-]$
\mathbf{N}	Matrix of element shape functions	$[-]$
p	Pressure	$[MPa]$
P	Contact penalty factor	$[-]$
P	Edge node (Chapter 5)	$[-]$
P	Parallel fraction (Chapter 6)	$[-]$
P_{ijkl}	Tensor defining anisotropic yield criterion	$[1 / MPa^2]$
\mathbf{P}	Matrix relating velocity and effective strain rate	$[mm^{-2}]$
q	Energy per unit volume	$[J / mm^3]$
\dot{q}	Energy rate per unit volume	$[W / mm^3]$
\mathbf{q}	Heat source vector	$[W]$
r	Radius	$[mm]$
R	Resistance	$[\Omega]$
R_a	Roughness	$[mm]$
R_c	Contact resistance	$[\Omega]$
R_e	Elastic fraction of load increment	$[-]$
R_p	Elastoplastic fraction of load increment	$[-]$
\mathbf{R}	Residual force vector	$[N]$
\mathbf{R}	Rotation matrix (Section 3.4.1)	$[-]$
S	Setdown (Chapter 9)	$[-]$
S	Surface	$[mm^2]$
S_k	Parametric distance along edge	$[-]$
\mathbf{s}	Skyline vector	$[/]$
t	Temporary center node (Chapter 4)	$[-]$
t	Time	$[s]$
T	Temperature	$[^{\circ}C]$
T_d^c	Temperature difference in contact pair c	$[K]$
T_f	Temperature of surroundings	$[^{\circ}C]$
T_{liq}	Liquidus temperature	$[^{\circ}C]$
T_N	Solution time on N threads	$[s]$
T_s	Surface temperature	$[^{\circ}C]$
T_{sol}	Solidus temperature	$[^{\circ}C]$

T_{tool}	Tool surface temperature	$[^{\circ}C]$
\dot{T}	Temperature rate	$[K/s]$
\mathbf{t}	Tangential vector	$[-]$
\mathbf{T}	Vector of nodal temperatures	$[^{\circ}C]$
u_i	Velocity (or displacement $[mm]$ in Chapter 2)	$[mm/s]$
\mathbf{u}	Vector of velocities (or displacements $[mm]$ in Chapter 2)	$[mm/s]$
ν	Valence	$[-]$
v_r	Relative sliding velocity	$[mm/s]$
V	Vertex node (Chapter 5)	$[-]$
V	Volume	$[mm^3]$
\mathbf{v}	Vector of nodal velocities	$[mm/s]$
x_i	Coordinate	$[mm]$
\mathbf{x}	Vector of coordinates	$[mm]$
Y	Yield stress	$[MPa]$

1. Introduction

The basics of resistance welding and examples of industrial applications are presented in the following for a brief introduction to the process that is dealt with in the present work. The background and the motivation of the project are described subsequently to give an overview of the starting point of the project and the overall goals and vision of the project. An outline of the remaining chapters is given in the end of the introduction.

1.1 Resistance welding

The resistance welding process is widely used in various industries to join two or more metal parts. Heat is generated by Joule heating in the parts and in the faying surfaces in particular to facilitate welding induced by an electric current and an applied compressive force. The following steps are included in a weld schedule.

- *Squeeze time* where a force is applied to form initial contact between the workpieces and between the workpieces and the electrodes.
- *Weld time* where an electric current is passed through while keeping the applied force. The current can be one or more pulses applied as alternating current (AC), direct current (DC) / middle frequency direct current (MFDC) or through a capacitor discharge (CD).
- *Hold time* where the force is kept until the weld has gained sufficient strength due to cooling and solidification.

Each weld takes in the order of half a second, say ranging from a few hundred microseconds to one second depending on the specific weld, and is thus a fast and efficient joining technique that has entered in many assembly lines. During the short process time, several physical phenomena interact and result in a complex process. The three basic physical aspects of resistance welding are mechanics, electricity and heat transfer. The mechanical aspects cover the applied force building up an initial contact area that is dynamically changed during the welding process and cover overall deformation and the resulting stress field. Electricity is the basic mechanism for the heat development. The electric current field generates heat according to Joule heating, which can be stated as follows,

$$q = \int_{t_1}^{t_2} \rho(t) J(t)^2 dt \quad (1.1)$$

where the generated energy per unit volume q during the time between t_1 and t_2 is a result of the electrical resistivity ρ in each material point that is experiencing a current density J of the electrical field. The current flow and the actual current density are governed by electricity. The induced heat results in temperature changes that are distributed over the involved material by heat transfer. On top of the three basic physical phenomena come material changes in terms of work hardening, direct influence from temperature changes and microstructural changes. This makes resistance welding a highly dynamic process.

Concentration of the developed heat is central to the process. Two typical ways of concentrating the current, and thereby the heat development, are illustrated in Fig. 1.1 by the two most common resistance welding processes. Spot welding is shown in Fig. 1.1a, where conical electrodes concentrate the current through two or more sheets. Fig. 1.1b illustrates projection welding, where the current is concentrated in the weld zones by the geometry of the parts to be welded. Projection welding takes various forms and covers many geometries, and Fig. 1.1b is just one of them.

Along the direction of current flow, the heat concentration is largely dependent on the distribution of electrical resistance. This holds especially for spot welding where the electrical contact resistance in the faying surfaces is playing a major role. To start with the bulk materials, the electrodes are in the majority of cases produced in a copper alloy with lower electrical resistance than e.g. steel, which is the most traditional material welded by spot welding. This already gives rise to a larger heat generation in the steel sheets than in the electrodes, and it is combined with the higher heat conduction in the electrodes than in the sheet material such that the temperature increases more in the sheets. In interaction with this is the effect of increased electrical resistance across material interfaces. Due to restriction of the current and additional resistance stemming from the actual metal surfaces in terms of surface films, oils and oxides, the electrical contact resistance can be an order of magnitude larger than the bulk resistivities in the early stages of the welding. As a result, the major heat generation and melt initiation

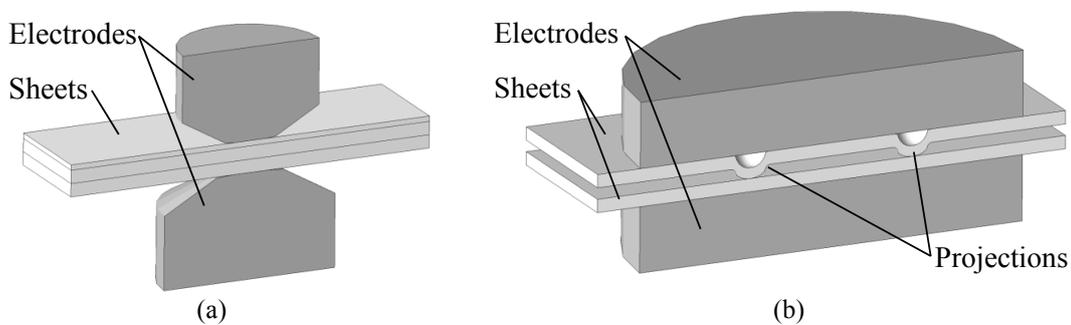


Fig. 1.1. Two types of resistance welding illustrated by (a) spot welding and (b) projection welding.

start in the faying surfaces between the sheets. Heat is also developed in the interfaces towards the electrodes, but because of the high thermal conductivity of the electrodes, the temperature remains moderate and the weld nugget forms across sheet metal interfaces to form a joint.

As regards projection welding, the above considerations for the electrical resistivity still apply; though with less direct influence because the current concentration due to the small contact area between the parts to be welded in many cases is determining the process. The initial heating caused by the current concentration can lead to good contact conditions due to elevated temperatures, softening of the asperities in real contact and squeeze out of initial surface contaminants. In this case the contact resistivity plays a lesser role.

The above presents the resistance welding process briefly. Later discussions will describe the process in more detail in relation to specific welding cases. Otherwise, for further general reading of the resistance welding process, reference is made to The Resistance Welding Manufacturers' Association [1] for a comprehensive description.

1.2 Applications of resistance welding

The resistance welding process is taking various forms and being applied in many industries. Using the words of Flax [2], *there is, in fact, scarcely an industry in which resistance welding has not entered in one form or another; indeed, in many instances economic production has only been attained, and maintained, by application of this welding technique.* A complete or systematic overview of applications is not provided in the following, but examples of the most typical applications are given for introduction purposes.

The most frequently used variations of the process, according to Schreiber [3], are the following:

- Spot welding.
- Seam welding.
- Projection welding.
- Flash and upset butt welding.

By number, the resistance spot welding in the automotive industry is the most applied welding process. It is widely used for joining two and three sheets. According to Zhu et al. [4], more than 200 sheet metal parts are welded, and 4000-7000 spot welds are carried out in the assembly of each car. In addition to the spot welds comes a number of projection welds. The spot welding process is also used for rail cars. As an example, Jaxa-Rozen [5] describes fabrication of passenger railcars of austenitic stainless steel. Here, spot welds of five sheets are carried out in assemblies having total thickness larger than 15mm , which is a fairly thick assembly compared to the assemblies in the automotive industry typically ranging from 1mm to 3.5mm .

The resistance welding process is used to join many different materials. In the automotive industry, steels are the primary material source, and because of the many types of steel this is providing welding engineers challenges in setting up weld schedules. The introduction of aluminum provides further challenges. Some illustrative examples are presented by Singh [6] showing that the resistance spot welding process meets competition from alternative joining processes. The so called all-aluminum Audi A8, 1993/94, was manufactured with the spot welding process being almost negligible. In the all-aluminum Audi A2, launched in 2000, no spot welds were used at all. Singh [6] also mentions the all-aluminum front-end of the BMW 5-series introduced in 2002 and the aluminum Jaguar S-series in 2003 as examples with limited use of spot welding. They were instead produced with substantial use of self-piercing and adhesive bonding. After these cases, the spot welding process was doubted for aluminum parts. However, the resistance welding engineers, including welding equipment producers, end-product manufacturers, and universities, have been innovative to solve the facing problems regarding welding of aluminum. According to Singh [6] the properties of the rivets did not follow the same developments, and thus spot welding is still relevant for aluminum parts.

The electronic industry makes use of resistance welding as well, but in this industry it is the projection welding that is of major interest. Based on the growth in the 1950s-1960s of the application of the process, Knowlson [7] describes the resistance welding in electronics. Lead wires, nickel-wires and gold wires were welded to each other, to connection pads and end-caps in resistors and small integrated circuits. Spot welding, cross-wire welding and projection welding are all used in electronics. Knowlson [7] specifically mentioned the benefit of packing components densely when joining by cross-wire welding. In production of electronic circuits, Fukumoto et al. [8] and their references mention resistance welding used for cross-wire welding of wires, welding of wires to substrate sheets and welding of foils. Fukumoto et al. [8] consider cross-wire welding of micro wires of nickel. Cross-wire welding of stainless steel was treated by Khan et al. [9], who mention the application of such welds to implantable medical devices, biosensors, stents, catheters, pacemakers, and surgical instruments.

1.3 Background and motivation

Modeling and simulation of resistance welding are tools to better understand the process and thereby a tool to solve new problems arising when welding new combinations of geometries and materials. Singh [6] points out that simulation cannot replace or substitute ingenuity or creativeness, but it can help in gaining understanding of the process, and hence reduce the amount of time spent during development. Volkswagen, as an example from industry, has experimented with an addition-material when welding high strength steels. They reached a larger joining area and a higher strength by introducing this material. In relation to this specific example, Graul [10] points out the val-

ue of being able to simulate the process. The number of iterations in such developments is kept minimal due to the simulations.

Such examples and statements motivate for modeling the resistance welding process, as it has been attempted over the past decades.

Some of the early numerical modeling of the resistance welding process is by Nied [11] and Cho and Cho [12]. Nied [11] is recognized as the first to apply finite element analysis in modeling of the process while Cho and Cho utilized the finite difference method. Because of the involved deformations, the finite element method is more suited and the work by Nied [11] gave the first early results using the commercial code ANSYS. He used solid elements for modeling the electrodes and workpieces, and he applied elements on the boundary for modeling the surface contact conditions. The surface elements were able to handle elastic mechanical contact (Herzian contact) by supporting compressive stresses but not tensile stresses. The electrical contact resistance and the thermal conduction of the contact layer were also applied to these elements. It was the first electro-thermo-mechanical coupling included in the analysis of resistance spot welding, and it was therefore for the first time possible to include the effect of the dynamically developing contact area. Good agreement with experiments was found and a motivation for further developments arose.

Another, later finite element study based on ANSYS was presented by Zhu et al. [4], who modeled projection welding of an automotive door hinge with two projections to a sheet metal. An electro-thermo-mechanical model was utilized with ability of modeling plasticity. Contact surface properties were modeled by applying a fictitious layer of elements with a finite thickness. Their simulations helped tool engineers, who were doing a large amount of measurements to find the distortion tendencies of the hinges after welding.

An electro-thermo-mechanical finite element program, SORPAS, has been developed at the Technical University of Denmark and later maintained and further developed by the spin-off company SWANTEC Software and Engineering ApS; see e.g. Zhang and Kristensen [13] and Zhang [14-15]. SORPAS is dedicated to simulation and optimization of resistance welding processes and has been widely sold in industry and academia worldwide with the majority of the users being related to the automotive industry focusing on spot welding. SORPAS has till now been available only in 2D, but the automotive industry and the electronic industry in particular have demanded the possibility of simulating resistance welding processes in 3D.

Developing a numerical tool based on the finite element method for simulating and analyzing complex welding cases is challenging and relevant for further understanding and improving the process. The motivation behind the current project is the development of a such computer program, which will open for a variety of cases that can be further analyzed.

The project is carried out in the environment spanned by the Technical University of Denmark with expertise and facilities related to the process, SWANTEC Software and Engineering ApS with knowledge, contacts in the field and with numerical expertise through the existing SORPAS 2D, and finally the Technical University of Lisbon with expertise in 3D numerical modeling of forming processes in terms of an in-house

computer program I-Form3 dedicated to thermo-mechanical metal forming processes; see e.g. Alves et al. [16]. Drawing on the experience and existence of SORPAS 2D and I-Form3, a new computer program SORPAS 3D is the overall goal with challenges spanning overall finite element modeling, contact modeling (see Song et al. [17-18] and Song [19], who developed the contact algorithms for SORPAS in 2D), meshing, remeshing and parallel computing from a numerical point of view. Process insight is a natural requirement before modeling and is a natural part of the project due to experimentation and verification and discussion when comparing simulations to real cases.

1.4 Outline of remaining chapters

The following Chapter 2 gives a brief overview of available finite element formulations, among which the flow formulation is chosen for modeling the resistance welding process. Chapter 3 presents in detail the flow formulation with implementation details and with overview of the thermal and electrical models that are all coupled in order to facilitate simulation of the process. Contact modeling, which is required to handle multiple objects, is presented in Chapter 4 including mechanical, electrical and thermal treatment. Hereafter follow two chapters on independent disciplines that are relevant in general finite element modeling and which accommodate the completeness of the presented finite element implementation as a complete standalone computer program. Chapter 5 deals with meshing and remeshing and Chapter 6 deals with parallel computing. Material characterization is described in Chapter 7 as an important area of simulation because accurate modeling of the materials is necessary to obtain useful results.

Simulations of pure mechanical contact are presented in Chapter 8 for verification of the contact algorithms before focusing on welding cases. Chapter 9 is devoted to present comparisons of simulations and experiments in terms of two of the most challenging resistance welding cases. The current implementation is hereby stretched to its limits such that they are clear before Chapter 10 dealing with a wide range of applications. Spot welding, projection welding and mechanical micro-joining by resistance heating exemplify Chapter 10, which is built upon a mixture of pure numerical studies and real industrial cases.

Finally, Chapter 11 presents the conclusions of the project and provides an outlook for the developed computer program SORPAS 3D.

References

- [1] Resistance Welding Manufacturers' Association (2003) Resistance welding manual. RWMA, Philadelphia.
- [2] Flax S (1971) Modern applications of resistance welding. *Welding and Metal Fabrication*:289-293.
- [3] Schreiber S (2000) State-of-the-art in resistance welding. Proceedings of the 1st International Seminar on: Advances in Resistance Welding, Copenhagen, Denmark, pp. 1-14.

- [4] Zhu W-F, Lin Z, Lai X-M, Luo A-H (2006) Numerical analysis of projection welding on auto-body sheet metal using a coupled finite element method. *International Journal of Advanced Manufacturing Technology* 28(1-2):45-52.
- [5] Jaxa-Rozen W (2008) Resistance welding in fabrication of passenger railcars. *Proceedings of the 5th International Seminar on: Advances in Resistance Welding*, Toronto, Canada, pp. 167-176.
- [6] Singh S (2004) Can simulation of the welding process help advance the state of the art in resistance welding? *Proceedings of the 3rd International Seminar on: Advances in Resistance Welding*, Berlin, Germany, pp. 5-11.
- [7] Knowlson PM (1967) Welding in electronics. *British Welding Journal* 14(7):398:400.
- [8] Fukumoto S, Tsubakino H, Zhou Y (2006) Heat input and deformation control in resistance micro-welding of fine nickel wires. *Welding International* 20(9):692-697.
- [9] Khan MI, Kim JM, Kuntz ML, Zhou Y (2009) Bonding mechanisms in resistance microwelding of 316 low-carbon vacuum melted stainless steel wires. *Metallurgical and Materials Transactions A* 40(4):910-919.
- [10] Graul M (2006) Development and optimization of resistance welding processes using process-simulation. *Proceedings of the 4th International Seminar on: Advances in Resistance Welding*, Wels, Austria, pp. 5-14.
- [11] Nied HA (1984) The finite element modeling of the resistance spot welding process. *Welding Research Supplement* 63(4):123-132.
- [12] Cho HS, Cho YJ (1989) A study of the thermal behavior in resistance spot welds. *Welding Journal* 68(6):236s-244s.
- [13] Zhang W, Kristensen L (1999) Finite element modeling of resistance spot and projection welding processes. *The 9th International Conference on Computer Technology in Welding*, Detroit, Michigan, pp. 15-23.
- [14] Zhang W (2003) Design and implementation of software for resistance welding process simulations. *Transactions: Journal of Materials and Manufacturing* 112(5):556-564.
- [15] Zhang W (2010) Weld planning with optimal parameters by computer simulations and optimizations. *Proceedings of the 6th International Seminar on: Advances in Resistance Welding*, Hamburg, Germany, pp. 119-127.
- [16] Alves ML, Rodrigues JMC, Martins PAF (2004) Three-dimensional modeling of forging processes by the finite element flow formulation. *Journal of Engineering Manufacture* 218:1695-1707.
- [17] Song Q, Zhang W, Bay N (2006) Contact modelling in resistance welding, part I: Algorithms and numerical verification. *Proceedings of the Institution of Mechanical Engineers, Part B (Journal of Engineering Manufacture)* 220(B5):599-606.
- [18] Song Q, Zhang W, Bay N (2006) Contact modelling in resistance welding, part II: Experimental validation. *Proceedings of the Institution of Mechanical Engineers, Part B (Journal of Engineering Manufacture)* 220(B5):607-613.
- [19] Song Q (2003) Testing and modeling of contact problems in resistance welding. PhD Thesis. Technical University of Denmark.

2. Finite Element Formulations

The governing equations for problems solved by the finite element method are typically formulated by partial differential equations in their original form. These are rewritten into a weak form, such that domain integration can be utilized to satisfy the governing equations in an average sense. A functional Π is set up for the system, typically describing the energy or energy rate and implying that the solution can be found by minimization. For a generic functional, this is written as

$$\frac{\partial \Pi}{\partial u} = \frac{\partial}{\partial u} \left(\int_V f(x_i, u_i) dV \right) = \frac{\partial}{\partial u} \left(\sum_j f(x_i, u_i) \Delta V_j \right) = 0 \quad (2.1)$$

where the functional is a function of the coordinates x_i and the primary variable u_i being e.g. displacements or velocities for mechanical problems depending on the formulation. The domain integration is approximated by a summation over a finite number of elements discretizing the domain. Fig. 2.1 illustrates a three-dimensional domain discretized by hexahedral elements with eight nodes. The variables are defined and solved in the nodal points, and evaluation of variables in the domain is performed by interpolation in each element. Shared nodes give rise to an assembly of elements into a global system of equations of the form

$$\mathbf{K}\mathbf{u} = \mathbf{f} \quad (2.2)$$

where \mathbf{K} is the stiffness matrix, \mathbf{u} is the primary variable and \mathbf{f} is the applied load, e.g. stemming from applied tractions F on a surface S_F in Fig. 2.1. The system of equations (2.2) is furthermore subject to essential boundary conditions, e.g. prescribed displacements or velocities u along a surface S_U .

The basic aspects of available finite element formulations in terms of modeling and computation are briefly reviewed in this chapter. This will support the choice of formulation to be detailed and applied in the remaining chapters, where an electrothermo-mechanical finite element formulation is presented together with a range of aspects to complete a computer program capable of modeling manufacturing processes such as metal forming and resistance welding. This chapter is focused on the mechanical formulations because they represent major differences and because the mechanical model plays a central role in the overall modeling strategy. From a process point of

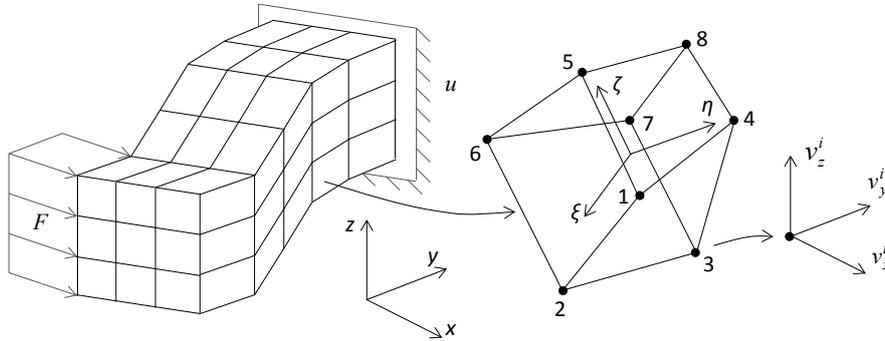


Fig. 2.1. Illustration of three-dimensional finite element model composed of isoparametric, hexahedral elements with eight nodes. Each node has three degrees of freedom for representation of vector fields and one degree of freedom for representation of scalar fields.

view the mechanical model is responsible for material flow, contact and stress distribution, and from a computational point of view is responsible for the largest amount of CPU time. In addition, the overall structure of the presented computer program is built upon the mechanical formulation with the remaining thermal and electrical modules integrated.

One fundamental difference between the finite element formulations is the governing equilibrium equation, being either quasi-static or dynamic in the modeling of manufacturing processes. Another fundamental choice to cover is the material model suited for describing the materials under consideration, bearing in mind the process to simulate and thereby the expected range of deformation and deformation rate. The available constitutive models to utilize in the material description are rigid-plastic/viscoplastic and elasto-plastic/viscoplastic.

Table 2.1, after Tekkaya and Martins [1], provides an overview of the quasi-static formulations and the dynamic formulation. The quasi-static formulations are represented by the flow formulation and the solid formulation, distinguishable by the underlying constitutive equations. The following two sections are devoted to give a brief overview of the quasi-static and dynamic formulations including their advantages and disadvantages.

Presentation of the quasi-static and dynamic formulations follows the general outline given by Tekkaya and Martins [1] and additional information can be found in major reference books by Zienkiewicz and Taylor [2], Banabic et al. [3], Wagoner and Chenot [4] and Dunne and Petrinic [5].

Table 2.1. Overview of finite element formulations and commercial computer programs applied in the metal forming industry.

	Quasi-static formulations		Dynamic formulation
	Flow formulation	Solid formulation	
Equilibrium equation:	Quasi-static	Quasi-static	Dynamic
Constitutive equations:	Rigid-plastic/ visco-plastic	Elasto-plastic/ visco-plastic	Elasto-plastic/ visco-plastic
Main structure:	Stiffness matrix and force vector	Stiffness matrix and force vector	Mass and damping matrices and internal and external force vectors
Solution scheme ^a :	Implicit	Implicit	Explicit
Size of incremental step:	Large	Medium to large	Very small
CPU time per incremental step:	Medium	Medium to long	Very short
Time integration scheme ^b :	Explicit	Implicit	Explicit
Accuracy of the results (stress and strain distributions):	Medium to high	High	Medium to low
Springback and residual stresses:	No (although the formulation can be modified to include elastic effects)	Yes	Yes/no
Commercial FEM computer programs related to metal forming	FORGE ^c , DEFORM ^c , QFORM, eesy-2-form	Abaqus (implicit), Simufact.forming, AutoForm, Marc	Abaqus (explicit), DYNA3D, PAM-STAMP

^a *Explicit / implicit if the residual force is not / is minimized at each incremental step.*
^b *Explicit / implicit if the algorithm does not / does need the values of the next time step to compute the solution.*
^c *Elasto-plastic options available.*

2.1 Quasi-static formulations

The quasi-static formulations are governed by the static equilibrium equation, which in the absence of body forces takes the following form,

$$\sigma_{ij,j} = 0 \quad (2.3)$$

where $\sigma_{ij,j}$ denotes the partial derivatives of the Cauchy stress tensor with respect to the Cartesian coordinates x_j . This equation expresses the equilibrium in the current configuration, i.e. in the mesh following the deformation.

By employing the Galerkin method, it is possible to write an integral form of equation (2.3) that fulfills the equilibrium in an average sense over the entire domain instead of satisfying the equilibrium point-wise. This formulation allows domain integra-

tion to substitute the more tedious solution of the original differential equations. The integral over domain volume V is

$$\int_V \sigma_{ij,j} \delta u_i dV = 0 \quad (2.4)$$

with δu_i being an arbitrary variation in the primary unknown u_i , which is either displacement or velocity depending on the implementation. Displacement is the primary unknown in rate independent formulations and velocity is the primary unknown in rate dependent formulations.

Applying integration by parts in equation (2.4), followed by the divergence theorem and taking into account the natural and essential boundary conditions, it is possible to rewrite equation (2.4) as follows,

$$\int_V \sigma_{ij} (\delta u_i)_{,j} dV - \int_S F_i \delta u_i dS = 0 \quad (2.5)$$

where $F_i = \sigma_{ij} n_j$ denotes the tractions with direction of the unit normal vector n_j applied on the boundary surface S . Equation (2.5) is the ‘‘weak variational form’’ of equation (2.3) because the static governing equilibrium equations are now only satisfied under weaker continuity requirements.

The above listed equations together with appropriate constitutive equations enable quasi-static finite element formulations to be defined by means of the following matrix set of non-linear equations,

$$\mathbf{K}^t \mathbf{u}^t = \mathbf{f}^t \quad (2.6)$$

which express the equilibrium condition at the instant of time t through the stiffness matrix \mathbf{K} , the generalized force vector \mathbf{f} resulting from the loads, pressure and friction stresses applied on the boundary. The equation system is non-linear due to the stiffness matrix’s dependency of the primary unknown \mathbf{u} to geometry and material properties.

The quasi-static finite element formulations utilized in the analysis of metal forming and resistance welding processes are commonly implemented in conjunction with implicit solution schemes. The main advantage of implicit schemes over alternative solutions based on explicit procedures is that equilibrium is checked at each increment of time by means of iterative procedures to minimize the residual force vector $\mathbf{R}(\mathbf{u})$, which is computed as follows in iteration number n ,

$$\mathbf{R}_n^t = \mathbf{K}_{n-1}^t \mathbf{u}_n^t - \mathbf{f}^t \quad (2.7)$$

The non-linear set of equations (2.6), derived from the quasi-static implicit formulations, can be solved by different numerical techniques such as the direct iteration (also known as “successive replacement”) and the Newton-Raphson methods. In the direct iteration method, the stiffness matrix is evaluated for the displacements of the previous iteration in order to reduce equation (2.6) to a linear set of equations. The method is iterative and converges linearly and unconditionally towards the solution during the earlier stages of the iteration procedure but becomes slow as the solution is approached. The standard Newton-Raphson method is an alternative iterative method based on a linear expansion of the residual $\mathbf{R}(\mathbf{u})$ near the velocity estimate at the previous iteration,

$$\mathbf{R}_n^t = \mathbf{R}_{n-1}^t + \left[\frac{\partial \mathbf{R}}{\partial \mathbf{u}} \right]_{n-1}^t \Delta \mathbf{u}_n^t = 0 \quad (2.8a)$$

$$\mathbf{u}_n^t = \mathbf{u}_{n-1}^t + \Delta \mathbf{u}_n^t \quad (2.8b)$$

This procedure is only conditionally convergent, but converges quadratically in the vicinity of the exact solution. The iterative procedures are designed in order to minimize the residual force vector $\mathbf{R}(\mathbf{u})$ to within a specified tolerance. Control and assessment is performed by means of appropriate convergence criteria.

The main advantage of the quasi-static implicit finite element formulations is that equilibrium conditions are checked at each increment of time in order to minimize the residual force vector $\mathbf{R}(\mathbf{u})$ to within a specified tolerance.

The main drawbacks in the quasi-static implicit finite element formulations are summarized as follows:

- Solution of linear systems of equations is required during each iteration;
- High computation times and high memory requirements;
- Computation time depends quadratically on the number of degrees of freedom if a direct solver is utilized, and with the Newton-Raphson method the solution is only conditionally convergent;
- The stiffness matrix is often ill-conditioned, which can turn the solution procedure unstable and deteriorate the performance of iterative solvers;
- Difficulties in dealing with complex non-linear contact and tribological boundary conditions are experienced, and that often leads to convergence problems.

2.2 Dynamic formulation

The dynamic finite element formulation is based on the dynamic equilibrium equation in the current configuration, here written in the absence of body forces with the inertia term expressed through the mass density ρ_m and the acceleration \ddot{u}_i ,

$$\sigma_{ij,j} - \rho_m \ddot{u}_i = 0 \quad (2.9)$$

Applying a mathematical procedure similar to that described in the previous section results in the following weak variational form of equation (2.9),

$$\int_V \rho_m \ddot{u}_i \delta u_i dV + \int_V \sigma_{ij} (\delta u_i)_{,j} dV - \int_S F_i \delta u_i dS = 0 \quad (2.10)$$

The above equation enables dynamic finite element formulations to be represented by the following matrix set of non-linear equations,

$$\mathbf{M}^t \ddot{\mathbf{u}}^t + \mathbf{f}_{\text{int}}^t = \mathbf{f}^t \quad (2.11)$$

which express the dynamic equilibrium condition at the instant of time t . The symbol \mathbf{M} denotes the mass matrix, $\mathbf{f}_{\text{int}} = \mathbf{K}\mathbf{u}$ is the vector of internal forces resulting from the stiffness, and \mathbf{f} is the generalized force vector.

The non-linear set of equations (2.11), derived from the dynamic formulation, is commonly solved by means of an explicit central difference time integration scheme,

$$\mathbf{M}^t \left(\frac{\dot{\mathbf{u}}^{t+1/2} - \dot{\mathbf{u}}^{t-1/2}}{\Delta t^{t+1/2}} \right) + \mathbf{f}_{\text{int}}^t = \mathbf{f}^t \quad (2.12a)$$

$$\dot{\mathbf{u}}^{t+1/2} = (\mathbf{M}^t)^{-1} (\mathbf{f}^t - \mathbf{f}_{\text{int}}^t) \Delta t^{t+1/2} + \dot{\mathbf{u}}^{t-1/2} \quad (2.12b)$$

$$\mathbf{u}^{t+1} = \mathbf{u}^t + \dot{\mathbf{u}}^{t+1/2} \Delta t^{t+1} \quad (2.12c)$$

If the mass matrix \mathbf{M} in equations (2.12a-b) is diagonalized (or lumped) its inversion is trivial, and the system of differential equations decouples. Its overall solution can then be performed independently and very fast for each degree of freedom. Further reductions of the computation time per increment of time stem from utilization of reduced integration schemes that are often applied even to the deviatoric parts of the stiffness matrix, and finally numerical actions related to mass scaling and load factoring contribute. Load factoring is described ahead.

Additional computational advantages result from the fact that dynamic explicit schemes, unlike quasi-static implicit schemes, do not check equilibrium requirements at the end of each increment of time. The analogy between the dynamic equilibrium equation (2.9) and the ideal mass-spring vibrating system allows concluding that explicit central difference time integration schemes (frequently referred as explicit integration schemes) are conditionally stable whenever the size of the increment of time Δt satisfies

$$\Delta t \leq \frac{L_e}{\sqrt{E/\rho_m}} = \frac{L_e}{c_e} \quad (2.13)$$

where L_e is the typical size of the finite elements discretizing the domain, E is the elasticity modulus and c_e is the velocity of propagation of a longitudinal wave in the material. In case of metal forming applications, the stability condition (2.13) requires the utilization of very small increments of time Δt , say microseconds, and millions of increments to finish a simulation because industrial metal forming processes usually take several seconds to be accomplished. This is the reason why computer programs often make use of the following numerical actions in order to increase the increment of time Δt and, consequently, reducing the overall computation time:

- Diagonalization of the mass matrix;
- Mass scaling - by increasing the density of the material and thus artificially reducing the speed c_e of the longitudinal wave;
- Load factoring - by changing the rate of loading through an artificial increase in the velocity of the tooling as compared to the real forming velocity;
- Reduced integration of the deviatoric part of the stiffness matrix, which is usually fully integrated.

The above-mentioned numerical actions can artificially add undesirable inertia effects, and it is therefore necessary to include a damping term $\mathbf{C}_d^t \dot{\mathbf{u}}^t$ in (2.11),

$$\mathbf{M}^t \ddot{\mathbf{u}}^t + \mathbf{C}_d^t \dot{\mathbf{u}}^t + \mathbf{f}_{\text{int}}^t = \mathbf{f}^t \quad (2.14)$$

The damping term $\mathbf{C}_d^t \dot{\mathbf{u}}^t$ is not only necessary because of the above-mentioned numerical actions to reduce the computation time but also to ensure fast convergence of the solution towards the static solution describing the actual process.

This turns dynamic explicit formulations into close resemblance with damped mass-spring vibrating systems and justifies the reason why these formulations lose efficiency whenever the material is strain-rate sensitive or thermo-mechanical phenomena need to be taken into consideration.

The main advantages of the dynamic explicit formulations are:

- Computer programs are robust and do not present convergence problems;
- The computation time depends linearly on the number of degrees of freedom while in alternative quasi-static implicit schemes the dependency is more than linear (in case of iterative solvers) and up to quadratic (in case of direct solvers).

The main drawbacks of the dynamic explicit formulation can be summarized as follows:

- Utilization of very small time increments;
- Equilibrium after each increment of time is not checked;
- Assignment of the system damping is rather arbitrary;
- The formulation needs experienced users for adequately designing the mesh and choosing the scaling parameters for mass, velocity and damping. Otherwise it may lead to inaccurate solutions for the deformation, prediction of forming defects and distribution of the major field variables within the workpiece;
- Springback calculations are very time consuming and may lead to errors. This specific problem is frequently overtaken by combining dynamic explicit with quasi-static implicit analysis.

The last two drawbacks apply if the dynamic explicit formulations are used in the “high-speed-mode”.

References

- [1] Tekkaya AE, Martins PAF (2009) Accuracy, reliability and validity of finite element analysis in metal forming: a user’s perspective. *Engineering Computations* 26:1026-1055.
- [2] Zienkiewicz OC, Taylor RL (2000) *The finite element method*, Butterworth-Heinemann, Oxford.
- [3] Banabic D, Bunge H-J, Pöhlandt K, Tekkaya AE (2000) *Formability of metallic materials*. Springer Verlag, Berlin.
- [4] Wagoner RH, Chenot JL (2001) *Metal forming analysis*. Cambridge University Press, Cambridge.
- [5] Dunne F, Petrinic N (2006) *Introduction to computational plasticity*. Oxford University Press, Oxford.

3. Coupled Finite Element Flow Formulation

This chapter presents a coupled finite element approach for thermo-mechanical modeling of metal forming and for electro-thermo-mechanical modeling of resistance welding. The parallel to metal forming is drawn because of the close relation between SORPAS 3D and IForm3 and because the mechanical part of resistance welding is of severe influence to the numerical modeling.

The finite element approach is based on the flow formulation which was described in Chapter 2 as one of the implicit quasi-static formulations. Direct comparison of the performance achieved with the implicit quasi-static formulations based on flow and solid approaches (refer to Table 2.1) are provided by Boer et al. [1] and Kobayashi et al. [2], who emphasize the advantages of the flow approach in modeling the mechanical response (plastic flow) of materials undergoing large deformations.

3.1 State-of-the-art

Taking a general view to the bibliographic retrieval by Brännberg and Mackerle [3] and Mackerle [4-5] it appears that the finite element flow formulation is one of the most widespread numerical methodologies for the analysis of metal forming processes.

In the flow formulation, the material is treated in a similar way to an incompressible fluid. Rigid-plastic/viscoplastic constitutive laws are utilized and the elastic response is neglected, simplifying the problem and offering additional computational advantages. The computer programs based on the flow formulation can successfully take into account the non-linearities in the geometry and material properties as well as the contact changes typical of metal forming and resistance welding processes to produce accurate predictions of plastic flow, temperature, current density and microstructure.

In order to calculate temperatures and its resulting effects, the flow formulation is coupled with heat transfer analysis to achieve complete thermo-mechanical modeling. In resistance welding this coupling is further extended to include electrical analysis with special treatment of contact interfaces and to account for Joule heating. The extended model is electro-thermo-mechanically coupled and enables utilization in a wide range of manufacturing applications by industry, research and education institutions with the aim of:

- Developing new products and processes in shorter time;

- Optimizing existing products and processes by cost and quality;
- Increasing process understanding and strengthening technological know-how;
- Performing more efficient experimentation by providing starting parameters and support to the analyses.

Modeling and simulation of manufacturing processes are tools to better understand and thereby solve new problems arising when forming or joining new materials and geometries. The interaction with industry has been the motivation for applying and continuously developing the finite element flow formulation for manufacturing applications over the past decades. A brief overview of the previous research in the field is given in what follows with the aim of providing a timeline of the major contributions and identifying the current state-of-the-art.

The finite element flow formulation was originally developed by Lee and Kobayashi [6], Cornfield and Johnson [7] and Zienkiewicz and Godbole [8] during the 1970s with the aim of simulating metal forming processes. During the 1980s, the flow formulation was primarily set up for modeling two-dimensional bulk forming processes and such efforts gave rise to the development of a first generation of commercial software with applicability limited to plane strain and axisymmetric conditions. Even so, authors such as Altan and Knoerr [9] were able to report case studies in which the two-dimensional constraint was ingeniously stretched out in order to obtain useful information regarding three-dimensional metal forming applications.

In order to extend applicability of the flow formulation to modeling conditions involving more than the mechanical behavior alone, a thermal model was introduced to simulate thermo-mechanical manufacturing processes. The first attempt to handle a coupled thermo-mechanical metal forming process was made by Zienkiewicz et al. [10] who used a finite element iterative procedure to solve the material flow for a given distribution of temperature, in conjunction with the heat transfer phenomenon, during plane strain extrusion. Later, Zienkiewicz et al. [11-12] modified the procedure to allow the temperature distribution within the workpiece to be obtained simultaneously with the solution of the velocity field. The modification, commonly known as “direct coupled thermo-mechanical” was applied to solve steady-state extrusion and rolling. The heat exchange with the tools was either neglected, as in the case of the extrusion problem, or taken into account by imposing a constant temperature on the tools, as in the case of steady-state rolling.

Direct coupled thermo-mechanical finite element algorithms were further developed by Rebelo and Kobayashi [13-14] to allow the numerical simulation of non-steady-state metal forming processes. The technique was applied to solid cylinder and ring compression testing.

As regards resistance welding, early contributions, being analytical or numerical, were focused on the temperature field under a given voltage potential. Nied [15] was the first to present electro-thermo-mechanical modeling of spot welding by finite elements using the commercial program ANSYS. The study was performed in two dimensions with assumed Hertzian elastic contact. Contact conditions are crucial for the numerical simulation of resistance welding due to dynamically developing contact area

and Nied [15] addressed this problem by means of surface elements that were capable of supporting compressive stresses, but not tensile stresses. Relative sliding was allowed assuming frictionless contact and electrical and thermal properties were included in the aforementioned surface elements.

The work of Nied [15] was the first numerical simulation of resistance welding being so complete. Subsequent published work in the field was also based on commercial finite element computer programs for general purpose modeling, e.g. Zhu et al. [16] modeled projection welding of an automotive door hinge with two projections to a sheet metal by means of a two-dimensional analysis based on ANSYS.

Newer developments in computers and reduction in the associated computational costs are presently extending the availability and effectiveness of finite element software to simulate three-dimensional manufacturing processes. As a consequence, complex processes are now being simulated precisely without the need to take advantage of possible geometrical and material flow simplifications. A detailed survey of the state-of-the-art regarding numerical simulation of metal forming processes is given by Brännberg and Mackerle [3] and Mackerle [4-5].

In resistance welding state-of-the-art is the prediction of weld parameters in spot welding with the only input being the geometries of sheets and electrodes as well as the desired weld nugget size (refer, for example, to the presentation of weld planning in SORPAS by Zhang [17]). SORPAS is a commercial finite element program dedicated to simulation and optimization of resistance projection and spot welding. The program is based on the finite element flow formulation and has been primarily developed for axisymmetric and plane-strain industrial applications. Recent developments within the present project have extended the capabilities of SORPAS to simulate complex three-dimensional resistance welding applications.

This chapter is aimed at describing the fundamentals and numerical implementation of the thermo-mechanical and electro-thermo-mechanical coupled approaches that are available in academic (e.g. I-Form [18]) and commercial (e.g. SORPAS [17]) computer programs that are based on the finite element flow formulation.

3.2 Theoretical background

3.2.1 Plastic flow

The flow formulation is based on the quasi-static equilibrium equations, which in the absence of body forces and after some mathematical treatment that takes into consideration the natural and essential boundary conditions, can be written as (refer to Equation (2.5) in Chapter 2)

$$\int_V \sigma_{ij} \delta \varepsilon_{ij} dV - \int_S F_i \delta u_i dS = 0 \quad (3.1)$$

where V is the domain volume, S is the boundary surface where tractions $F_i = \sigma_{ij}n_j$ are applied and $\dot{\varepsilon}_{ij}$ are the components of the strain rate tensor,

$$\dot{\varepsilon}_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i}) \quad (3.2)$$

In the flow formulation, velocities u_i are the primary unknown instead of displacements, there is no strain tensor and the stress σ_{ij} is directly related to the strain rate by means of rigid-plastic/viscoplastic constitutive equations.

In case of using the von Mises yield criterion, also called the “distortion energy criterion”,

$$f(\sigma_{ij}) = \frac{1}{2}\sigma'_{ij}\sigma'_{ij} \quad (3.3)$$

where f is the yield function and σ'_{ij} is the deviatoric stress tensor, the constitutive equations (also known as the “Levy-Mises equations”) are written as

$$\dot{\varepsilon}_{ij} = \sigma'_{ij}\dot{\lambda} \quad (3.4)$$

The proportionality factor $\dot{\lambda}$ in the above equation is given by

$$\dot{\lambda} = \frac{3}{2} \frac{\dot{\bar{\varepsilon}}}{\bar{\sigma}} \quad (3.5)$$

with effective strain rate $\dot{\bar{\varepsilon}}$ and effective stress $\bar{\sigma}$ obtained from

$$\dot{\bar{\varepsilon}} = \sqrt{\frac{2}{3}} \left\{ \dot{\varepsilon}_{ij}\dot{\varepsilon}_{ij} \right\}^{\frac{1}{2}} \quad (3.6)$$

$$\bar{\sigma} = \sqrt{\frac{3}{2}} \left\{ \sigma'_{ij}\sigma'_{ij} \right\}^{\frac{1}{2}} \quad (3.7)$$

The variational principle associated with (3.1) requires that among admissible velocities u_i , satisfying the conditions of compatibility and incompressibility as well as the velocity boundary conditions, the actual solution gives the following functional a stationary value (minimum of the total energy rate),

$$\Pi = \int_V \bar{\sigma}\dot{\bar{\varepsilon}} dV - \int_S F_i u_i dS \quad (3.8)$$

where $\overline{\sigma \dot{\varepsilon}} = \sigma_{ij} \dot{\varepsilon}_{ij}$ according to (3.6) and (3.7). Equation (3.1) corresponds to a zero first order variation of the total energy rate of the system (3.8) and is accordingly re-written as follows,

$$\delta \Pi = \int_V \overline{\sigma \dot{\varepsilon}} dV - \int_S F_i \delta u_i dS = 0 \quad (3.9)$$

This is a weak form of the quasi-static equilibrium condition (2.3) because it lowers the continuity requirements on the stress field and allows solving the equilibrium condition by domain integration instead of the more tedious direct solving of differential equations.

In order to guarantee that the flow formulation is capable of providing a geometrically self-consistent velocity field that ensures the incompressibility condition it is necessary to ensure a zero first order variation of the functional Π (3.8), subject to a general constraint, $\dot{\varepsilon}_{kk} = 0$ over the entire domain. This can be done in several different ways, where the two most widespread techniques are based on the utilization of Lagrange multipliers (treating incompressibility as a mixed velocity-pressure approach) or penalties.

The utilization of a Lagrange multiplier λ_L , corresponding to the mean stress σ_m , modifies (3.9) to the following form,

$$\delta \Pi = \int_V \overline{\sigma \dot{\varepsilon}} dV + \int_V \lambda_L \delta \dot{\varepsilon}_{jj} dV + \int_V \delta \lambda_L \dot{\varepsilon}_{jj} dV - \int_S F_i \delta u_i dS = 0 \quad (3.10)$$

whereas the utilization of a penalty K , which is a large positive number related to the mean stress through $K \dot{\varepsilon}_{kk} = 2\sigma_m$, modifies (3.9) to the following form,

$$\delta \Pi = \int_V \overline{\sigma \dot{\varepsilon}} dV + K \int_V \dot{\varepsilon}_{ii} \delta \dot{\varepsilon}_{jj} dV - \int_S F_i \delta u_i dS = 0 \quad (3.11)$$

The advantage of the Lagrange multipliers is the exact solution, but on the expense of prolonged computation time due to additional unknowns in form of the mean stress $\sigma_m = \sigma_{kk}/3$ pressure terms.

The penalty approach does not introduce additional unknowns but suffers from a dilemma in the selection of the value of the penalty factor K . It has to be as large as possible to enforce incompressibility but it cannot be chosen too large because the system of equations becomes ill-conditioned with increasing penalty factor and leads to locking (trivial solution) whenever the penalty constraint takes a dominant role. The penalty based approach (also named as the ‘‘irreducible finite element flow formulation’’) is applied hereafter, such that the variational equation (3.11) is utilized.

3.2.2 Heat transfer

The purpose of simulating heat transfer and heat generation is to model the effects of the temperature increase due to plastic work, to heat generated by electrical Joule heating and to temperature variation due to exchange of heat with the tools and the surrounding environment. In an arbitrary volume, the energy rate balance requires

$$\dot{q}_{in} - \dot{q}_{out} + \dot{q}_{generate} = \dot{q}_{store} \quad (3.12)$$

where \dot{q}_{in} and \dot{q}_{out} are the energy rates per unit volume into the volume and out of the volume, respectively. The heat rate per unit volume due to generation inside the volume is $\dot{q}_{generate}$, and \dot{q}_{store} is the rate of stored energy per unit volume giving rise to a temperature gradient \dot{T} according to

$$\dot{q}_{store} = \rho_m c_m \dot{T} \quad (3.13)$$

where ρ_m is the mass density and c_m is the heat capacity.

In the temperature range of melting and solidification, i.e. $T_{sol} < T < T_{liq}$ with solidus temperature T_{sol} and liquidus temperature T_{liq} , an effective heat capacity is defined to include an approximation of the latent heat L as follows [19],

$$\tilde{c}_m = c_m + \frac{L}{T_{liq} - T_{sol}} \quad (3.14)$$

Applying Fourier's law for heat conduction, $\dot{q} = -kT_{,i}$ with thermal conductivity k , and assuming the control volume to be infinitesimal, the transient heat diffusion equation can be obtained from (3.12) as

$$\left(kT_{,i}\right)_{,i} + \dot{q}_{generate} = \rho_m c_m \dot{T} \quad (3.15)$$

The heat generation has several contributions. In the material volume, heat generation exists due to dissipated energy by the plastic work and the electrical heat source due to Joule heating. On the boundary surface, the contributions are convection and radiation to the surroundings and to the tools as well as friction generated heat in contact interfaces with relative sliding.

The contribution from the plastic work, is the fraction of the plastic deformation energy dissipated as heat,

$$\dot{q}_{plastic} = \beta \sigma_{ij} \dot{\epsilon}_{ij} = \beta \bar{\sigma} \dot{\bar{\epsilon}} \quad (3.16)$$

where $\beta \approx 0.85 - 0.95$.

The generated Joule heating due to electrical resistivity ρ and current density J is given by

$$\dot{q}_{electrical} = \rho J^2 \quad (3.17)$$

which will be analyzed more detailed in Section 3.2.3 dealing with the electrical field and resulting heat generation.

Newton's law for convection, applying to all free surfaces, is given by

$$\dot{q}_{convection} = h(T_s - T_f) \quad (3.18)$$

with heat transfer coefficient h , surface temperature T_s and temperature of the surroundings T_f .

Similarly Stefan-Boltzmann's law for radiation, applying to all free surfaces, is given by

$$\dot{q}_{radiation} = \varepsilon_{emis} \sigma_{SB} (T_s^4 - T_f^4) \quad (3.19)$$

with emission coefficient ε_{emis} , Stefan-Boltzmann coefficient σ_{SB} and temperatures in Kelvin.

At surfaces contacting the tools, convection follows

$$\dot{q}_{tool} = h_{lub} (T_s - T_{tool}) \quad (3.20)$$

where T_{tool} is the tool temperature and h_{lub} is the relevant convection coefficient, typically taken for an applied lubricant.

Finally, the heat generated by friction shear stresses τ_f in the contact interfaces with relative sliding v_r is given by

$$\dot{q}_{friction} = \tau_f |v_r| \quad (3.21)$$

The transient heat diffusion equation, (3.15), was firstly implemented by Rebelo and Kobayashi [13-14] in a finite element computer program for modeling thermo-mechanical metal forming processes, and subsequently implemented by Zhang et al. [19] for modeling the heat developed by Joule heating in resistance welding.

From this point of the presentation the thermal conductivity will be assumed constant within each integration domain, implying that $(kT_{,i})_{,i}$ simplifies to $kT_{,ii}$. Under these circumstances and applying the classical Galerkin method, the heat transfer equation (3.15) can be written as follows,

$$\int_V kT_{,i} \delta T_{,i} dV + \int_V \rho_m c_m \dot{T} \delta T dV - \int_V \dot{q}_{generate} \delta T dV - \int_S kT_{,n} dS = 0 \quad (3.22)$$

where $T_{,n}$ is the gradient of T along the outward normal to the surface S . The third term in (3.22) is the heat generated from plastic deformation (3.16) and Joule heating (3.17), and the fourth term is the heat flux on boundary surfaces. Along free surfaces S_{free} conduction and radiation follow (3.18) and (3.19), and along surfaces in contact with the tools S_{tool} , convection and friction generated heat follow (3.20) and (3.21). All these terms can be summarized as follows,

$$\begin{aligned} & \int_V kT_{,i} \delta T_{,i} dV + \int_V \rho_m c_m \dot{T} \delta T dV - \int_V (\dot{q}_{plastic} + \dot{q}_{electrical}) \delta T dV \\ & + \int_{S_{free}} (\dot{q}_{convection} + \dot{q}_{radiation}) dS + \int_{S_{tool}} (\dot{q}_{tool} - \dot{q}_{friction}) dS = 0 \end{aligned} \quad (3.23)$$

3.2.3 Electricity

The distribution of electric potential Φ utilized in the coupled electro-thermo-mechanical finite element implementation is based on Laplace's equation,

$$\Phi_{,ii} = 0 \quad (3.24)$$

Although this approach considers the distribution of the electric potential to be solely determined by geometry under steady conditions ($\dot{\Phi} = 0$) [20], it is generally considered a good approach because an electric field has a much faster reaction rate than a temperature field.

Along boundaries with power supply, the electric potential is the supplied potential, $\Phi = \Phi_0$, and along free surfaces electric potential is zero. Integrating Laplace's equation for an arbitrary variation in the electric potential Φ and applying the divergence theorem, equation (3.24) becomes

$$\int_V \Phi_{,i} \delta \Phi_{,i} dV = \int_S \Phi_{,n} dS \quad (3.25)$$

where $\Phi_{,n}$ is the normal gradient of the electric potential to the free surfaces. The right hand side of (3.25) can be omitted because $\Phi_{,n} = 0$ along free surfaces. Having solved the electric potential, the current density J in any direction is available through

$$J_i = \frac{1}{\rho} \Phi_{,i} \quad (3.26)$$

Defining the squared current density as $J^2 = J_i J_i$, the heat generation rate due to Joule heating (3.17) is available through $\dot{q}_{electrical} = \rho J^2$.

3.3 Numerical implementation

The above presented models for the mechanical, thermal and electrical responses are combined and implemented in a finite element computer program based on the flow formulation. This section describes the coupling of the models and the details of computer implementation for each individual model.

3.3.1 Basic coupling procedures

Fig. 3.1 includes a schematic outline of the couplings of the presented models. The thermal and mechanical models are generally coupled as shown in Fig. 3.1a for the purpose of modeling thermo-mechanical metal forming processes, whereas the electrical, thermal and mechanical models are coupled as shown in Fig. 3.1b for the electro-thermo-mechanical modeling of resistance welding processes.

Besides the immediate difference due to the electrical model, the two implementations differ by the number of times the mechanical model is applied during each step. In both cases the mechanical model is applied at the beginning of each step to setup a velocity field and a stress response.

The next step in the thermo-mechanical modeling of metal forming processes is to run the thermal model, and in this case it is run fully coupled with the mechanical

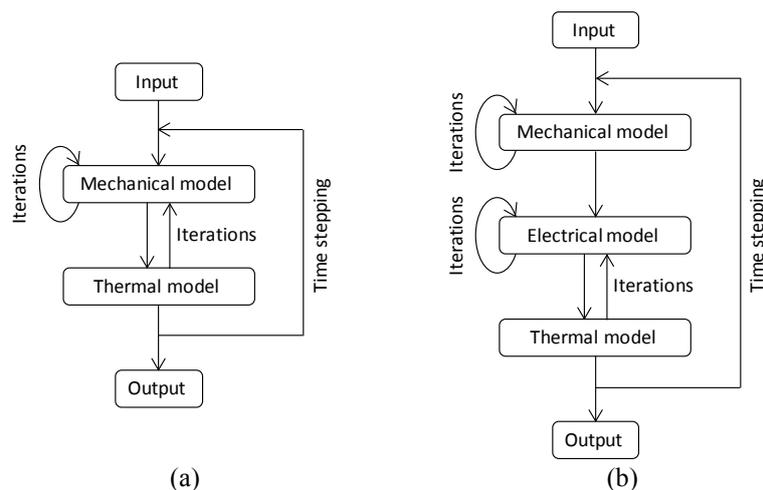


Fig. 3.1. Numerical coupling of mechanical, thermal and electrical models for (a) thermo-mechanical modeling of metal forming processes and (b) electro-thermo-mechanical modeling of resistance welding processes.

model, such that the new temperature field and resulting changes in material properties are converged with the mechanical response including the heat generation at the end of each step.

When it comes to the electro-thermo-mechanical modeling of resistance welding, this strong coupling between the thermal and mechanical models is loosened due to the very small time steps in order to capture the effects of the welding process. For example, when using alternating current with frequency $50Hz$ as energy source, each half period has duration $10ms$, and proper modeling therefore requires time steps of $1ms$ or preferably less. Instead of having a strong coupling, the implementation is relying on the small time steps in a weaker coupling, where the new material properties of the resulting temperature is only affecting the mechanical response from the following time step, and the corresponding change in the heat generation due to plastic work is ignored due to the insignificant influence compared to the electrically generated heat.

In this type of implementation, the electrical model is applied after the mechanical model to supply the thermal model with the current density giving rise to the heat generation. The electrical model is linear and thus inexpensive compared to the mechanical model; hence the electrical and thermal models are strongly coupled such that the electrical model is run during each of the iterations of the thermal model. The implemented coupling is outlined in Fig. 3.1b and follows the work of Zhang et al. [19].

3.3.2 Finite elements

The discretization of the main equations dealing with the physics of plastic flow, heat transfer and electricity is based on 8-node hexahedral elements under three-dimensional conditions. Other elements could be employed in the discretization as will be discussed in Chapter 5 in relation to mesh generation.

The 8-node hexahedral element provides three degrees of freedom in each node for the velocity components of plastic flow in the mechanical model and one degree of freedom for modeling the scalar fields of temperature and potential in the thermal and electrical models, respectively.

In the mechanical model, discretization by hexahedral elements implies that velocity inside an element is interpolated from its nodal values as follows,

$$\mathbf{u} = \mathbf{N}^T \mathbf{v} \quad (3.27a)$$

$$\mathbf{u} = \{u_x, u_y, u_z\}^T \quad (3.27b)$$

$$\mathbf{v} = \{u_x^1, u_y^1, u_z^1, u_x^2, u_y^2, u_z^2, \dots, u_z^8\}^T \quad (3.27c)$$

where \mathbf{u} is the vector containing the velocity components in an arbitrary location within the element, \mathbf{v} is the vector of nodal velocities and \mathbf{N} is a matrix including the shape functions N_i at the corresponding arbitrary location in natural coordinates ξ, η, ζ (e.g. $N_i = (1 + \xi_i \xi)(1 + \eta_i \eta)(1 + \zeta_i \zeta)/8$).

The temperature and electric potential are interpolated similarly, except that the interpolation is for scalars rather than vectors of components. In all cases, the formulation is isoparametric, such that coordinates and field variables are interpolated by the same shape functions.

Matrix notation is introduced in what follows for better describing the computer implementation of the discretized finite element equations.

3.3.3 Mechanical model

Finite element discretization

The strain rate matrix \mathbf{B} relating strain rates to nodal velocities is built from the derivatives of the shape functions in the following manner,

$$\dot{\boldsymbol{\varepsilon}} = \mathbf{B}\mathbf{v} = \mathbf{L}\mathbf{N}^T \mathbf{v}, \quad \mathbf{L} = \begin{bmatrix} \frac{\partial}{\partial x} & 0 & 0 \\ 0 & \frac{\partial}{\partial y} & 0 \\ 0 & 0 & \frac{\partial}{\partial z} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} & 0 \\ 0 & \frac{\partial}{\partial z} & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} & 0 & \frac{\partial}{\partial x} \end{bmatrix} \quad (3.28)$$

Introducing a diagonal matrix $\mathbf{D} = \text{diag}\{\frac{2}{3}, \frac{2}{3}, \frac{2}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$ the effective strain rate (3.6) is written as

$$\left(\dot{\boldsymbol{\varepsilon}}\right)^2 = \dot{\boldsymbol{\varepsilon}}^T \mathbf{D} \dot{\boldsymbol{\varepsilon}} \quad (3.29)$$

or, in the following alternative matrix form after introducing (3.28) and defining $\mathbf{P} = \mathbf{B}^T \mathbf{D} \mathbf{B}$,

$$\left(\dot{\boldsymbol{\varepsilon}}\right)^2 = \mathbf{v}^T \mathbf{B}^T \mathbf{D} \mathbf{B} \mathbf{v} = \mathbf{v}^T \mathbf{P} \mathbf{v} \quad (3.30)$$

The volumetric strain rate $\dot{\varepsilon}_{ii}$ is expressed as follows,

$$\dot{\varepsilon}_{ii} = \mathbf{C}^T \mathbf{B} \mathbf{v} \quad (3.31)$$

with \mathbf{C} being the vectorial form of the Kronecker delta δ_{ij} .

Newton-Raphson iterative procedure

By insertion of the above equations into (3.11), the first derivative of the energy rate functional is obtained as

$$\frac{\partial \Pi}{\partial \mathbf{v}} = \int_V \frac{\bar{\sigma}}{\dot{\varepsilon}} \mathbf{P} \mathbf{v} dV + K \int_V \mathbf{B}^T \mathbf{C} \mathbf{C}^T \mathbf{B} \mathbf{v} dV - \int_S \mathbf{N} \mathbf{F} dS \quad (3.32)$$

where \mathbf{F} is the matrix form of the applied boundary surface tractions $F_i = \sigma_{ij} n_j$.

The second derivative of the energy rate functional is obtained as

$$\frac{\partial^2 \Pi}{\partial \mathbf{v}^2} = \int_V \frac{\bar{\sigma}}{\dot{\varepsilon}} \mathbf{P} dV + \int_V \left(\frac{1}{\dot{\varepsilon}} \frac{\partial \bar{\sigma}}{\partial \dot{\varepsilon}} - \frac{\bar{\sigma}}{\dot{\varepsilon}^2} \right) \frac{1}{\dot{\varepsilon}} \mathbf{P} \mathbf{v} \mathbf{v}^T \mathbf{P} dV + K \int_V \mathbf{B}^T \mathbf{C} \mathbf{C}^T \mathbf{B} dV \quad (3.33)$$

A second order linearization of (3.11) by Taylor expansion near an initial guess $\mathbf{v} = \mathbf{v}_0$ of the velocity field leads to

$$\underbrace{\frac{\partial \Pi}{\partial \mathbf{v}} \Big|_{\mathbf{v}=\mathbf{v}_0}}_{\equiv \mathbf{f}} + \underbrace{\frac{\partial^2 \Pi}{\partial \mathbf{v}^2} \Big|_{\mathbf{v}=\mathbf{v}_0}}_{\equiv \mathbf{K}} \Delta \mathbf{v} = 0 \quad (3.34)$$

which can be discretized by M finite elements and assembled to the system of equations

$$\sum_{m=1}^M \{ \mathbf{K} \Delta \mathbf{v} - \mathbf{f} \} = \mathbf{0} \quad (3.35)$$

From (3.34) and (3.35) it is seen that (3.33) is the stiffness matrix \mathbf{K} and that (3.32) is the load vector \mathbf{f} except for the sign. The stiffness matrix and the load vector are integrated in each element by Gauss integration and assembled into the global system of equations (3.35), which is solved for the velocity increment $\Delta \mathbf{v}$. The velocity \mathbf{v} is updated according to

$$\mathbf{v}_n = \mathbf{v}_{n-1} + \alpha_{NR} \Delta \mathbf{v}_n \quad (3.36)$$

where n is the iteration number and $\alpha_{NR} \in]0;1[$ is a deceleration coefficient to avoid overshooting and oscillations in the solution. The update is carried out until convergence,

$$\frac{|\Delta \mathbf{v}_n|}{|\mathbf{v}_{n-1}|} < \alpha_{conv} \quad (3.37)$$

that is, until the velocity field \mathbf{v} is not changed considerable by including one more iteration. A typical value of α_{conv} is taken around 10^{-5} .

Direct iterations

When applying direct iterations, the constitutive relation is evaluated at the previous converged velocity field, such that the iterations become linear. By insertion of (3.27), (3.30) and (3.31) into the variation of the functional (3.11) and canceling out the virtual velocity field $\delta \mathbf{v}^T$ due to arbitrariness, the following system of equations is obtained,

$$\underbrace{\left(\int_V \frac{\bar{\sigma}}{\varepsilon} \mathbf{P} dV + K \int_V \mathbf{B}^T \mathbf{C} \mathbf{C}^T \mathbf{B} dV \right)}_{\mathbf{K}} \mathbf{v} = \underbrace{\int_S \mathbf{N} \mathbf{F} dS}_{\mathbf{f}} \quad (3.38)$$

where the stiffness matrix \mathbf{K} and the load vector \mathbf{f} are defined as well. Discretization by M finite elements and assembling into a global system of equations, (3.38) lead to

$$\sum_{m=1}^M \{\mathbf{K} \mathbf{v} - \mathbf{f}\} = \mathbf{0} \quad (3.39)$$

with update following

$$\mathbf{v}_n = \alpha_D \mathbf{v}_n + (1 - \alpha_D) \mathbf{v}_{n-1} \quad (3.40)$$

In the above equation n is the iteration number and $\alpha_D \in]0;1[$ is a measure of the degree of updating, which acts as a stabilizer to avoid the solution to overshoot.

Combination of direct and Newton-Raphson iterative procedures

The Newton-Raphson iterative procedure usually results in fast convergence near the actual solution, i.e. when a good estimate of the initial guess $\mathbf{v} = \mathbf{v}_0$ is provided. The initial velocity field can, however, be difficult to obtain and, therefore, the procedure employed in direct iterations is often applied to generate a velocity field close to the actual solution before the Newton-Raphson solution is applied for a fast convergence towards the required tolerance (3.37).

In the first step (that is, at the beginning of the numerical simulation), a velocity field corresponding to a constant strain rate in all elements may serve as the starting point for the direct iterations.

Schematic illustrations of the two iterative procedures are provided in Fig. 3.2 for a simplified one-dimensional velocity field. Fig. 3.2a illustrates the fast convergence of the direct iterations in the early stages and Fig. 3.2b shows the fast convergence of the Newton-Raphson iterations near the solution. Fig. 3.2c shows divergence with the Newton-Raphson iterative procedure in case of an initial guess for the velocity field further away from the actual solution or in case of a sudden complication due to nonlinearities such as contact.

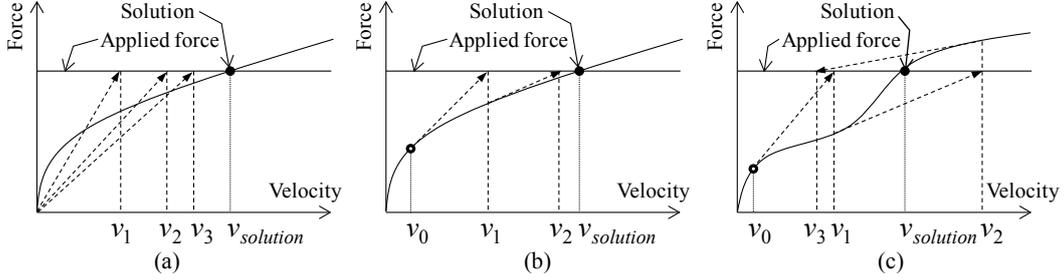


Fig. 3.2. Convergence schemes with subscript numbers referring to iteration number. (a) Direct iterations. (b) Newton-Raphson iterations with convergence. (c) Newton-Raphson iterations with divergence. Subscript zero identifies the initial guess for Newton-Raphson iterations.

In case of divergence of the Newton-Raphson iterative procedure, convergence may be sought with direct iterations.

Selection of deceleration coefficients

As mentioned previously the deceleration coefficients α_D and α_{NR} control the degree of updating of both direct and Newton-Raphson iterative procedures. In case of direct iterations the selection of α_D is obtained after analyzing the ratios $\|\mathbf{v}_n - \mathbf{v}_{n-1}\|/\|\mathbf{v}_{n-1}\|$ and $\|\mathbf{R}_{n-1}\|/\|\mathbf{f}\|$ of the velocity \mathbf{v} and residual \mathbf{R} at iterations n and $n-1$, where

$$\mathbf{R}_{n-1} = \sum_{m=1}^M \{\mathbf{K}_{n-1} \mathbf{v}_n - \mathbf{f}\} \quad (3.41)$$

A similar approach is performed in case of Newton-Raphson iterative procedures, where the residual \mathbf{R} is obtained from a Taylor expansion of the residual near the velocity estimate at the previous iteration,

$$\mathbf{R}(\mathbf{v}_n) \approx \mathbf{R}_n = \mathbf{R}_{n-1} + \left[\frac{\partial \mathbf{R}}{\partial \mathbf{v}} \right]_{n-1} \Delta \mathbf{v}_n = \mathbf{0} \quad (3.42)$$

In addition, it is also a good choice to determine the deceleration coefficient α_{NR} by means of a line search procedure that consider the residual \mathbf{R} at the end of each iteration to be orthogonal to the velocity correction term $\Delta \mathbf{v}$ [21],

$$\Delta \mathbf{v}_n^T \cdot \mathbf{R}(\mathbf{v}_{n-1} + \alpha_{NR} \Delta \mathbf{v}_n) = \mathbf{0} \quad (3.43)$$

Domain integration

The integration of the integrals in (3.32), (3.33) and (3.38) is performed by means of a selective Gauss integration scheme. Volume integrals are integrated by full integration

(2^3 Gauss points) except for the second term in (3.32), the last term in (3.33) and the second term in (3.38), which are related to the volumetric part of the stiffness matrix \mathbf{K} . These terms are integrated by reduced Gauss integration (one Gauss point) to avoid locking. The surface integrals that include boundary pressure and friction along the tools are integrated by 5^2 Gauss points [22].

Stress calculation

The direct results of (3.35) and (3.39) are the velocities and the strain rates. The strains are accumulated at the end of each simulation step by multiplying the strain rates by the increment of time and the effective strain allows determination of the effective stress directly from the applied material law.

The distribution of stress at the end of each simulation step requires determining the mean stress σ_m (refer to Section 3.2.1),

$$\sigma_m = \frac{K}{2} \dot{\epsilon}_{kk} \quad (3.44)$$

and adding this value to the corresponding deviatoric stress obtained from the constitutive equations via the strain rate values (3.4) and (3.5),

$$\sigma_{ij} = \sigma'_{ij} + \delta_{ij} \sigma_m \quad (3.45)$$

The penalty K may be chosen as a constant value or as an adaptive value that changes for each element. If an adaptive value is chosen, small elements take larger penalty values because small elements are generally placed in the regions of higher interest. The accuracy is thereby increased in the regions with refined mesh, while keeping the overall penalization as low as possible in order to diminish ill-conditioning of the matrix systems.

An option is to scale the penalty K according to the ratio of the maximum element volume to the actual element volume. If any scaling factor is above 10, all scaling factors are rescaled such that the maximum scaling is 10. This is to avoid very large penalty factors resulting in increased ill-conditioning.

Rigid regions

To avoid singularities in the system of equations when having rigid regions, where the strain rates approach zero, a cut-off strain rate $\dot{\epsilon}_0$ is introduced [2]. Whenever $\dot{\epsilon} < \dot{\epsilon}_0$, the cut-off strain rate $\dot{\epsilon}_0$ replaces the actual strain rate to overcome the problem of singularities. The cut-off strain rate is taken as a value considerably smaller than the average strain rate of the deforming body. A too large value will model rigid regions poorly, and a too small value may lead to numerical inaccuracies.

An improvement of the above approach has been implemented to avoid excessive strain accumulation in rigid regions. The strain is only accumulated if the equivalent strain rate is increasing (which will not be the case if it is constantly equal to the cut-off strain rate) or if the equivalent strain has already exceeded a certain strain level meaning that the region should not be treated as rigid.

3.3.4 Thermal model

Using the same shape functions as for the mechanical model, the temperature can be interpolated as

$$T = \mathbf{N}^T \mathbf{T} \quad (3.46)$$

where \mathbf{T} contains the nodal temperatures and \mathbf{N} contains the shape functions at positions to realize the summation over nodal values. Similarly a matrix \mathbf{N}' is defined such that

$$T_{,i} = \mathbf{N}'^T \mathbf{T} \quad (3.47)$$

by having $N'_{ij} = N_{i,j}$.

Inserting (3.46) and (3.47) into (3.22) and canceling out the arbitrary temperature variation, the system of equations for the thermal model, $\mathbf{K}_c \mathbf{T} + \mathbf{C}_c \dot{\mathbf{T}} = \mathbf{q}$, becomes

$$\underbrace{\int_V k \mathbf{N}' \mathbf{N}'^T dV}_{\equiv \mathbf{K}_c} \mathbf{T} + \underbrace{\int_V \rho_m c_m \mathbf{N} \mathbf{N}^T dV}_{\equiv \mathbf{C}_c} \dot{\mathbf{T}} = \underbrace{\int_V \dot{q}_{generate} \mathbf{N} dV + \int_S k T_{,n} \mathbf{N} dS}_{\equiv \mathbf{q}} \quad (3.48)$$

where \mathbf{K}_c is the heat conduction matrix, \mathbf{C}_c is the heat capacity matrix and \mathbf{q} includes the boundary flux and the source term. The right hand side \mathbf{q} is expanded as follows to include the heat sources and heat losses due to equations (3.16)-(3.21) as in (3.23),

$$\begin{aligned} \mathbf{q} = & \int_V (\dot{q}_{plastic} + \dot{q}_{electrical}) \mathbf{N} dV - \int_{S_{free}} (\dot{q}_{convection} + \dot{q}_{radiation}) \mathbf{N} dS \\ & - \int_{S_{tool}} (\dot{q}_{tool} - \dot{q}_{friction}) \mathbf{N} dS \end{aligned} \quad (3.49)$$

The domain integration of the thermal system of equations (3.48) is performed over hexahedral elements in the usual manner, whereas the time integration is more complicated. The presence of the term including $\dot{\mathbf{T}}$ makes the system of equations differ from typical forms utilized in the mechanical models, e.g. (3.39). Details regarding the solution of the system of equations can be found in several references, e.g. in the pioneer-

ing work of Rebelo and Kobayashi [13-14], which requires the utilization of the following time-stepping scheme,

$$\mathbf{T}_{t+\Delta t} = \mathbf{T}_t + \Delta t \left[(1-\theta)\dot{\mathbf{T}}_t + \theta\dot{\mathbf{T}}_{t+\Delta t} \right] \quad (3.50)$$

where θ is a parameter varying between 0 and 1. A value of $\theta = 0.75$ is typically chosen.

3.3.5 Electrical model

The shape functions and shape function derivatives are introduced similarly to (3.46) and (3.47) in the electrical model, such that they interpolate the potential and its derivatives as follows,

$$\Phi = \mathbf{N}^T \mathbf{\Phi} \quad (3.51)$$

$$\Phi_{,i} = \mathbf{N}'^T \mathbf{\Phi} \quad (3.52)$$

Inserting (3.52) in (3.25) and canceling out the right hand side and the arbitrary potential variation, the discretized form of the electrical model (3.25) can be written as

$$\underbrace{\int_V \mathbf{N}' \mathbf{N}'^T dV}_{\equiv \mathbf{K}_e} \mathbf{\Phi} = \mathbf{0} \quad (3.53)$$

where \mathbf{K}_e is the electrical conductance matrix to be integrated over elements and assembled into the global system of equations.

3.4 Incorporation of anisotropy

Finite element modeling of manufacturing processes often treats materials as isotropic but when it comes to materials supplied as sheets, anisotropic behavior can be important due to the effect of prior rolling of the material. This section describes the implementation of Hill's quadratic anisotropic yield criterion [23] and subsection 3.4.1 describes the necessary rotation between global axes and local material axes as they in general differ after deformation.

Hill's quadratic anisotropic yield criterion takes the following form,

$$f^a = \frac{1}{2} \frac{3}{2(F+G+H)} \left[F(\sigma_{22} - \sigma_{33})^2 + G(\sigma_{33} - \sigma_{11})^2 + H(\sigma_{11} - \sigma_{22})^2 + 2L\sigma_{23}^2 + 2M\sigma_{31}^2 + 2N\sigma_{12}^2 \right] \quad (3.54)$$

where the anisotropic parameters, F , G , H , L , M and N , are to be determined from material testing through the following relations involving uniaxial and shear effective stresses $\bar{\sigma}_{ij}$,

$$\begin{aligned} \frac{1}{\bar{\sigma}_{11}^2} &= G + H, & \frac{1}{\bar{\sigma}_{22}^2} &= H + F, & \frac{1}{\bar{\sigma}_{33}^2} &= F + G, \\ \frac{1}{\bar{\sigma}_{12}^2} &= 2N, & \frac{1}{\bar{\sigma}_{23}^2} &= 2L, & \frac{1}{\bar{\sigma}_{31}^2} &= 2M \end{aligned} \quad (3.55)$$

The yield function, (3.54), can be written as

$$f^a = \frac{1}{2} \frac{3}{2(F + G + H)} \sigma_{ij} P_{ijkl} \sigma_{kl} \quad (3.56)$$

where

$$\begin{aligned} P_{ijkl} &= \begin{bmatrix} P_{1111} & P_{1122} & P_{1133} & P_{1112} & P_{1123} & P_{1131} \\ P_{2211} & P_{2222} & P_{2233} & P_{2212} & P_{2223} & P_{2231} \\ P_{3311} & P_{3322} & P_{3333} & P_{3312} & P_{3323} & P_{3331} \\ P_{1211} & P_{1222} & P_{1233} & P_{1212} & P_{1223} & P_{1231} \\ P_{2311} & P_{2322} & P_{2333} & P_{2312} & P_{2323} & P_{2331} \\ P_{3111} & P_{3122} & P_{3133} & P_{3112} & P_{3123} & P_{3131} \end{bmatrix} \\ &= \begin{bmatrix} G + H & -H & -G & 0 & 0 & 0 \\ -H & F + H & -F & 0 & 0 & 0 \\ -G & -F & F + G & 0 & 0 & 0 \\ 0 & 0 & 0 & 2N & 0 & 0 \\ 0 & 0 & 0 & 0 & 2L & 0 \\ 0 & 0 & 0 & 0 & 0 & 2M \end{bmatrix} = \begin{bmatrix} [\mathbf{P}^{UL}] & [\mathbf{0}] \\ [\mathbf{0}] & [\mathbf{P}^{LR}] \end{bmatrix} \end{aligned} \quad (3.57)$$

The upper left (UL) and lower right (LR) submatrices are identified for later use.

The yield function f^a (3.56) is defined as counterpart to the isotropic yield function f associated with the von Mises yield criterion (3.3) and the effective stress is defined as

$$\bar{\sigma} = \sqrt{\frac{3}{2(F + G + H)}} \left\{ \sigma_{ij} P_{ijkl} \sigma_{kl} \right\}^{\frac{1}{2}} \quad (3.58)$$

which is the counterpart to the effective stress associated with von Mises isotropic yield criterion (3.7).

The proportionality factor associated with Hill's criterion is given by

$$\dot{\lambda} = \frac{\dot{\bar{\epsilon}}}{\bar{\sigma}} \quad (3.59)$$

Insertion of (3.56) and (3.59) into the flow rule leads to the counterpart of the Levy-Mises constitutive equations; namely the relation between strain rates and deviatoric stresses that is consistent with Hill's criterion,

$$\dot{\epsilon}_{ij} = \frac{3}{2(F+G+H)} \frac{\dot{\bar{\epsilon}}}{\bar{\sigma}} P_{ijkl} \sigma_{kl} = \frac{3}{2(F+G+H)} \frac{\dot{\bar{\epsilon}}}{\bar{\sigma}} P_{ijkl} \sigma'_{kl} \quad (3.60)$$

where the last equality is seen by insertion of $\sigma_{ij} = \sigma'_{ij} + \delta_{ij} \sigma_m$ and recognition of $P_{ijkl} \delta_{kl} = 0$ for any ij .

The deviatoric stress components are available through

$$\sigma'_{ij} = \frac{2(F+G+H)}{3} \frac{\bar{\sigma}}{\dot{\bar{\epsilon}}} M_{ijkl} \dot{\epsilon}_{kl} \quad (3.61)$$

which follows from (3.60), except for the fact that P_{ijkl} is singular and therefore cannot be inverted. The tensor M_{ijkl} is therefore introduced instead of the non-existing inversion of P_{ijkl} . The structure of M_{ijkl} is

$$[\mathbf{M}] = \begin{bmatrix} [\mathbf{M}^{UL}] & [\mathbf{0}] \\ [\mathbf{0}] & [\mathbf{M}^{LR}] \end{bmatrix} \quad (3.62)$$

due to the structure of P_{ijkl} . The two non-zero submatrices are independent inversions of the corresponding submatrices defined in (3.57) as long as they would be regular. The lower right submatrix in (3.57) is regular, so $[\mathbf{M}^{LR}] = [\mathbf{P}^{LR}]^{-1}$. The upper left submatrix in (3.57) is singular, so $[\mathbf{P}^{UL}]^{-1}$ does not exist. Instead, $[\mathbf{M}^{UL}]$ is introduced such that

$$[\mathbf{M}^{UL}] [\mathbf{P}^{UL}] \begin{Bmatrix} \sigma'_{11} \\ \sigma'_{22} \\ \sigma'_{33} \end{Bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{Bmatrix} \sigma'_{11} \\ \sigma'_{22} \\ \sigma'_{33} \end{Bmatrix} = \begin{Bmatrix} \sigma'_{11} \\ \sigma'_{22} \\ \sigma'_{33} \end{Bmatrix} \quad (3.63)$$

since this will have the same effect as if $[\mathbf{M}^{UL}]$ was equal to $[\mathbf{P}^{UL}]^{-1}$. This is possible due to the last equality sign where it is utilized that $\sigma'_{ii} = 0$. The matrix $[\mathbf{M}^{UL}]$ satisfying (3.63) is written out together with $[\mathbf{M}^{LR}]$ to form the entire tensor M_{ijkl} with positions as defined in (3.57),

$$M_{ijkl} = \begin{bmatrix} Fk & -(F+G)k & -(H+F)k & 0 & 0 & 0 \\ -(F+G)k & Gk & -(G+H)k & 0 & 0 & 0 \\ -(H+F)k & -(G+H)k & Hk & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2N} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2L} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2M} \end{bmatrix}, \quad (3.64)$$

$$k = \frac{1}{3(FG + FH + GH)}$$

Insertion of (3.61) into $\overline{\sigma \dot{\varepsilon}} = \sigma_{ij} \dot{\varepsilon}_{ij}$ yields an expression for the effective strain rate similar to that of the isotropic formulation based on von Mises' yield criterion (3.6),

$$\dot{\varepsilon} = \sqrt{\frac{2(F+G+H)}{3}} \left\{ \dot{\varepsilon}_{ij} M_{ijkl} \dot{\varepsilon}_{kl} \right\}^{\frac{1}{2}} \quad (3.65)$$

which by insertion of (3.64) and utilization of $\dot{\varepsilon}_{ii} = 0$ leads to

$$\dot{\varepsilon} = \sqrt{\frac{2(F+G+H)}{3}} \left\{ \frac{F\dot{\varepsilon}_{11}^2 + G\dot{\varepsilon}_{22}^2 + H\dot{\varepsilon}_{33}^2}{FG + FH + GH} + \frac{(2\dot{\varepsilon}_{12})^2}{2N} + \frac{(2\dot{\varepsilon}_{23})^2}{2L} + \frac{(2\dot{\varepsilon}_{31})^2}{2M} \right\}^{\frac{1}{2}} \quad (3.66)$$

From (3.66) it is seen that an equation similar to (3.29) can be set up by defining another diagonal \mathbf{D} -matrix, namely

$$\mathbf{D}^a = \frac{2(F+G+H)}{3} \text{diag} \left\{ \frac{\{F, G, H\}}{FG + FH + GH}, \frac{1}{2N}, \frac{1}{2L}, \frac{1}{2M} \right\} \quad (3.67)$$

which is related to the effective strain rate like in (3.29). The anisotropic finite element formulation follows the derivations in Section 3.3 with substitution of (3.67) into (3.29) and (3.30).

3.4.1 Rotation between global axes and material axes

In the above formulation, \mathbf{D}^a refers to the global coordinate system, which may not be the same as the material coordinate system. In general, part of the deformation is rigid body rotation, which gives rise to misalignment between material axes and global axes. Therefore an incremental rotation matrix is set up to rotate \mathbf{D}^a in each step according to the rigid body rotations associated with the previous step.

From the updated nodal velocities, the spin rate tensor,

$$\dot{\omega}_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i} \right) = \begin{bmatrix} 0 & \omega_{12} & \omega_{13} \\ -\omega_{12} & 0 & \omega_{23} \\ -\omega_{13} & -\omega_{23} & 0 \end{bmatrix} \quad (3.68)$$

can be calculated in each step in each element. It is set up for the central point (in natural coordinates) of each element through the shape function derivatives. Assuming small incremental rigid body rotations, the incremental rotation matrix is approximated by adding the unit matrix and the incremental spin matrix, i.e.

$$\Delta \mathbf{R} = \mathbf{I} + \omega \Delta t \quad (3.69)$$

This incremental rotation matrix is a 3×3 matrix, which rotates another 3×3 matrix $\tilde{\mathbf{D}}^a$ through

$$\tilde{\mathbf{D}}_k^a = \Delta \mathbf{R} \tilde{\mathbf{D}}_{k-1}^a \Delta \mathbf{R}^T \quad (3.70)$$

where k represents step number. \mathbf{D}^a in (3.67) is transferred into a 3×3 format during the rotation by the following translations between positions, pos_i , (used towards right before rotation and used towards left after rotation),

$$\text{diag}\{pos_1, pos_2, pos_3, pos_4, pos_5, pos_6\} \leftrightarrow \begin{bmatrix} pos_1 & pos_4 & pos_6 \\ pos_4 & pos_2 & pos_5 \\ pos_6 & pos_5 & pos_3 \end{bmatrix} \quad (3.71)$$

This translation follows the translation between the stress vector and the stress matrix defined as

$$\boldsymbol{\sigma} = \{\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12}, \sigma_{23}, \sigma_{31}\}^T \leftrightarrow \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{31} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{23} & \sigma_{33} \end{bmatrix} \quad (3.72)$$

The stress calculation should also be carried out with attention to rotation. The deviatoric stress is related to the strain rate according to (3.61). In this equation, M_{ijkl} is defined in a material coordinate system, whereas the available strain rates are defined relative to the global coordinate system. Due to possible material rotation, these two systems may not coincide. It is therefore necessary to rotate one of the quantities from one system to the other. The implemented procedure is as follows; the strain rate in each element is rotated from global axes to material axes through the following rotation,

$$\dot{\boldsymbol{\varepsilon}}_{mat} = \mathbf{R} \dot{\boldsymbol{\varepsilon}} \mathbf{R}^T \quad (3.73)$$

where \mathbf{R} is the accumulated rotation during N_{rot} rotation steps defined as

$$\mathbf{R} = \mathbf{I} + \sum_{i=1}^{N_{rot}} (\boldsymbol{\omega} \Delta t)_i \quad (3.74)$$

Now in material axes, the deviatoric stress can be computed by insertion of (3.73) into (3.61), and finally the deviatoric stress in the global system is available by rotation,

$$\boldsymbol{\sigma}' = \mathbf{R}^T \boldsymbol{\sigma}'_{mat} \mathbf{R} \quad (3.75)$$

Rotation of the Cauchy stress is possible without introduction of artificial contributions from rigid body rotation since it is an objective stress measure. The remaining stress calculation follows (3.44) and (3.45).

As a final remark to the anisotropic formulation, it should be mentioned that with $F = G = H = 1$ and $L = M = N = 3$, the anisotropic formulation reduces to the isotropic formulation.

3.5 Incorporation of elastic effects

The core of the mechanical model is the rigid-plastic/viscoplastic flow formulation as presented so far. In this formulation the elastic effects are neglected due to the large deformations typically simulated. The elastic effects may, however, be of importance in some cases when only part of the volume is heavily deformed. In these cases, the

remaining volume will only deform slightly, such that the elastic part should not be neglected.

An example where elastic deformation is of importance is resistance welding including bending of a sheet (gap between sheets or welding of a component to a sheet structure). In this case the overall deformation is governed by elastic deformation and only local deformation is governed by plasticity. The amount of elastic bending can be of importance to the actual contact area, which is essential for the welding process.

Elastic effects can be included in computer programs based on the finite element flow formulation following the procedure or variants of the procedure proposed by Mori et al. [24]. By doing this, the elastic effects are captured while the advantages of the flow formulation are kept for the remaining elements considered rigid-plastic due to large deformations. A possible implementation of this procedure can be implemented as described in what follows.

All elements are initialized as elastic elements before loading. After loading to the vicinity of the yield stress Y , the relevant elements are turned into elastoplastic elements, and after further loading the relevant elements are turned into rigid-plastic elements ignoring any further elastic deformation. In order for the programs to be more efficient, a range of stress is assigned to define the elastoplastic behavior of the elements. With reference to Fig. 3.3a, the constitutive laws are applied as follows,

$$\text{constitutive law} = \begin{cases} \text{elastic,} & \bar{\sigma} \leq f_l Y \\ \text{elastoplastic,} & f_l Y < \bar{\sigma} < f_u Y \\ \text{rigid - plastic/viscoplastic,} & \bar{\sigma} \geq f_u Y \end{cases} \quad (3.76)$$

where typical factors are chosen around $f_l = 0.95$ and $f_u = 1.01$. The flattened curve after yielding is reflected in the upper factor being closer to unity than the lower factor.

A stress situation in the vicinity of yielding is illustrated in Fig. 3.3b, where a stress path is exceeding the yield stress of the material causing strain hardening. The present stress state P is elastic with effective stress less than the yield stress, $\bar{\sigma}_t < Y$. The assumed load increment will cause a stress path through yielding (point Q) followed by

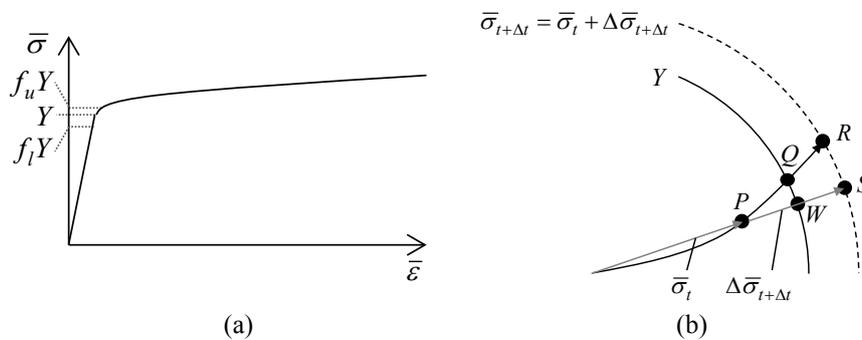


Fig. 3.3. Definitions in the vicinity of the yield stress Y . (a) Limits defining elastic, elastoplastic and rigid-plastic/viscoplastic regions. (b) Stress path for definition of elastic and elastoplastic fractions of stress increment.

strain hardening to a stress state in point R with effective stress, $\bar{\sigma}_{t+\Delta t} = \bar{\sigma}_t + \Delta\bar{\sigma}_{t+\Delta t}$, equal to the new flow stress. A ratio R_e of the elastic part of the stress to the total stress increment is defined and approximated, respectively, as follows with reference to Fig. 3.3b,

$$R_e = \frac{\overline{PQ}}{\overline{PR}} \approx \frac{\overline{PW}}{\overline{PS}} = \frac{Y - \bar{\sigma}_t}{\bar{\sigma}_{t+\Delta t} - \bar{\sigma}_t} \quad (3.77)$$

Yamada et al. [25] presented the correct solution corresponding to $\frac{\overline{PQ}}{\overline{PR}}$, but the approximation by $\frac{\overline{PW}}{\overline{PS}}$ is considered sufficient for the present purpose.

The ratio R_e was originally used to scale the load increment according to the elastic element closest to yielding to achieve a situation where it just reaches the yield stress. Hereafter, the element will be considered plastic. Another approach is to avoid splitting the time step (corresponding to the load increment). The ratio of the elastic contribution to the stress increment is instead used to scale the amount of the stress-strain matrix stemming from either the elastic relation or the elastoplastic relation according to

$$\Delta\boldsymbol{\sigma} = \mathbf{D}^{ep} \Delta\boldsymbol{\varepsilon} = (R_e \mathbf{D}^e + (1 - R_e) \mathbf{D}^p) \Delta\boldsymbol{\varepsilon} \quad (3.78)$$

where e refers to elasticity and p to plasticity.

For pure elasticity (3.78) reduces to Hooke's generalized law after inversion and the elastic stress-strain matrix is written as follows,

$$\mathbf{D}^e = \frac{E}{(1+\nu)(1-2\nu)} \begin{bmatrix} 1-\nu & \nu & \nu & 0 & 0 & 0 \\ \nu & 1-\nu & \nu & 0 & 0 & 0 \\ \nu & \nu & 1-\nu & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1-2\nu}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1-2\nu}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1-2\nu}{2} \end{bmatrix} \quad (3.79)$$

For elastoplasticity, (3.78) resembles the inverse Prandtl-Reuss equations. The starting point is taken by the deviatoric part of the Prandtl-Reuss equations,

$$d\varepsilon'_{ij} = \sigma'_{ij} d\lambda + \frac{d\sigma'_{ij}}{2G} \quad (3.80a)$$

$$d\lambda = \frac{3}{2} \frac{d\bar{\varepsilon}^p}{\bar{\sigma}} = \frac{3}{2} \frac{d\bar{\sigma}}{\bar{\sigma}H'} \quad (3.80b)$$

where $G = \frac{E}{2(1+\nu)}$ is the shear modulus and $H' = \frac{d\bar{\sigma}}{d\bar{\varepsilon}^p}$ is the slope of the stress-strain curve. The corresponding elastoplastic stress-strain matrix originally obtained by Yamada et al. [25] is built by inverting (3.80) and can be written as follows,

$$\mathbf{D}^p = \frac{1}{1+\nu} \begin{bmatrix} \frac{1-\nu}{1-2\nu} - \frac{\sigma_{11}^2}{S} & \frac{\nu}{1-2\nu} - \frac{\sigma_{11}\sigma'_{22}}{S} & \frac{\nu}{1-2\nu} - \frac{\sigma_{11}\sigma'_{33}}{S} & -\frac{\sigma'_{11}\sigma_{12}}{S} & -\frac{\sigma'_{11}\sigma_{23}}{S} & -\frac{\sigma'_{11}\sigma_{31}}{S} \\ \frac{\nu}{1-2\nu} - \frac{\sigma_{11}\sigma'_{22}}{S} & \frac{1-\nu}{1-2\nu} - \frac{\sigma_{22}^2}{S} & \frac{\nu}{1-2\nu} - \frac{\sigma'_{22}\sigma'_{33}}{S} & -\frac{\sigma'_{22}\sigma_{12}}{S} & -\frac{\sigma'_{22}\sigma_{23}}{S} & -\frac{\sigma'_{22}\sigma_{31}}{S} \\ \frac{\nu}{1-2\nu} - \frac{\sigma_{11}\sigma'_{33}}{S} & \frac{\nu}{1-2\nu} - \frac{\sigma'_{22}\sigma'_{33}}{S} & \frac{1-\nu}{1-2\nu} - \frac{\sigma_{33}^2}{S} & -\frac{\sigma'_{33}\sigma_{12}}{S} & -\frac{\sigma'_{33}\sigma_{23}}{S} & -\frac{\sigma'_{33}\sigma_{31}}{S} \\ -\frac{\sigma_{11}\sigma_{12}}{S} & -\frac{\sigma'_{22}\sigma_{12}}{S} & -\frac{\sigma_{33}\sigma_{12}}{S} & \frac{1}{2} - \frac{\sigma_{12}^2}{S} & -\frac{\sigma'_{12}\sigma_{23}}{S} & -\frac{\sigma'_{12}\sigma_{31}}{S} \\ -\frac{\sigma_{11}\sigma_{23}}{S} & -\frac{\sigma'_{22}\sigma_{23}}{S} & -\frac{\sigma_{33}\sigma_{23}}{S} & -\frac{\sigma'_{12}\sigma_{23}}{S} & \frac{1}{2} - \frac{\sigma_{23}^2}{S} & -\frac{\sigma'_{23}\sigma_{31}}{S} \\ -\frac{\sigma_{11}\sigma_{31}}{S} & -\frac{\sigma'_{22}\sigma_{31}}{S} & -\frac{\sigma_{33}\sigma_{31}}{S} & -\frac{\sigma'_{12}\sigma_{31}}{S} & -\frac{\sigma'_{23}\sigma_{31}}{S} & \frac{1}{2} - \frac{\sigma_{31}^2}{S} \end{bmatrix} \quad (3.81)$$

with

$$S = \frac{2}{3} \bar{\sigma}^2 \left(1 + \frac{1}{3G} \frac{d\bar{\sigma}}{d\bar{\varepsilon}^p} \right) \quad (3.82)$$

The elastoplastic solution presented by (3.78) with elastic and elastoplastic stress relations by (3.79) and (3.81) requires the stress to be incremented in each step, which is not the case in the flow formulation where the stress is given solely by the accumulated effective strain and the strain rate of the current step. In the flow formulation, the stress is therefore not necessarily saved between steps unless written to result files. On the contrary, in the solid formulations, the stress field of the previous step is of importance as the new step is only solving a stress increment. The stress of the previous step enters the equations as an initial stress, and in the end of the step it is incremented by the solution obtained in (3.78).

In general, the deformation will include rigid body motion between simulation steps. It is therefore necessary at each step to rotate the stress from the previous step into the new configuration, both for the role of initial stress and for the incremental update in the end of the step. With incremental rotation as defined in (3.69) and calculated stress increment $\Delta\boldsymbol{\sigma}_{t+\Delta t}$, the stress after the new time step is

$$\boldsymbol{\sigma}_{t+\Delta t} = \Delta\boldsymbol{\sigma}_{t+\Delta t} + \Delta\mathbf{R}\boldsymbol{\sigma}_t\Delta\mathbf{R}^T \quad (3.83)$$

where the last term is identical to the stress field of the previous time step rotated into the new configuration. This term is also applied as the initial stress.

The presented formulation includes a mixture of elastic, elastoplastic and rigid-plastic/viscoplastic elements. Whenever elastic effects are relevant, all the elements are initialized as elastic as mentioned previously. They are changed to elastoplastic elements in the vicinity of yielding according to (3.76) and later changed to rigid-plastic/viscoplastic elements. The different states of the elements are working simulta-

neously, implying that typical situations will include a local deformation zone with rigid-plastic/viscoplastic elements, a transition zone of elastoplastic elements while the remaining elements are elastic.

Elastic unloading at the end of a simulation is performed by changing all elements to the elastic state and performing one more iteration step with the actual stress field as the initial stress. Dynamic elastic unloading was covered by Mori et al. [24] by changing elements back to the elastic state according to (3.76) and Fig. 3.3a.

References

- [1] Boer CR, Gudmundson P, Rebelo N (1982) Comparison of elastoplastic FEM, rigid-plastic FEM and experiments for cylinder upsetting. Numerical Methods for Industrial Forming Processes, Pittman JFT (Ed.), Pineridge Press, Swansea.
- [2] Kobayashi S, Oh SI, Altan T (1989) Metal forming and the finite element method. Oxford University Press, Oxford.
- [3] Brännberg N, Mackerle J (1994) Finite element methods and material processing technology. Engineering Computations 11:413-455.
- [4] Mackerle J (1998) Finite element methods and material processing technology, an addendum (1994-1996). Engineering Computations 15:616-690.
- [5] Mackerle J (2004) Finite element analyses and simulations of sheets metal forming processes. Engineering Computations 21:891-940.
- [6] Lee CH, Kobayashi S (1973) New solutions to rigid plastic deformation problems using a matrix method. Journal of Engineering for Industry, ASME, 95:865-873.
- [7] Cornfield GC, Johnson RH (1973) Theoretical prediction of plastic flow in hot rolling including the effect of various temperature distributions. The Journal of the Iron and Steel Institute 211:567-573.
- [8] Zienkiewicz OC, Godbole PN (1974) Flow of plastic and viscoplastic solids with special reference to extrusion and forming processes. International Journal for Numerical Methods in Engineering 8:3-16.
- [9] Altan T, Knoerr M (1992) Application of 2D finite element method to simulation of cold forging process. Journal of Materials Processing Technology 35:275-302.
- [10] Zienkiewicz OC, Jain PC, Oñate E (1978) Flow of solids during forming and extrusion. Some aspects of numerical solutions. International Journal of Solids and Structures 12:15-38.
- [11] Zienkiewicz OC, Oñate E, Heinrich JC (1978) Plastic flow in metal forming – I. Coupled thermal behavior in extrusion – II. Thin sheet forming. Proceedings of Winter Annual Meeting of ASME on Application of Numerical Methods to Forming Processes, San Francisco, 28:107.
- [12] Zienkiewicz OC, Oñate E, Heinrich JC (1981) A general formulation for coupled thermo flow of metals using finite elements. International Journal for Numerical Methods in Engineering 17:1497-1514.
- [13] Rebelo N, Kobayashi S (1980) A coupled analysis of viscoplastic deformation and heat transfer – I. Theoretical considerations. International Journal of Mechanical Sciences 22:699-706.
- [14] Rebelo N, Kobayashi S (1980) A coupled analysis of viscoplastic deformation and heat transfer – II. Applications. International Journal of Mechanical Sciences 22:707-718.

- [15] Nied HA (1984) The Finite Element Modeling of the Resistance Spot Welding Process. *Welding Research Supplement* 63:123-132.
- [16] Zhu W-F, Lin ZQ, Lai X-M, Luo A-H (2006) Numerical analysis of projection welding on auto-body sheet metal using a coupled finite element method. *International Journal of Advanced Manufacturing Technology* 28:45-52.
- [17] Zhang W (2010) Weld planning with optimal welding parameters by computer simulations and optimizations. *Proceedings of the 6th International Seminar on Advances in Resistance Welding*, Hamburg, Germany, pp. 119-127.
- [18] Alves ML, Rodrigues JMC, Martins PAF (2004) Three-dimensional modelling of forging processes by the finite element flow formulation. *Journal of Engineering Manufacture* 218:1695–1707.
- [19] Zhang W, Jensen HH, Bay N (1997) Finite element modeling of spot welding similar and dissimilar metals. *The 7th International Conference on Computer Technology in Welding*, San Francisco, USA, pp. 364-373.
- [20] Greenwood JA, Williamson JBP (1958) Electrical conduction in solids II. Theory of temperature-dependent conductors. *Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences* 246:13-31.
- [21] Alves ML (2004) *Modelação numérica e análise experimental de operações de forjamento* (in Portuguese). PhD Thesis, IST-Technical University of Lisbon.
- [22] Barata Marques MJM, Martins PAF (1990) Three-dimensional finite element contact algorithm for metal forming. *International Journal for Numerical Methods in Engineering*, 30:1341-1354.
- [23] Hill R (1950) *The mathematical theory of plasticity*. Oxford University Press, London.
- [24] Mori K, Wang CC, Osakada K (1996) Inclusion of elastic deformation in rigid-plastic finite element analysis. *International Journal of Mechanical Sciences* 38:621-631.
- [25] Yamada Y, Yoshimura N, Sakurai T (1968) Plastic stress-strain matrix and its application for the solution of elastic-plastic problems by the finite element method. *International Journal of Mechanical Sciences* 10:343-354.

4. Contact Modeling

Due to the highly non-linear behavior, contact modeling remains among the more difficult disciplines within finite element simulations. Contact between workpieces and tooling and in-between workpieces defines the shape of formed components in metal forming as well as the contact conditions in resistance welding between the components to be joined and the welding electrodes. Section 4.1 presents a direct contact algorithm to handle the contact between a deformable workpiece and rigid tools and Section 4.2 presents a variational approach to the contact between deformable objects. Section 4.3 includes a description of the physical contact properties to be modeled in special interface elements handling electrical and thermal contact resistances.

Descriptions are given based on mechanical contact while thermal and electrical contacts are included by simplification of the mechanical description. The mechanical contact conditions can be separated into normal constraints and tangential constraints. The normal constraint is always that the contacting surfaces cannot penetrate into each other. The tangential constraints depend on the treatment of friction. In case of a frictionless approach, there are no tangential constraints and in case of full sticking, the tangential constraints are similar to the normal constraint since relative sliding is prohibited. In case of frictional conditions (including combined sticking and sliding), the constraints are governed by the employed friction law. At low, medium and high normal pressures, the following three friction laws are commonly employed:

- Amonton-Coulomb $\tau_f = \mu p$, typically assumed for normal pressure below $\sim 1.5 p/Y$.
- Law of constant friction (Tresca) $\tau_f = mk$, typically assumed above $\sim 3 p/Y$.
- Wanheim-Bay general friction model $\tau_f = f\alpha k$, applicable over the entire range of normal pressure and especially relevant in the range between the two aforementioned models above $\sim 1.5 p/Y$ and below $\sim 3 p/Y$.

In the above friction models, the friction shear stress is τ_f , p is the normal pressure and Y is the material flow stress of the softest contact surface, k is the shear flow stress, μ is the friction coefficient, f and m are the friction factors and α is the ratio of the real contact area to the nominal contact area.

4.1 Contact between workpiece and tooling

During non-stationary processes, boundary conditions are progressively modified as a result of the interaction between workpiece and tooling. The contact algorithm implemented in both I-Form3 and SORPAS 3D is based on Barata Marques and Martins [1] and requires the workpieces to be discretized by hexahedral elements and tools (treated as rigid) to be discretized by spatial triangular surface elements. This discretization of tool surfaces had originally been proposed by Chenot [2], while Shiau and Kobayashi [3] and Yoon and Yang [4] preferred to describe the tool geometry by Bezier surfaces. However, the choice of a discretization by spatial triangles is somewhat natural in finite element modeling which is already based on discretization procedures.

The resulting contact formulation is based on node-to-triangle contact as illustrated in Fig. 4.1 by a workpiece node contacting a triangular element of the tool. Boundary nodes, like N_p in Fig. 4.1, are analyzed for each triangular surface element of the tool. The orthogonal projection N_{p^*} of node N_p to the plane spanned by the triangle is calculated. Fig. 4.1b shows an example of the orthogonal projection being inside the considered triangular element, which is one of the conditions for being in contact. Fig. 4.1c shows an example of the orthogonal projection lying outside, and hence node N_p and this triangular element are not in contact. The evaluation of whether or not the projection lies inside the triangle is based on a comparison of the total area of the triangle $A_{N_1N_2N_3}$ and the area sum $A_1 + A_2 + A_3$ of the triangles spanned by the projection point and two of the triangle vertices. If the point is inside, the two areas are identical. Another condition for being in contact is that the distance between N_p and N_{p^*} is less than a specified value in order to avoid nodes far from the tool to be considered in contact.

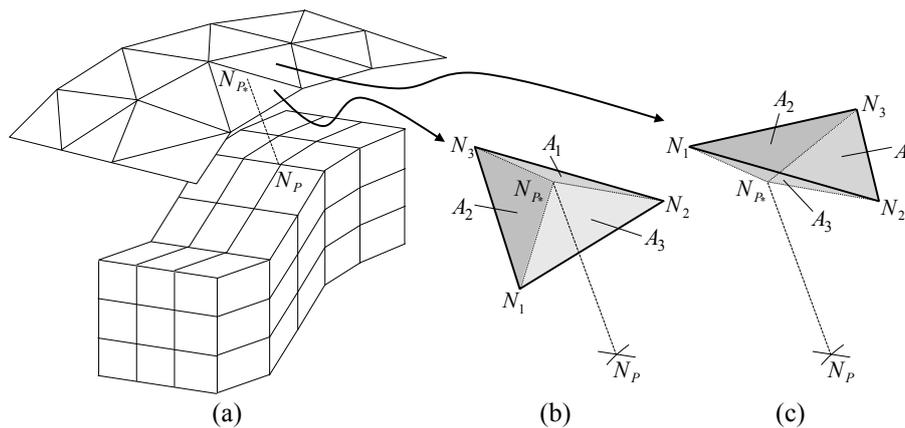


Fig. 4.1. Contact between the hexahedral mesh and the triangular surface mesh of a rigid tool. (a) Node N_p and its projection N_{p^*} in a triangular element, (b) normal projection N_{p^*} of node N_p lying inside triangular surface element and (c) normal projection N_{p^*} of node N_p lying outside triangular surface element.

The time increment necessary for a nodal point to get in contact with the tools is evaluated implicitly ($\theta = 0$) or explicitly ($\theta = 1$) according to

$$\mathbf{x}_{t+\Delta t} = \mathbf{x}_t + [\theta \mathbf{u}_t + (1 - \theta) \mathbf{u}_{t+\Delta t}] \Delta t \quad (4.1)$$

The implemented computer program is calculating the time increment based on the explicit approach, such that the time needed for each of the potential nodes to get in contact with a tool is calculated according to

$$\Delta t_p = \frac{\overline{N_p N_{P^*}}}{v_{N_p} - v_{N_{P^*}}} \quad (4.2)$$

where the denominator is the normal velocity difference between the candidate node N_p and its projection N_{P^*} on the tool, which if it is negative corresponds to an increasing gap and in that case it is discarded as a candidate. Among the candidates, the minimum time Δt_p^{\min} from (4.2) is decisive for the following time increment. If the time step is larger than the minimum time for a contact point to arise, it is split into $\Delta t = \Delta t_p^{\min}$. All points getting in contact to the tools within a specified tolerance in the following step are projected to the tool and assigned boundary conditions to enforce the points to follow the movement of the tool.

Taking the constant friction law as an example, the friction stress $\tau_f = mk$ acts in the opposite direction of the relative velocity \mathbf{u}_r between workpiece material and tool and can therefore be written as

$$\tau_f = -mk \frac{\mathbf{u}_r}{|\mathbf{u}_r|} \quad (4.3)$$

This friction model is illustrated in Fig. 4.2a at the vicinity of a neutral point (no relative velocity). The derivative of the friction stress with respect to the relative velocity is also shown as it is relevant for the finite element implementation, and it is seen that the derivative goes to infinity. To avoid this singularity, Chen and Kobayashi [5] proposed the following approximation,

$$\tau_f \cong -\frac{2}{\pi} mk \cdot \arctan\left(\frac{\mathbf{u}_r}{u_0}\right) \quad (4.4)$$

which resembles the friction stress as shown in Fig. 4.2b when u_0 is a constant much smaller than the magnitude of the relative sliding velocity. The friction contribution to the functional Π (3.8) is

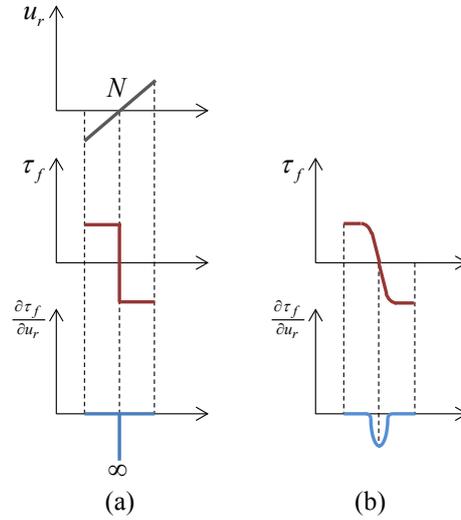


Fig. 4.2. Friction between workpiece and rigid tools. (a) Relative velocity (upper), corresponding friction stress according to the constant friction law (middle) and the derivative of the friction stress with respect to the relative velocity. (b) Modified friction stress according to (4.4) and its derivative.

$$\Pi_f = \int_{S_{tool}} \left[\int_0^{|u_r|} \tau_f d\mathbf{u}_r \right] dS \quad (4.5)$$

The first and second variations of this term are evaluated and added to equations (3.32) and (3.33), thereby entering (3.34) and (3.38) also. The derivatives of (4.5) are integrated by 5×5 Gauss quadrature following Barata Marques and Martins [1].

Once a node is in contact with the tools, it is kept in contact until the normal stress eventually becomes positive, which corresponds to a release of contact. Whenever a node is in contact it is treated mechanically as above, but also thermal and electrical effects may be relevant. The thermal effects are due to heat exchange with the tool (3.20) and friction generated heat (3.21). The electrical boundary conditions are either an applied potential or isolation (isolation is similar to a free surface).

4.2 Contact between deformable objects

An indirect, variational approach is taken to the modeling of contact between deformable objects. A modification to the variation of the functional expressing the total energy-rate of the system is performed by adding a term due to the contact constraints. Traditionally, Lagrange multipliers or the penalty method has been applied. The method of Lagrange multipliers solves the problem exactly, but at the cost of additional unknowns.

The penalty method does not include additional unknowns, but suffers from a compromise in choosing high penalty factors for improving accuracy and low penalty fac-

tors for avoiding ill-conditioned stiffness matrices. Taking advantages from both strategies, the augmented Lagrangian method has become popular; see e.g. Wriggers et al. [6] for an early presentation of the augmented Lagrangian method. This method does however imply longer computation time than the pure penalty method due to iterations involving solution of the main system of equations in order to find the Lagrange multipliers. These iterations do not always converge fast, cf. Zavarise and Wriggers [7] who proposed an improved convergence scheme. Fast convergence is particularly critical for complex finite element computer programs involving non-linearities due to mechanical, thermal and electrical constitutive models. Many solutions assume frictionless or sticking contact, but friction has been included as well. Among the pioneers in frictional modeling are Simo and Laursen [8] using the augmented Lagrangian method.

In relation to resistance welding, Song et al. [9-10] modeled contact in two dimensions by the penalty method. The contact between deformable objects in three dimensions to be presented in this section follows the work of Nielsen et al. [11] and is based on penalties for avoiding penetration of one object into another object or self-penetration of an object. All boundary nodes are analyzed for potential contact to another element face in each simulation step. If a certain node and a corresponding element face are identified as a potential contact pair, a normal gap velocity g_n^c is set up, such that if it is positive, the given velocity field will result in a gap in the contact pair, and if it is negative, the velocity field will result in penetration of the node and the element face. Depending on the mesh and the contact conditions, a node may be a contacting node in one contact pair, and at the same time take part in target surfaces in other contact pairs. This introduces symmetry in the contact algorithm naturally.

Fig. 4.3a shows an example of a node N_p contacting an element face $N_1 - N_2 - N_3 - N_4$ of another element, in this case from another object. Identification of such contact pairs is based on a distance criterion by a small tolerance and that the relative velocity of N_p to the element face is orthogonally projecting to the element face. The definition of a plane is necessary from the element face in order to evaluate the orthogonal projection, but from four nodes, it generally does not exist, since a plane is defined by only three points. Therefore, the quadrilateral surface element is divided into triangles by one of the following two algorithms:

- *Algorithm 1*: The face is divided into two triangles by division through a diagonal as shown in Fig. 4.3b. Doghri et al. [12] experienced loss of symmetry when applying this method. Division by a diagonal leaves two choices, thus resulting in two potential pairs of two triangles, and the problem is which pair to choose. In the present work, both divisions are evaluated, resulting in two overlapping triangles both containing the contact node projection. Among these, the triangle where the projection point results in most equal area coordinates is chosen. Area coordinates are defined as

$$\alpha_j = \frac{A_j}{\sum_{i=1}^3 A_i} \quad (4.6)$$

with areas A_i defined in Fig. 4.1b. This selection of triangle has resulted in better representation of symmetry.

- *Algorithm II*: The face is divided into four triangles by a temporary center node t in the face as shown in Fig. 4.3c. This method was adopted by Doghri et al. [12] to overcome their loss of symmetry with the above method due to the unique choice of triangle. This algorithm is computationally more demanding due to larger expansion of the skyline of the stiffness matrix as the target face is represented by all four nodes compared to three nodes in the above algorithm.

When applying algorithm *I*, the normal gap velocity for contact pair c is defined as

$$\mathbf{g}_n^{c(I)} = (\mathbf{v}_i^p - \alpha_j \mathbf{v}_i^j) \mathbf{n}_i \quad (4.7)$$

where α_j are the area coordinates (4.6), \mathbf{v}_i^j is the velocity of the j 'th node of the se-

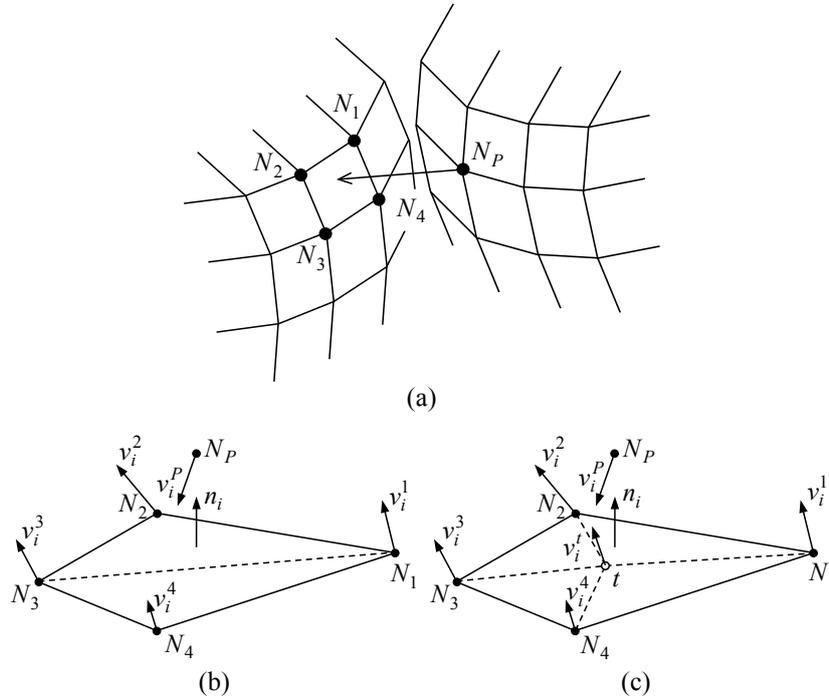


Fig. 4.3. Definition of contact pairs between deformable objects. (a) Node N_p contacting a quadrilateral element face $N_1 - N_2 - N_3 - N_4$ of another element. (b) Division of element face by diagonal. (c) Division of element face by temporary center node t .

lected triangle, and n_i is the normal to the triangle spanned by three of the element face nodes. Note the summation in i and j . Similarly for algorithm *II*, the normal gap velocity becomes

$$\mathbf{g}_n^{c(II)} = \left(v_i^P - \alpha_1 v_i^1 - \alpha_2 v_i^2 - \alpha_t v_i^t \right) n_i \quad (4.8)$$

where index t refers to the temporary center node. Approximation of the velocity in the temporary center node by linear interpolation from the four face nodes, i.e. averaging, leads to

$$\mathbf{g}_n^{c(II)} = \left(v_i^P - \left(\alpha_1 + \frac{\alpha_t}{4} \right) v_i^1 - \left(\alpha_2 + \frac{\alpha_t}{4} \right) v_i^2 - \frac{\alpha_t}{4} v_i^3 - \frac{\alpha_t}{4} v_i^4 \right) n_i \quad (4.9)$$

The normal gap velocity can be written in compact notation to ease subsequent derivations of the variational contribution to the energy rate functional. The following parameters are introduced for algorithms *I* and *II* to assist the compact notation,

$$\alpha_m^I = \begin{cases} 1 & \text{for } m = P \\ -\alpha_m & \text{for } m = 1, 2, 3 \end{cases} \quad (4.10)$$

$$A_I = \text{diag} \{ \alpha_P^I, \alpha_P^I, \alpha_P^I, \alpha_1^I, \alpha_1^I, \alpha_1^I, \alpha_2^I, \alpha_2^I, \alpha_2^I, \alpha_3^I, \alpha_3^I, \alpha_3^I \} \quad (4.11)$$

$$\mathbf{v}_I^T = \{ \mathbf{v}_P, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \}^T \quad (4.12)$$

$$\mathbf{n}_I^T = \{ \mathbf{n}, \mathbf{n}, \mathbf{n}, \mathbf{n} \}^T \quad (4.13)$$

$$\alpha_m^{II} = \begin{cases} 1 & \text{for } m = P \\ -\left(\alpha_m + \frac{\alpha_t}{4} \right) & \text{for } m = 1, 2 \\ -\frac{\alpha_t}{4} & \text{for } m = 3, 4 \end{cases} \quad (4.14)$$

$$A_{II} = \text{diag} \{ \alpha_P^{II}, \alpha_P^{II}, \alpha_P^{II}, \alpha_1^{II}, \alpha_1^{II}, \alpha_1^{II}, \alpha_2^{II}, \alpha_2^{II}, \alpha_2^{II}, \alpha_3^{II}, \alpha_3^{II}, \alpha_3^{II}, \alpha_4^{II}, \alpha_4^{II}, \alpha_4^{II} \} \quad (4.15)$$

$$\mathbf{v}_{II}^T = \{ \mathbf{v}_P, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4 \}^T \quad (4.16)$$

$$\mathbf{n}_{II}^T = \{ \mathbf{n}, \mathbf{n}, \mathbf{n}, \mathbf{n}, \mathbf{n} \}^T \quad (4.17)$$

The velocity gap functions can then be written in the following compact notation for each of the algorithms φ ,

$$\mathbf{g}_n^{c(\varphi)} = \mathbf{v}_\varphi^T \mathbf{A}_\varphi \mathbf{n}_\varphi \quad \varphi = I, II \quad (4.18)$$

which by definition are equivalent to (4.7) and (4.9).

4.2.1 Frictionless contact

According to the definition, action has to be taken only when $g_n^c < 0$ corresponding to penetration in the contact pair. In these cases, the velocity field is constrained by penalizing the penetration, through

$$\delta\Pi_C = \sum_{c=1}^{N_c} P g_n^c \delta g_n^c \quad (4.19)$$

which is to be added to the variation of the energy rate functional (3.11). The total number of contact pairs to be constrained is N_c , and P is a large positive constant. Equation (4.19) handles frictionless contact. In order to handle friction or full sticking, tangential velocity terms should be included.

Evaluation of (4.19) is accomplished for both algorithms by inserting the gap velocity g_n^c , while at the same time replacing \mathbf{v} by $\mathbf{v}_0 + \Delta\mathbf{v}$ resembling the incremental finite element solution when using Newton-Raphson iterations. It is also noted that $\mathbf{g}_n^{c(\varphi)} = \mathbf{n}_\varphi^T \mathbf{A}_\varphi \mathbf{v}_\varphi$ is equivalent to (4.18), since it is simply the transpose of a scalar. Note also that $\mathbf{A}_\varphi^T = \mathbf{A}_\varphi$, since \mathbf{A}_φ is a diagonal matrix. The substitution is shown in the following, where it has been utilized that the variation of the constant \mathbf{v}_0 is zero and $\mathbf{g}_n^c \delta \mathbf{g}_n^c = \delta \mathbf{g}_n^c \mathbf{g}_n^c$,

$$\delta\Pi_C = \sum_{c=1}^{N_c} P \delta \Delta \mathbf{v}_\varphi^T \mathbf{A}_\varphi \mathbf{n}_\varphi \mathbf{n}_\varphi^T \mathbf{A}_\varphi (\mathbf{v}_{0\varphi} + \Delta \mathbf{v}_\varphi), \quad \varphi = I, II \quad (4.20)$$

Utilizing that $\delta \Delta \mathbf{v}_\varphi$ is to be chosen arbitrarily, it is possible to recognize the contributions to the stiffness matrix and the load vector after rearranging terms,

$$\delta\Pi_C = \sum_{c=1}^{N_c} \left[\delta \Delta \mathbf{v}_\varphi^T \left(\underbrace{P \mathbf{A}_\varphi \mathbf{n}_\varphi \mathbf{n}_\varphi^T \mathbf{A}_\varphi}_{\equiv \mathbf{K}_c} \Delta \mathbf{v}_\varphi + \underbrace{P \mathbf{A}_\varphi \mathbf{n}_\varphi \mathbf{n}_\varphi^T \mathbf{A}_\varphi \mathbf{v}_{0\varphi}}_{\equiv -\mathbf{f}_c} \right) \right], \quad \varphi = I, II \quad (4.21)$$

The contribution from the c 'th contact pair to the stiffness matrix is \mathbf{K}_c , and the corresponding contribution to the load vector is \mathbf{f}_c . For algorithm *I*, the dimensions will be 12×12 for \mathbf{K}_c and 12×1 for \mathbf{f}_c , whereas for algorithm *II*, the dimensions will be 15×15 and 15×1 , respectively.

Regarding the assembly, an overview is best given by writing the contributions to the stiffness matrix and the load vector in the following forms, where for the load vector it is recognized that the initial gap velocity is $\mathbf{g}_{n0}^c = \mathbf{n}_\varphi^T \mathbf{A}_\varphi \mathbf{v}_{0\varphi}$,

$$K_c^{ijmn} = P \alpha_m^\varphi \alpha_n^\varphi n_i n_j, \quad i, j = 1, 2, 3$$

$$m, n = \begin{cases} P, 1, 2, 3 & \text{for } \varphi = I \\ P, 1, 2, 3, 4 & \text{for } \varphi = II \end{cases} \quad (4.22)$$

$$f_c^{jm} = \begin{cases} -P \alpha_m^\varphi n_j g_{n0}^c, & \text{for Newton - Raphson iterations} \\ 0, & \text{for direct iterations} \end{cases} \quad (4.23)$$

$$m, n = \begin{cases} P, 1, 2, 3 & \text{for } \varphi = I \\ P, 1, 2, 3, 4 & \text{for } \varphi = II \end{cases}$$

When using direct iterations, \mathbf{v} is solved directly, rather than the incremental velocity $\Delta \mathbf{v}$ when using Newton-Raphson iterations. For direct iterations it follows (similar to (4.20) and (4.21)) that \mathbf{K}_c is identical, but $\mathbf{f}_c = \mathbf{0}$.

The factor P is the penalty, α_m^φ is given by either (4.10) or (4.14), and n_i is the unit normal vector to the contact face. The position of each of the components, K_c^{ijmn} and f_c^{jm} , in the global system of equations is shown by Fig. 4.4, where blocks of 3×3 positions, ij , are identified as the relation between nodal points m and n . Each block is symmetric ($ij = ji$) and the blocks are symmetric around the diagonal ($mn = nm$). In cases ($mn = \{2P, 23, (43)\}$) in Fig. 4.4) where the penalty blocks lie above the skyline, the skyline profile has to be expanded to allow the additional penalty blocks.

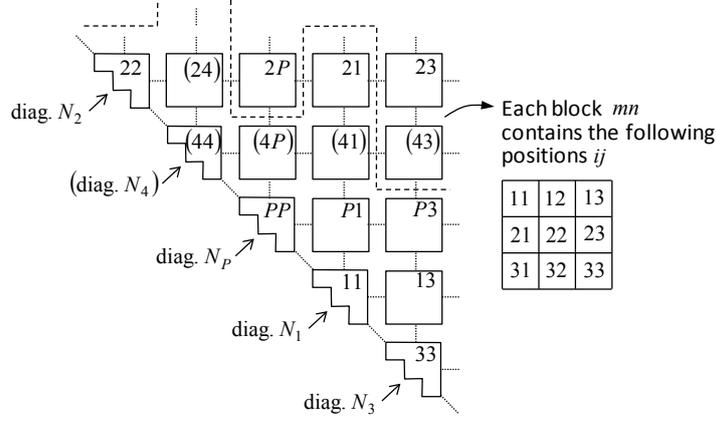


Fig. 4.4. Position of penalty terms in global system of equations with random order of the numbers $P, 1, 2, 3, (4)$ as the position depends on their relative node numbers. All positions related to 4 are in parentheses as they are only active for contact algorithm $\varphi = II$.

4.2.2 Sticking contact

In sticking contact there is no sliding between the surfaces in contact. The tangential velocity difference is therefore penalized in addition to the normal gap velocity. The variational penalty term stemming from the tangential velocity difference g_t^c is given by

$$\delta \Pi_C^t = \sum_{c=1}^{N_c} P g_t^c \delta g_t^c \quad (4.24)$$

which is similar to (4.19). The derivations are also identical except that the tangential contributions result in two sets of penalty terms corresponding to the two tangential components of the tangential velocity difference written as

$$g_{t1}^{c(\varphi)} = \mathbf{v}_\varphi^T \mathbf{A}_\varphi \mathbf{t}_{\varphi 1}, \quad g_{t2}^{c(\varphi)} = \mathbf{v}_\varphi^T \mathbf{A}_\varphi \mathbf{t}_{\varphi 2}, \quad \varphi = I, II \quad (4.25)$$

with notations following (4.10)-(4.17) and $\mathbf{t}_{\varphi 1}$ and $\mathbf{t}_{\varphi 2}$ being vectors of the two tangential unit vectors. The resulting terms after insertion into (4.24) are similar to (4.22)-(4.23) with the normal vector exchanged by each of the tangential vectors.

4.2.3 Frictional contact

As for the contact between workpiece and rigid tools, the constant friction law, $\tau_f = mk$, will be taken as an example. To avoid the derivatives going to infinity cf. the discussion related to Fig. 4.2, the friction stress is written as (4.4), here with the tangential velocity difference defined by (4.25),

$$\tau_f \cong -\frac{2}{\pi} mk \cdot \arctan\left(\frac{\mathbf{g}_t^{c(\varphi)}}{u_0}\right) \quad (4.26)$$

The contribution to the energy rate functional due to friction and its corresponding variation are

$$\Pi_f = \sum_{c=1}^{N_c} \tau_f A_c \mathbf{g}_t^c \quad (4.27)$$

$$\delta\Pi_f = \sum_{c=1}^{N_c} [\delta\tau_f A_c \mathbf{g}_t^c + \tau_f A_c \delta\mathbf{g}_t^c] \quad (4.28)$$

where frictional force is introduced by the product of the frictional stress τ_f and the area of the contact pair A_c . Insertion of the friction stress and the tangential velocity difference components (4.25) into the variational form (4.28) results in the following additional terms to the stiffness matrix K_f^{ijmn} and generalized load vector f_f^{jm} for each of the tangential components (one set for each of the inserted tangential unit vectors),

$$K_f^{ijmn} = -\frac{2}{\pi} mk \frac{u_0}{(\mathbf{g}_t^c)^2 + u_0^2} \alpha_m^\varphi \alpha_n^\varphi t_i t_j, \quad i, j = 1, 2, 3 \quad (4.29)$$

$$m, n = \begin{cases} P, 1, 2, 3 & \text{for } \varphi = I \\ P, 1, 2, 3, 4 & \text{for } \varphi = II \end{cases}$$

$$f_f^{jm} = \begin{cases} -\frac{2}{\pi} mk A_c \left(\frac{u_0 \mathbf{g}_{t0}^c}{(\mathbf{g}_t^c)^2 + u_0^2} + \arctan\left(\frac{\mathbf{g}_t^c}{u_0}\right) \right) \alpha_m^\varphi t_j, & \text{for Newton - Raphson iterations} \\ -\frac{2}{\pi} mk A_c \arctan\left(\frac{\mathbf{g}_t^c}{u_0}\right) \alpha_m^\varphi t_j, & \text{for direct iterations} \end{cases} \quad (4.30)$$

$$m = \begin{cases} P, 1, 2, 3 & \text{for } \varphi = I \\ P, 1, 2, 3, 4 & \text{for } \varphi = II \end{cases}$$

4.2.4 Electrical and thermal contact

Electrical and thermal contact properties are included in contact interface elements on one or both of the objects in contact and eventual drops over the interface due to contact resistances are included in these elements, see Section 4.3 for a description of the physical properties. The contact implementation here is therefore limited to ensure that the electrical potential and the temperature are identical on both sides of the contacting finite elements. This is ensured by penalizing electrical potential difference Φ_d and temperature difference T_d by

$$\delta\Pi_{\Phi} = \sum_{c=1}^{N_c} P\Phi_d^c \delta\Phi_d^c \quad \delta\Pi_T = \sum_{c=1}^{N_c} PT_d^c \delta T_d^c \quad (4.31)$$

Both the potential and the temperature are scalar fields, and the derivation is therefore a reduced form of the frictionless contact derivation in the absence of the normal vector. The contributions to the system matrices are

$$K_{\Phi,T}^{mn} = P\alpha_m^{\varphi}\alpha_n^{\varphi}, \quad m,n = \begin{cases} P,1,2,3 & \text{for } \varphi = I \\ P,1,2,3,4 & \text{for } \varphi = II \end{cases} \quad (4.32)$$

while there are no contribution to the right hand sides.

4.3 Physical contact properties

In the above electrical and thermal contact models, the numerical procedures ensure that the electrical potential and the temperature are the same on the two surfaces of finite element meshes in contact, cf. the penalty terms in (4.31). These are pure numerical constraints not taking into account the electrical and thermal contact resistances stemming from the actual surface properties. A thin layer of interface elements is introduced on one or both of the contacting surfaces to include the physical properties of the surface in terms of electrical and thermal responses. Fig 4.5 shows the local mesh of a resistance spot welding case with interface layers between contacting objects, in this case between an electrode and a sheet and between two sheets.

4.3.1 Electrical contact resistance

Taking the electrical contact resistance as a starting point, Fig. 4.6 shows the two un-

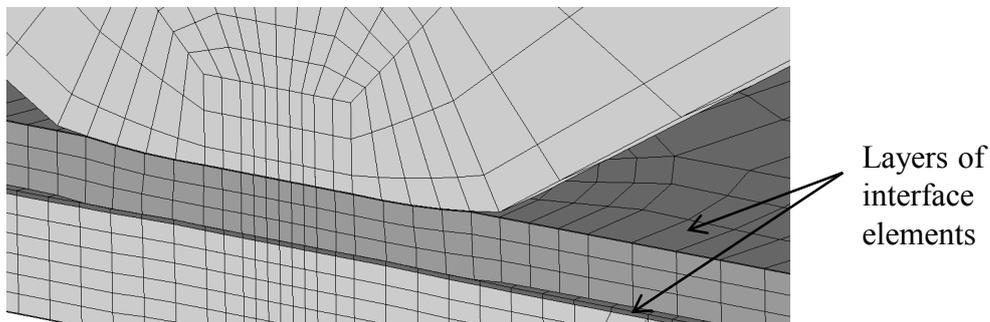


Fig. 4.5. Example of finite element mesh of different objects with interface layers of elements to simulate the physical contact properties. Three objects are visible; an electrode (upper) and two sheets (middle and lower). Interface layers are added between the sheets and between the upper sheet and the electrode.

derlying contributions to the increase of electrical resistance over a contact interface. When two surfaces contact each other, only a fraction of the apparent area will be in real contact due to surface roughness. The contact area develops between the surface asperities until the load can be carried. Hence, the fraction of the area being in real contact increases with contact pressure as the surface asperities deform and it depends on the strength of the materials in contact.

Fig. 4.6a, after Timsit [13], illustrates current flow across an interface between two surfaces. The current can only pass from one surface to the other through the contact spots in real contact. This restricts the current and increases the resistance. Fig. 4.6b shows a close up of a contact spot, where the second source of increased resistance is illustrated by a layer of contaminants on the surfaces. These layers can stem from oxide layers, surface films, grease and dirt and introduce additional resistance due to typical higher resistivity of these layers.

The contact resistance is highly dynamic during resistance welding and is among the more complex and important factors for the process. The level of the contact resistance can vary a lot between material batches and process variations due to surface conditions and natural scatter. The current restriction due to the real contact area being only a fraction of the real contact area varies with material strength, prior processing due to local strain hardening as well as with the first touching of the welding electrodes because of the different normal pressures imposed by soft touching or impacts. The contribution from the contaminant layer naturally varies because of different oxide layers and surface treatments in terms of eventual degreasing or cleaning. Variation within batches or even single welds also occurs due to variations in the contaminant layer both in terms of thickness and composition.

The dynamic behavior of the contact resistance stems from the changing contact pressure and temperature during the resistance welding process. The contact pressure directly influences the real contact area and hence the restriction of the current. The contact pressure also influences the contaminant layer due to eventual break down of oxide layers and squeeze out of surface films and grease.

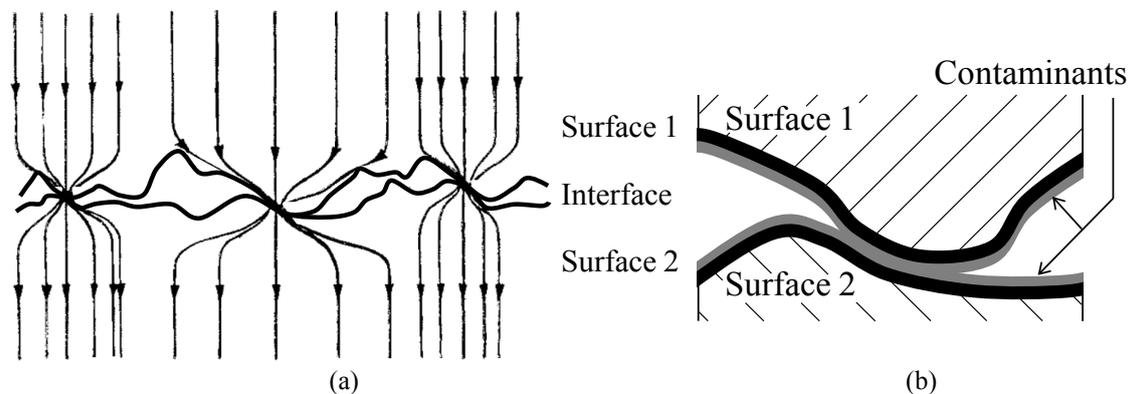


Fig. 4.6. Origins of electrical contact resistance by (a) restriction of the current flow due to the real contact area being only a fraction of the apparent area and by (b) surface contaminants such as oxide layers, surface films, grease and dirt (illustrated by the grey layers).

Temperature also affects the contact resistivity. When the temperature increases, the material softens and an increase in the real contact area follows to carry the applied force. Furthermore, the bulk resistivity changes with temperature, which affects the contact resistivity due to the longer path resulting from constriction. The properties of the contaminant layer also changes with temperature both in terms of resistivity and level of difficulty for the squeeze out.

Following the theory by Bowden and Tabor [14], the ratio of the real contact area to the apparent area is given by

$$\alpha = \frac{\sigma_n}{3\sigma_{soft}} \quad (4.33)$$

where σ_{soft} is the flow stress of the softer material in contact and σ_n is the contact normal pressure. In the contact resistance model in SORPAS (see e.g. Zhang [15]), the electrical conductance over the interface is assumed proportional to this area fraction, and hence the contact resistance is inversely proportional to (4.33) applying the theory by Bowden and Tabor [14]. The model for the electrical contact resistance takes the following form,

$$\rho_c = \frac{3\sigma_{soft}}{\sigma_n} \left(\frac{\rho_1 + \rho_2}{2} + \rho_{contaminants} \right) \quad (4.34)$$

where ρ_1 and ρ_2 are the electrical bulk resistivities of the two contacting materials and $\rho_{contaminants}$ is the resistivity stemming from surface contaminants such as oxides, surface films and dirt.

Alternative models

The electrical contact resistance has drawn continuous attention for at least a century, but due to the complexity imposed by the various conditions and the process dependency, a generic model describing all experiments was never proposed. The implemented model (4.34) is therefore only one option among several proposed models in literature, and it is naturally chosen due to the affinity to the 2D version of SORPAS, which has been using this model successfully through more than a decade.

The alternative models can be divided into two groups stemming from either fitting of experimental data or theoretical modeling of the surfaces. Both groups suffer from the fact that a change in material, surface conditions or process parameters will make the actual response differ from the model.

Studer [16] tested and discussed the electrical contact resistance in relation to spot welding back in 1939 and provides in the discussion reported models for the contact resistance based on power laws fitted to experimental curves. Some data sets are fitted by $R = Cp^{-1}$ or by $R = Cp^{-1/2}$, while yet other data sets are not fitted by any power

law. In these models, R is the resistance, p is the contact pressure and C is a constant. The implemented model (4.34) is a variant of $R = Cp^{-1}$ where material properties are introduced to replace the constant C , though leaving the contaminant resistivity $\rho_{contaminants}$ as a variable to resemble the actual surface conditions. The empirical models can be combined into a single relation, $R = Cp^{-\beta}$, where β is the dependence of the pressure, e.g. $\beta = 1$ and $\beta = 1/2$ as above. Many other contributions present empirical investigations of the electrical contact resistance; see e.g. Song et al. [17] and their references.

Other studies are based on theoretical analysis and assumptions. They typically take the starting point in the work by Holm [18] and Greenwood [19], where the constriction resistance can be calculated by assuming the real contact area is either a single circular spot or a cluster of uniformly distributed spots. A collection of later contributions based on these assumptions are presented by Friis [20] in terms of references and models. An overall comment to those models is that none of them fits all data sets and that parameters have to be fitted for each data set.

4.3.2 Thermal contact resistance

The contact interface presents resistance to heat transfer as it does to the electric current flow. The effect to the welding process is considered less because of limited temperature gradients compared to direct importance of the electrical contact resistance to the heat generation. The major influence of the thermal contact resistance is in interfaces where the temperature is not mainly generated in the interface but conducted to and through the interface. An example is spot welding of three sheets, where in some cases the heat is generated between two thicker, high strength steels while the heat generation is limited between the third thin, low carbon steel sheet and one of the two high strength steels. In this case the heat transfer towards the thin sheet (and through the sheet to the colder electrode) plays a major role and the thermal contact resistance may be determining whether the weld nugget forms into the thin sheet.

The implemented model for the thermal contact resistance in SORPAS is based on a similar concept as (4.34). With thermal resistance, κ , the expression becomes

$$\kappa_c = \frac{3\sigma_{soft}}{\sigma_n} \left(\frac{\kappa_1 + \kappa_2}{2} \right) \quad (4.35)$$

The exception when comparing to (4.34) is that a term corresponding to the contaminant layer is not included. Such a term could have been included as well as the opposite working radiation could have been included.

References

- [1] Barata Marques MJM, Martins PAF (1990) Three-dimensional finite element contact algorithm for metal forming. *International Journal for Numerical Methods in Engineering* 30:1341-1354.
- [2] Chenot JL (1987) Finite element calculation of unilateral contact with friction in non-steady-state processes. *Proceedings of the NUMETA Conference*, Swansea, United Kingdom.
- [3] Shiau YC, Kobayashi S (1988) Three dimensional finite element analysis of open die forging. *International Journal for Numerical Methods in Engineering* 25:67-86.
- [4] Yoon JH, Yang DY (1988) Rigid plastic finite element analysis of three dimensional forging by considering friction on continuous curved dies with initial guess generation. *International Journal of Mechanical Sciences* 30:887-898.
- [5] Chen C, Kobayashi S (1978) Rigid plastic finite element analysis of ring compression. *Application of Numerical Methods in Forming Processes*, AMD 28, ASME, New York, pp. 163-174.
- [6] Wriggers P, Simo JC, Taylor RL (1985) Penalty and augmented Lagrangian formulations for contact problems. *Proceedings of the NUMETA Conference*, 97-105.
- [7] Zavarise G, Wriggers P (1999) A superlinear convergent augmented Lagrangian procedure for contact problems. *Engineering Computations* 16:88-119.
- [8] Simo JC, Laursen TA (1992) An augmented Lagrangian treatment of contact problems involving friction. *Computers and Structures* 42:97-116.
- [9] Song Q, Zhang W, Bay N (2006) Contact modelling in resistance welding. Part 1: Algorithms and numerical verification. *Journal of Engineering Manufacture* 220(5):599-606.
- [10] Song Q, Zhang W, Bay N (2006) Contact modelling in resistance welding. Part 2: Experimental validation. *Journal of Engineering Manufacture* 220(5):607-613.
- [11] Nielsen CV, Martins PAF, Zhang W, Bay N (2011) Mechanical contact experiments and simulations. *Steel Research International* 82:645-650.
- [12] Doghri I, Muller A, Taylor RL (1998) A general three-dimensional contact procedure for implicit finite element codes. *Engineering Computations* 15:233-259.
- [13] Timsit RS (1999) Electrical contact resistance: Properties of stationary interfaces. *IEEE Transactions on Components and Packaging Technology* 22(1):85-98.
- [14] Bowden FP, Tabor D (1950) *The Fabrication and Lubrication of Solids*. Oxford University Press.
- [15] Zhang W (2003) Design and implementation of software for resistance welding process simulations. *Transactions: Journal of Materials and Manufacturing* 112(5):556-564.
- [16] Studer FJ (1939) Contact resistance in spot welding. *Welding Journal* 18(10):374-380.
- [17] Song Q, Zhang W, Bay N (2005) An experimental study determines the electrical contact resistance in resistance welding. *Welding Journal* 84(5):73s-76s.
- [18] Holm R (1967) *Electric Contacts*. Springer-Verlag, Berlin/Heidelberg, Germany.
- [19] Greenwood JA (1966) Constriction resistance and the real area of contact. *British Journal of Applied Physics* 17(12):1621-1632.
- [20] Friis KS (2010) Resistance welding of advanced materials and micro components. PhD Thesis, Technical University of Denmark.

5. Meshing and Remeshing

A significant amount of time in finite element modeling of manufacturing processes is spent in mesh generation. Setting up three-dimensional meshes is a cumbersome task due to complexity of the processes and the involved geometries. Moreover, additional meshing challenges often appear due to the fact that manufacturing processes based on large plastic deformations present progressive mesh distortion (or degeneracy), potential interference between mesh and contour of the tools and possible contact of the mesh with itself. This poses the need for robust, automatic, mesh generation and regeneration (remeshing) procedures in order to ensure that complex processes are modeled from the beginning to the end with high levels of accuracy both in terms of geometry and distribution of field variables.

The choice of element type has large impact on the simulations, and the typical dilemma in three dimensions arises from the selection between tetrahedral and hexahedral elements. The arguments for the tetrahedral elements are the robustness, versatility and availability of meshing algorithms. Based on Delaunay tessellation, Coupez et al. [1] opened the possibility of effectively and automatically simulating the whole forming process of complex three-dimensional parts from beginning to the end. On the other hand, the argument for the hexahedral elements is the accuracy. Furthermore, standard tetrahedral elements suffer from locking due to the incompressibility constraint in plasticity. Second-order tetrahedral elements overcome this problem but perform poorly in the tool-workpiece contact interfaces, often leading to stability problems in the contact algorithms as stated by Tekkaya and Martins [2]. As a result of this, special tetrahedral elements with interior nodes have been developed for preventing locking. These elements, however, still suffer from some of the typical drawbacks of tetrahedral elements: They are overly stiff, very sensitive to mesh orientation and frequently require up to an order of magnitude more elements to achieve the same level of accuracy as hexahedral elements. Benzley et al. [3] also noticed that meshes based on tetrahedral elements result in larger models, and therefore in larger computational requirements, than meshes based on hexahedra for the same level of accuracy. Kraft [4] observed that tetrahedral elements cause critical errors when distorted, whereas hexahedra have better behavior even when distorted. Hexahedral elements are used in the present work and are therefore dealt with in this chapter.

Meshes based on hexahedral elements can be divided into two groups. One group is structured meshes, which can be recognized by all interior nodes of the mesh having equal number of adjacent elements. The simplest geometries are easily meshed and the

more complicated can be handled by isoparametric meshing of superelements as described in Section 5.2. The utilization of this method is limited to geometries that can be divided into hexahedral superelements. The second group is unstructured meshes, which, in principle, should cover all three-dimensional geometries.

The simplest unstructured meshing by hexahedral elements is performed by means of an indirect approach, where the geometry is first meshed by tetrahedral elements using Delaunay tessellation. Each tetrahedron is subsequently decomposed into four hexahedral elements. This approach is robust but always leads to distorted elements with only a fraction of the quality of an ideal hexahedron. Furthermore, the indirect meshing by decomposition always leads to nodal points with high valence, which artificially increases the overall stiffness of the finite element models. The poor quality obtained by this approach is considered the reason why some well-known commercial finite element programs currently utilized in metal forming do not offer hexahedral elements as an option, or do not provide automatic remeshing if hexahedral elements are available. The alternative approach for the automatic generation of good quality hexahedral elements in arbitrary domains was originally proposed by Schneiders and Bünten [5] and will hereafter be referred to as “all-hexahedral meshing”.

This chapter is divided into sections as follows. Because description of tooling is relevant for meshing and remeshing, a brief review of the techniques that are utilized for the description of tool surfaces is given in Section 5.1. Section 5.2 presents structured meshing based on superelements and Section 5.3 presents all-hexahedral meshing. To conclude the meshing techniques, Section 5.4 presents the strategies for adding interface layers as those shown in Fig. 4.5 to include the physical properties of contacting surfaces being flat or curved. Finally, all-hexahedral remeshing is presented in Section 5.5.

5.1 Description of tooling

Tools can be described by analytical or parametric surfaces, surface meshes and clouds of points, Santos and Makinouchi [6]. In most of the commercial finite element computer programs, the surfaces are described by means of surface meshes (e.g. triangular elements, Fig. 5.1a). The utilization of a grid of triangular elements instead of alternative approaches based on analytical functions, parametric surfaces (Fig. 5.1b) or clouds of points, is due to the fact that the former always guarantees successful discretization of the surfaces while other techniques often face difficulties whenever complex shapes and/or small geometrical details are to be discretized.

However, triangular elements fail to ensure smoothness and, therefore, introduce artificial roughness on the surface of tooling. This can bring in geometrical errors, for instance in case of small fillet radii which may be poorly captured if the discretization is too coarse.

Tools deform elastically but are commonly modeled as rigid in three-dimensional finite element modeling of metalworking processes because their deformations are

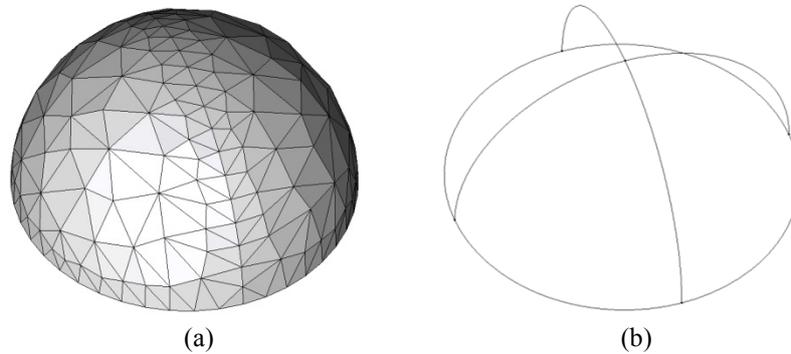


Fig. 5.1. Two main approaches utilized in the discretization of a hemispherical tool: (a) surface meshes and (b) analytical functions or parametric surfaces.

negligible when compared to the plastic deformation of the workpieces. However, the simplification is not always feasible as shown by Tekkaya and Martins [2] by means of a metal forming example displaying significant erroneous tool force when assuming the tool rigid instead of elastic. In such situations there is a need to take the elastic deformation of tooling into consideration. When it comes to resistance welding applications, the tools, in terms of boundary conditions, act as the coupling between electrodes and the welding machine and are therefore sufficiently modeled as rigid.

The majority of the applications reported in the literature that deal with the elastic deformation of tools is restricted to the utilization of finite elements both in the workpiece material and tools, Boussetta et al. [7] and Behrens and Kerkeling [8]. This results in limitations in terms of the size and complexity of the overall computer models when the tools, having complex geometrical shapes, are to be discretized and included in the overall set of finite-element computations. Some of these limitations can be solved by alternative approaches based on combination of finite element and boundary element methods; see Fernandes et al. [9].

The utilization of boundary elements for performing the elastic deformation of the dies not only avoids over-sizing the resulting computer models as it offers significant computational advantages over the existing approaches fully based on finite elements. The first advantage is due to the fact that boundary elements only require discretization of the die surfaces. The second advantage is seen by taking into consideration that numerical simulation of manufacturing processes is generally accomplished through a succession of displacement increments, each modeling a small percentage of the initial height of the preform. In practical terms, this means that a simulation based on several hundreds of increments will require the elastic deformation of the dies to be also calculated hundreds of times. This is the reason why alternative approaches based on boundary elements make a difference against fully finite element based solutions. Similar finite element – boundary element combined approaches can also be utilized for solving thermo-mechanical coupling in the tool-workpiece interface, as shown by Rodrigues and Martins [10].

5.2 Isoparametric structured meshing

Structured meshes of hexahedral elements can be created by a method based on isoparametric meshing of superelements as first shown by Zienkiewicz and Phillips [11]. Martins and Barata Marques [12] developed a three-dimensional mesh generator based on this technique and published the source code, which is adopted in the present work.

The method is applicable when the geometry to be meshed can be divided into a number of sub-blocks, the so-called superelements. An example is shown in Fig. 5.2a in terms of a quarter of an electrode for spot welding. The top face shows a typical division of solid cylindrical faces in order to achieve well shaped superelements resulting in well-shaped 8-node hexahedral elements.

The superelements are 20-node elements specified by the user by the coordinates of the eight corner points and 12 mid-side points. The mid-side points are automatically placed half distance on the straight line between two corner points if not specified. Otherwise, the edges of the superelements are represented parabolic by the mid-side nodes and their two respective corner points. Any point within the superelements is given by interpolation using the standard shape functions for a 20-node hexahedral element. Division of the superelements into 8-node elements is based on specified number of divisions along each superelement side along with a corresponding grading of the element division. Note that divisions should only be specified for the three mutual orthogonal (with respect to natural coordinates) directions due to the structured nature of the created meshes. Fig. 5.2b-c show an example of the resulting mesh.

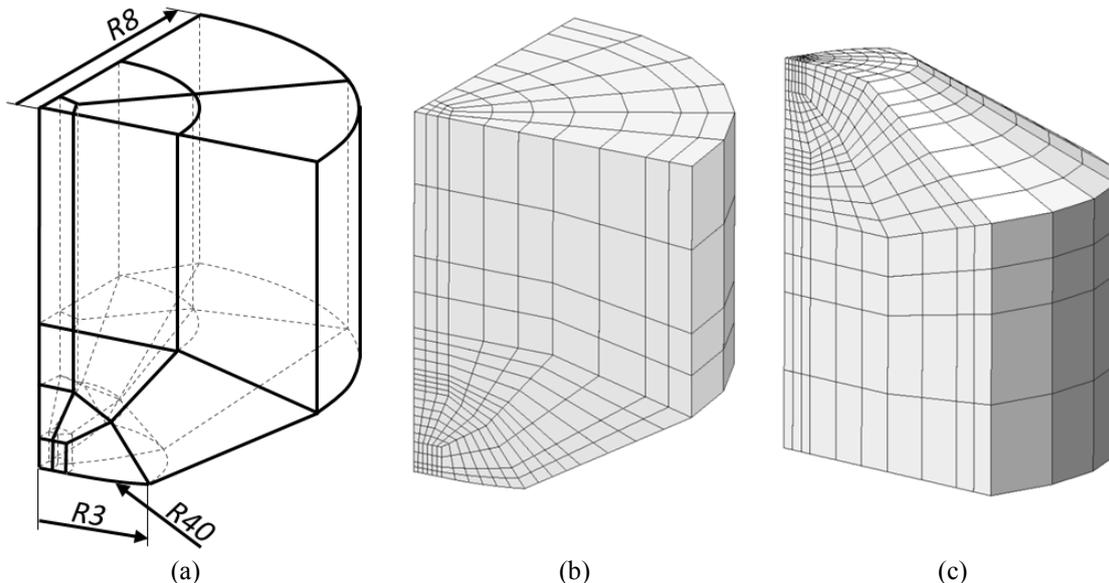


Fig. 5.2. Isoparametric structured meshing of a quarter of an ISO type B0 electrode for spot welding. (a) Subdivision of geometry into 20-node superelements. (b-c) Subsequent automatic division of superelements into 8-node hexahedral elements.

In combination with a user interface for setting up the superelements, this meshing technique is a powerful tool in setting up initial geometries and meshes. The method is not automatically applicable in remeshing procedures because the underlying geometries of the meshes with need for remeshing usually cannot be identified by a reasonable number of superelements. Remeshing is therefore solely accomplished by meshing techniques based on unstructured meshing.

5.3 All-hexahedral unstructured meshing

All-hexahedral meshing is a grid based approach that involves the construction of a structured three-dimensional mesh of hexahedra in the interior of the volume (core mesh) followed by subsequent generation of an extra layer of elements for linking the core with its projection on the boundary of the workpiece. The method proposed by Schneiders and Bünthen [5] is an extension of the two-dimensional approach based on quadrilateral elements that was previously developed by Schneiders et al. [13]. Among other contributors to the all-hexahedral meshing techniques are e.g. Kraft [4], Zhu and Gotoh [14], Karadogan and Tekkaya [15] and Kwak and Im [16].

The all-hexahedral meshing algorithm to be presented in what follows was originally developed by Fernandes and Martins [17], who provided a detailed description of the major procedures and programming solutions. The procedures are further developed in the present work (see also Nielsen et al. [18]) by introducing adaptive core meshes and the possibility of handling multiple objects besides enhancing the overall robustness and versatility.

5.3.1 Identification of geometric features

The starting point of the meshing procedure is a triangular surface mesh of the geometry, e.g. provided by a CAD program, and the all-hexahedral meshing procedure is then responsible for supplying a hexahedral mesh within the surface. An important step before the meshing itself is the recognition of geometrical features in form of vertices and edges that must be kept during meshing. Fig. 5.3a shows a triangular surface mesh and Fig. 5.3b shows the geometrical features that were identified after applying algebraic algorithms based on the evaluation of surface normals to the triangles and analyzing nodal valences. As illustrated in Fig. 5.3a, a typical segment 1–2 shared by two adjacent triangular elements '123' and '142' is taken as an edge segment if the angle α between the normals \mathbf{n}_{123} and \mathbf{n}_{142} to the triangular elements is greater than a specified threshold angle (say $\theta = 45^\circ$),

$$\mathbf{n}_{123} \cdot \mathbf{n}_{142} > \cos \theta \quad (5.1)$$

The summation and sorting of adjacent edge segments before and after 1–2 in a sequential manner leads to the edge $A-B$ (Fig. 5.3b).

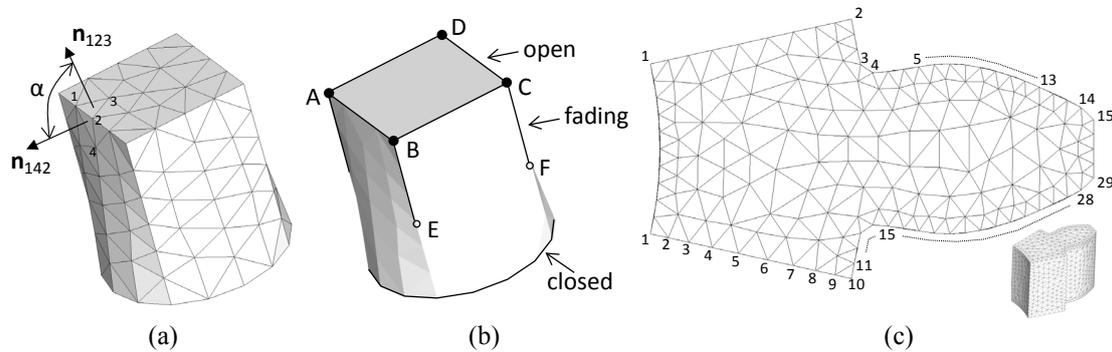


Fig. 5.3. Identification of geometrical features. (a) Identification of edge segments from triangular surface mesh. (b) Identification of edges and vertices. (c) Merging of collinear edge segments.

Vertices are collected from the end points of edge segments that are connected to at least three neighboring edge segments. Edges are classified into three main groups (Fig. 5.3b): (i) open edges, (ii) closed edges and (iii) fading edges. Open edges connect two different vertices ($A-B$, $B-C$, $C-D$ and $D-A$), closed edges start and end in the same point and do not contain vertices, and fading edges start in a vertex but smoothly vanish along the surface (e.g. $B-E$ and $C-F$).

In order to simplify the overall size of the point-edge array utilized for collecting the edge segments in a sequential manner, consecutive and collinear edge segments (with a threshold tolerance of say 2°) are merged. In Fig. 5.3b, this implies that a typical edge $A-B$ consisting of four different edge segments will be represented by a single edge segment connecting end points A and B , which are vertices. Fig. 5.3c explains the procedure in more detail showing the representation of two identical edges (1-10 and 1-2 before and after being merged). As a result of this procedure the upper edges in Fig. 5.3c are made of 15 different points (14 edge segments) whereas the lower edges, which are not merged, consist of 29 different points (28 edge segments).

The benefits of merging collinear edge segments go beyond the aforementioned reduction in the size of point-edge arrays because it considerably enlarges the choice of candidate nodes of the hexahedral mesh to be projected to edges of the surface triangular mesh. The selection of candidate nodes is discussed in Section 5.3.4.

5.3.2 Adaptive core mesh and reconstruction of surfaces

After recognition and organization of the geometrical features, the first step of the meshing procedure is the generation of the core mesh of hexahedral elements. The algorithm for generating the core of all-hexahedral finite element meshes is commonly made of four major procedures [5, 17]: (i) identification of the bounding box, (ii) construction of a hexahedral mesh within the bounding box, (iii) removal of elements placed outside the workpiece to be discretized, and (iv) topology list packing for the remaining elements.

However, in what concerns identification of the bounding box, published work in the field is usually limited to the utilization of cuboids that circumscribe the geometry.

The ‘identification’ is, therefore, limited to manual or automatic selection of the size and number of hexahedra along three concurrent edges (Fig. 5.4a). An exception to this is the work of Kwak and Im [16] suggesting the utilization of wedge type bounding boxes for generating core meshes in applications with angular symmetries. Octree based core meshes as used by Kraft [4] presents another alternative, which has its strength by the possibility of controlling mesh density.

The present work (see also Nielsen et al. [18]) introduces the concept of adaptive selection between cuboids and circular cylindrical or tubular bounding boxes as a function of dominant displaying geometrical features of the workpiece. Core meshes derived from cuboids are generated by standard grid based procedures [5] and those resulting from rotationally or near rotationally symmetric geometries are generated by means of isoparametric based procedures as presented in Section 5.2.

Fig. 5.4 shows three types of bounding boxes that may be utilized for generating adaptive core meshes. The bounding boxes and the corresponding core meshes are free to take arbitrary positions along x-y-z directions and the overall concept is not limited to the basic elementary geometries that are included in the figure. Other shapes resulting from additional elementary geometries or from the assembly of elementary geometries may be employed for better adapting the core mesh to a specific geometry being meshed or remeshed. The procedures outlined in Section 5.2 can potentially be utilized for creating any shapes of bounding boxes to surround the actual geometry.

The influence of the core mesh on the overall quality of the final hexahedral mesh has proven to be significant. In particular, as will be seen later in the presentation, they avoid or minimize the need for mesh repairment. Taking the case shown in Fig. 5.5 as an example, it follows that all-hexahedral meshing starting from a circular cylindrical core mesh produces a much better final mesh than that obtained from a cuboid core mesh. The left sides of Fig. 5.5b-c show the core meshes and the right sides show the final meshes.

The remaining procedures for generating the core mesh; namely the removal of elements placed outside the workpiece and the topology list packing of the remaining elements are comprehensively explained by Fernandes and Martins [17]. In particular,

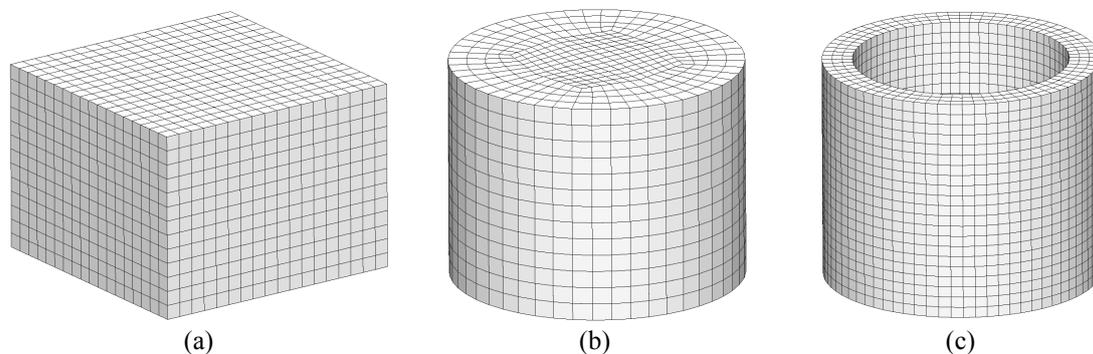


Fig. 5.4. Adaptive bounding boxes for domains displaying dominant (a) cuboid, (b) circular cylindrical or (c) tubular geometrical features.

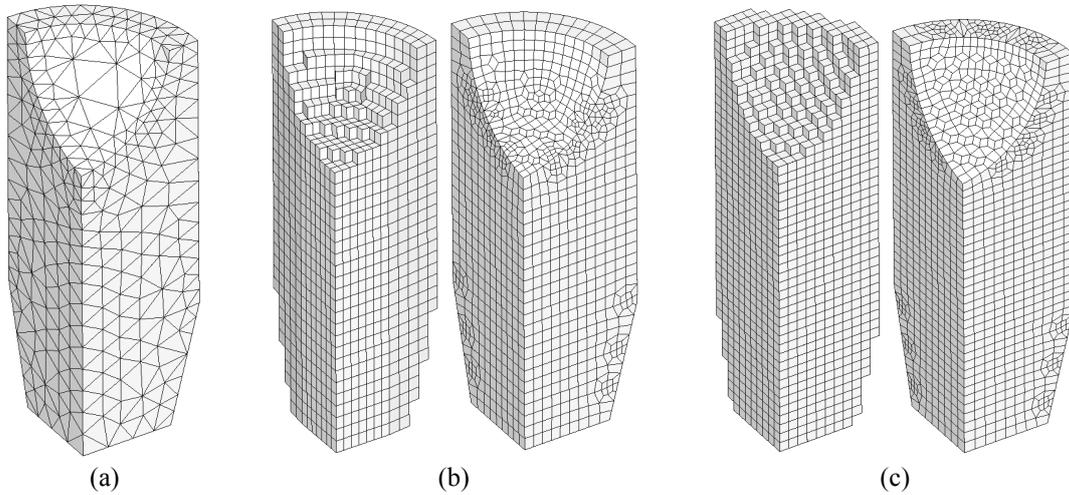


Fig. 5.5. Influence of the adaptive core mesh to the final hexahedral mesh. (a) Triangular surface mesh to be meshed by hexahedral elements. (b) Circular cylindrical core mesh and corresponding final hexahedral mesh. (c) Cuboid core mesh and corresponding final hexahedral mesh.

the ray-tracing algorithm described by O'Rourke [19] is applied to determine whether a node is inside or outside the triangular surface mesh that is utilized to represent the workpiece. According to this procedure, following the direction of an arbitrary vector starting in each node and counting the number of intersections with triangular elements of the surface allows determining if a node is inside (odd number of intersections) or outside (even number of intersections) the workpiece. Elements with at least one node outside the workpiece are removed, and the topology list packing completes the construction of the core mesh.

After the generation of the core mesh, it is necessary to reconstruct the geometrical features identified above. The reconstruction includes introduction of an additional layer of elements on the core mesh with projection to the surface, projection of selected nodes to vertices and projection of nodes to reconstruct edges. The reconstruction of surfaces is successfully performed with the isomorphism technique originally proposed by Schneiders and Bünten [5] and later modified by Fernandes and Martins [17]. The isomorphism technique is based on the generation of a layer of elements between the core mesh and the triangular surface mesh that defines the contour of the workpiece. The core mesh is smoothened before projecting the outmost nodal points to the triangular surface mesh in order to avoid crossing of adjacent surface normals that would create projection problems.

5.3.3 Reconstruction of vertices

The objective in vertex reconstruction is to project nodal points of the aforementioned layer of hexahedral elements onto the vertices of the surface triangular mesh. The algorithm proposed within the present work (see also Nielsen et al. [18]) suggests that instead of using a purely distance based criterion, priority must be given to combination

of distance d and valence v (that is, the number of element edges attached to a node) in order to avoid creating degenerated elements that will need subsequent repairment. Moreover, the proposed algorithm is based on an iterative search for the best candidate A to be projected onto vertex V with the examining radius r being progressively increased in order to guarantee that the selection is always made between the closest nodal points,

$$A = f(d, v, r) \quad (5.2)$$

Fig. 5.6 illustrates the differences between conventional and the abovementioned proposed procedures. The corner of the hexahedral mesh marked in Fig. 5.6a is taken as the working region of the example and Fig. 5.6b-h provide magnified details. The candidates to be projected onto the existing vertex are the nearest node k_4 and the second nearest node k_3 , which have valences 4 and 3, respectively.

By pure distance criterion, node k_4 would be projected (Fig. 5.6c) and vertex reconstruction would lead to a degenerated hexahedron (Fig. 5.6d) that after repairment would provide the final mesh shown in Fig. 5.6e. If, instead, selection is prioritized by a combination of distance and valence matching the number of edges meeting in the existing vertex, node k_3 would be chosen. Projection of k_3 is shown in Fig. 5.6f, corresponding reconstruction of edges is illustrated in Fig. 5.6g, and the final mesh is provided in Fig. 5.6h. The mesh resulting from the new proposed algorithm is more regular than that generated by vertex reconstruction based on pure distance.

Iterative search of the best candidate A is illustrated in Fig. 5.7 for a test case consisting of a square pyramid and a vertex V , where candidate nodes are sought itera-

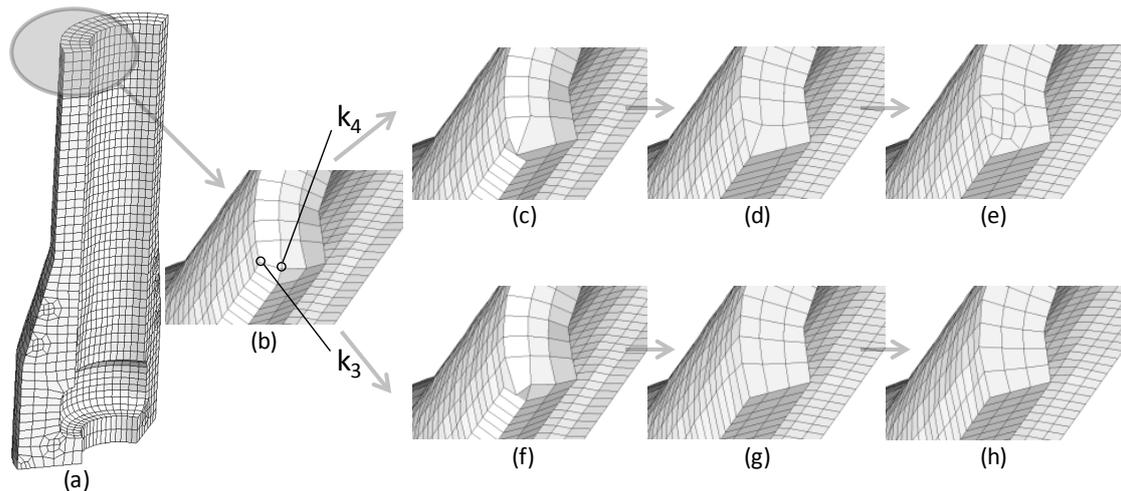


Fig. 5.6. Selection of the nodal point to be projected onto an existing vertex. (a) Working region of the example. (b) Candidate nodal points to be projected onto the existing vertex. (c-e) Vertex reconstruction by pure distance and its influence on the final hexahedral mesh. (f-h) Vertex reconstruction by combining distance and valence and its influence on the final hexahedral mesh.

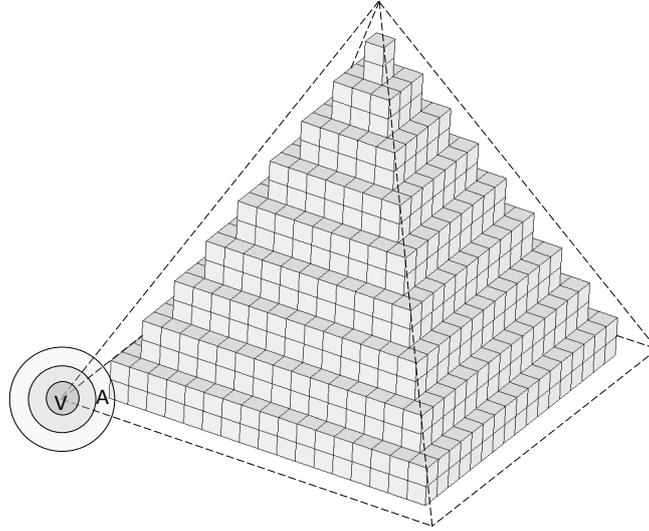


Fig. 5.7. Combining iterative searching distance with valence to select the best candidate A to be projected onto the existing vertex V .

tively by increasing the search radius r until the best candidate is found. A valence criterion is applied among candidates, in this case selecting node A to be projected onto vertex V . The iterative searching procedure is important for reconstructing sharp corners, where the distance to the core mesh can be large and no candidates are likely found at first. On the other hand, the iterative procedure allows the search radius to grow from small values in order to avoid candidates located far away to be projected onto the existing vertex.

5.3.4 Reconstruction of edges

Edge reconstruction is the most critical step in all-hexahedral based meshing. The procedure is illustrated in Fig. 5.8 and is based on the algorithm by Kwak and Im [16] modified in the present work (see also Nielsen et al. [18]) to include additional geometrical features and topology based constraints. Fig. 5.8a shows the final mesh of the example and Fig. 5.8c-e show magnified details of the intermediate meshes. The mesh included in Fig. 5.8c was plotted after vertex reconstruction while the meshes in Fig. 5.8d-e were taken after partial (from vertex node V to edge node P) and final reconstruction of edges. It is important to notice that ‘final reconstruction’ of an edge should not be confused with its ‘completeness’, as can be easily observed in Fig. 5.8e. Completion of edges is discussed in Section 5.3.5.

In the selection of the best candidate to project to the edge after P , the first step is the identification and sorting of up to three candidate nodes based on the distance to the last projected node P . Nodal point $k_{(1)}$ is the first potential candidate, but it is discarded because its angle with the previous part of the edge is larger than a critical threshold value (say 60°), implying that $k_{(1)}$ is not considered as a node ahead of P .

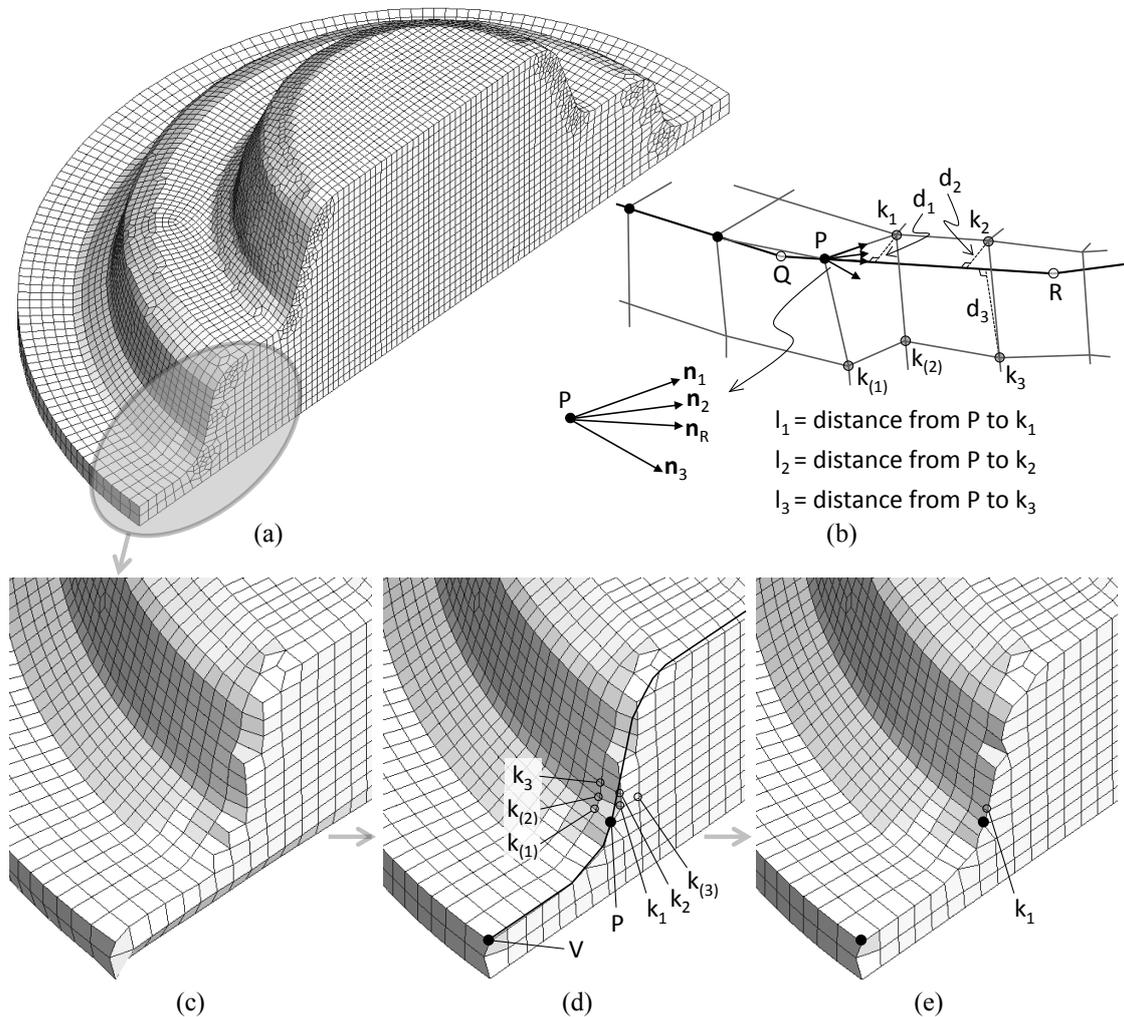


Fig. 5.8. Reconstruction of edges in a typical forged flange component. (a) Final hexahedral mesh. (b) Schematic illustration of reconstruction process on edge segment QR . (c) Mesh after reconstruction of vertices. (d) Mesh after partial reconstruction of edges. (e) Mesh after final reconstruction of edges showing k_1 projected on the edge and evidence of lack of completeness.

The next nodal point to be evaluated is k_1 , which fulfills all the necessary conditions and is accepted as candidate number 1. Node $k_{(2)}$ is discarded as second candidate because it is diagonally opposed to P (that is, line segment drawn from $k_{(2)}$ to P is a diagonal of the quadrilateral face). Node k_2 is selected as candidate 2, and during selection of candidate 3, $k_{(3)}$ is discarded for also being diagonally opposed to P , and k_3 is selected instead.

The introduction of the topology based criterion avoiding diagonally opposite nodes on the edge prevents the occurrence of degenerated hexahedral elements along the edges. However, this type of constraint should not be confused with the necessity of

having diagonals on the edges for ensuring its completeness as addressed in Section 5.3.5.

The second step is to choose between the three identified candidates. Fig. 5.8b illustrates such a situation with respect to the latest projected node P for candidate nodes k_1 , k_2 and k_3 . The first priority is obtained after applying the following function (hereafter named g -function),

$$g = (1 - \mathbf{n}_R \cdot \mathbf{n}_i) + \frac{d_i}{\max(d_1, d_2, d_3, C_0)} + \frac{l_i}{\max(d_1, d_2, d_3, C_0)} \quad (5.3)$$

where \mathbf{n}_R is the unit vector from P to R and \mathbf{n}_i are the unit vectors from P towards the candidates. The distances from the candidates to P are denoted l_i and the distances from the candidates to the edge segment $Q-R$ of the triangular surface mesh are represented by d_i , where index i refers to the candidates. The constant C_0 refers to the characteristic element side of the core mesh.

In choosing between candidates, first priority is given to the candidate minimizing the g -function while the candidate maximizing the g -function is directly discarded. In the example in Fig. 5.8, candidate k_1 is selected from the minimization of the g -function, and the projection of k_1 is shown in Fig. 5.8e together with the remaining projections based on the application of the proposed algorithm.

The first two terms of the g -function were originally suggested by Kwak and Im [16] and account for the selection of candidates that minimizes collinearity and distance to the edge segment. The third term is added (Nielsen et al. [18]) to force minimization of the g -function to be dependent on the distance to the latest projected node P . The importance of the new term is best illustrated by an example where two candidates are equidistant to a straight edge, such that the second term in (5.3) is of no importance. In this case the first term alone would prioritize the candidates further away from P due to collinearity, although the nearer node may fulfill all other criteria for being chosen. The third term adds robustness by compromising between collinearity and distance to node P .

Fig. 5.9 illustrates the importance of topology based prioritization rules that exclude diagonally opposite candidates to the last projected node P (or, first node V in case of vertices). From the vertex V in Fig. 5.9a the first node to be projected to the edge towards the upper right corner has to be selected between candidates k_1 and k_2 , with the former candidate being the closest to V . Fig. 5.9b shows the reconstruction of the edges and the final mesh after repairs when k_1 is chosen. As seen, the resulting mesh is irregular and presents lower quality than that plotted in Fig. 5.9c, where the diagonally opposed node to the vertex is discarded and the neighbor is prioritized.

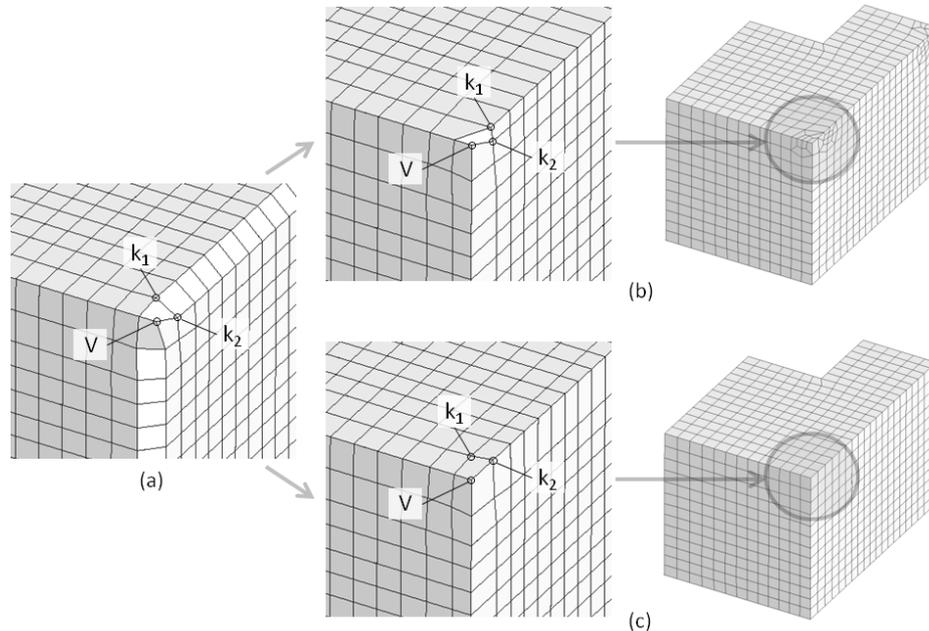


Fig. 5.9. The importance of topology based prioritization during reconstruction of edges. (a) Mesh after reconstruction of vertices. (b) Selection of the nearest candidate k_1 exclusively based on the g -function and resulting final mesh. (c) Selection of candidate k_2 and final mesh if diagonally opposed candidates to a nodal point (or vertex) V are discarded before application of the g -function.

5.3.5 Edge repairment

As it was mentioned in relation to reconstruction of edges in Fig. 5.8 and exemplified further in Fig. 5.10 by an extreme geometry in form of a hexahedron, there is often necessity of performing repairment of the edges in order to ensure completeness of the edges. Fig. 5.10a shows the core mesh generated from a cuboid bounding box and Fig. 5.10b shows the intermediate mesh after reconstruction of surfaces, vertices and edges. The necessity of repairment to complete the edges by element sides and to improve the element quality is obvious, and procedures to eliminate these geometrical inconsistencies are crucial for further utilization of the mesh.

Topology based procedures are implemented to resolve the lacking completeness of edges after the main reconstruction of edges according to the above procedures. Projection of nodes, such as $1a$ and $1b$ in Fig. 5.10b, is one of the topology based repairment procedures applied. Nodes $1a$ and $1b$ are characterized by being neighbors of two consecutive nodes on the edge that do not share an element side. Their projections will locally complete the corresponding edges. Another topology based repairment is the projection of neighboring pairs of nodes such as $2a$ and $2b$ in Fig. 5.10b. Each of these nodes is neighboring one of two consecutive nodes on the edge, which do not share an element side. Again, the repairment ensures local completion of the edge. The

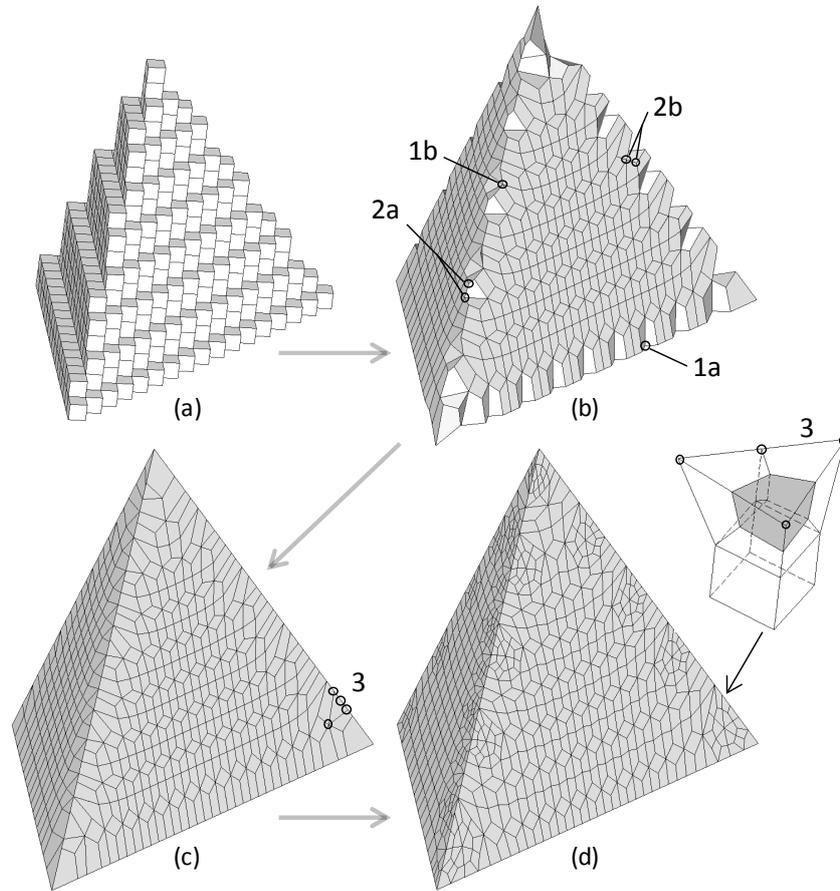


Fig. 5.10. Selected overview of topology based repair of edges in all-hexahedral meshing of a tetrahedron. (a) Core mesh obtained from a cuboid bounding box. (b) Mesh after reconstruction of surfaces, vertices and edges. (c) Mesh after topology based repairs illustrated by node pairs $1a$, $1b$ and $2a$, $2b$. (d) Final mesh after application of templates for eliminating degenerated hexahedra (e.g. '3') by means of its decomposition into well-shaped hexahedra.

mesh of the tetrahedron after performing the two previously mentioned types of repairment is shown in Fig. 5.10c.

At this stage the edges are complete, but additional repairment is still necessary to resolve degenerated elements, such as that labeled '3' in the figure. The element has three nodes along the edge and can be split into four elements of better quality by means of the template proposed by Schneiders and Bünthen [5]. The template is illustrated by the detail in Fig. 5.10d, where also the resulting hexahedral finite element mesh is shown.

Another procedure for solving geometric inconsistency along edges and improving the mesh quality has been utilized by Kraft [4] and Kwak and Im [16] who add a thin layer of extra elements with high quality along edges.

5.3.6 Smoothing

Smoothing procedures are applied with the purpose of repairing distorted elements and improving their shape in different stages of meshing and remeshing. In general terms, smoothing is accomplished by changing the position of the nodal points to new positions given by a weighted average of the neighboring nodal points without modifying the topology of the mesh.

Several constraints must be taken into account to preserve the geometrical consistency of the hexahedral meshes. Vertices are excluded from smoothing as their positions are fixed. Edge nodes stay on the edges and the surface nodes remain on the surfaces. To overcome these constraints, edges are smoothed first by means of a parametric based procedure developed within the present work (see Nielsen et al. [18]). Surfaces are smoothed next while excluding the edge nodes, and finally, the volume is smoothed while excluding all nodal points located on edges and surfaces.

Edge smoothing can be performed by means of three different concepts that are schematically shown in the example provided in Fig. 5.11a. Local smoothing and global smoothing with re-projection are classical relaxation techniques that are comprehensively described in the work of Karadogan and Tekkaya [15] and Fernandes and Martins [17]. Edge smoothing by means of a parametric based procedure is a new relaxation technique proposed in the present project that is capable of providing more regular and uniform hexahedral meshes.

An understanding of the main differences in the abovementioned edge smoothing procedures is provided in the following. The physical position of a node k , placed on the edge between vertices V_1 and V_2 , is defined as the distance L_k from vertex V_1 to k along the individual edge segments of the triangular mesh. From the example shown in Fig. 5.11a (labeled as ‘before smoothing’) it follows that,

$$L_k = L_1 + L_2 + 0.25 \cdot L_3 \quad (5.4)$$

The parametric position of node k is given by counting the number of edge segments from V_1 to k ; in this case two edge segments plus a fraction of the third segment lead to $S_k = 2.25$. A direct translation between the two types of measurements is possible.

In parametric edge smoothing, the physical position L_i is identified for each node, and the k 'th node is repositioned after satisfying the following condition,

$$L_k = \frac{L_{k-1} + L_{k+1}}{2} \quad (5.5)$$

The new physical position is translated into a corresponding parametric position that will identify the updated position of node k along the edge segments. The procedure is iterative, inexpensive, and offers the advantage of node k being able to move along

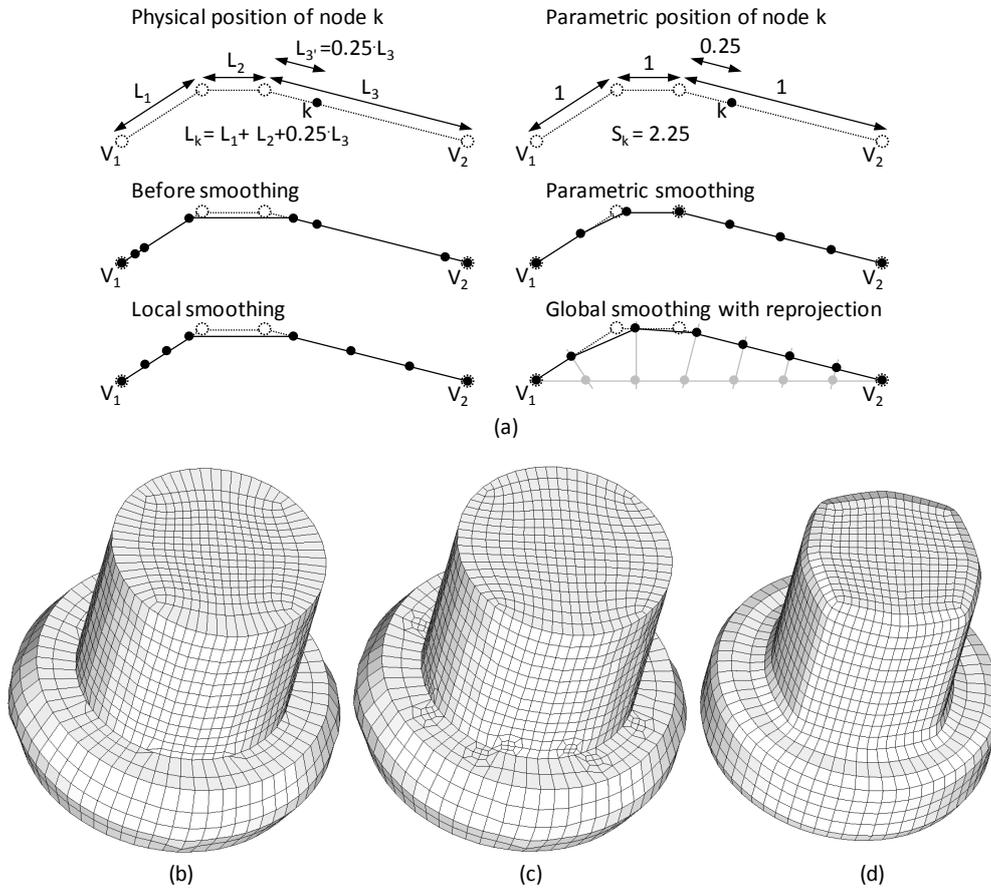


Fig. 5.11. Improvement of hexahedral mesh by smoothing. (a) Schematic representation of the three different procedures utilized in edge smoothing. (b) Edge smoothing of a headed forged model undergoing remeshing. (c) Surface smoothing of the part shown in (b) after solving degenerated hexahedral elements. (d) Volume smoothing of the core mesh.

the edge segments and from one edge segment to another. The final result is shown in Fig. 5.11a (labeled ‘parametric smoothing’) and, as can be easily observed, the distance between consecutive nodes along the edge is constant.

The advantages of the parametric edge smoothing against the abovementioned local and global edge smoothing are schematically resumed in Fig. 5.11a. Local edge smoothing is similar to (5.5), but only works among groups of nodes on a straight line. Global edge smoothing moves node k to an average position of nodes $k - 1$ and $k + 1$ that is located outside the edge. In the extreme case of applying a large number of iterations, all the nodes would be positioned on the straight line connecting vertices V_1 and V_2 and subsequent orthogonal re-projection to the edge would be essential to recover geometrical consistency. The consequence of re-projection is the impossibility of ensuring that nodes will be positioned on equally spaced positions along the edges.

Because parametric smoothing is the only edge smoothing technique that is capable of ensuring equally spaced nodes along the edges, it helps improving the overall quali-

ty of meshing and remeshing. The application of parametric edge smoothing to a head-forged component is shown in Fig. 5.11b.

Surfaces are smoothed by averaging nodal positions according to the weighted areas of the neighboring surface quadrilaterals (Fig. 5.11c) and volume smoothing is performed by averaging nodal positions of the core mesh according to weighted volumes of neighboring hexahedral elements (Fig. 5.11d). Both surface and volume smoothing procedures are comprehensively described elsewhere; see Karadogan and Tekkaya [15] and Fernandes and Martins [17].

5.3.7 Application of all-hexahedral meshing

A connecting rod is presented in Fig. 5.12 to illustrate the applicability of the presented all-hexahedral meshing technique. The shape of the connecting rod is provided by a triangularized surface from AutoCAD. Following the above presented procedures from identification of geometric features, generation of core mesh and reconstruction of edges, the smoothed mesh with improved element conditions is obtained as shown in Fig. 5.12.

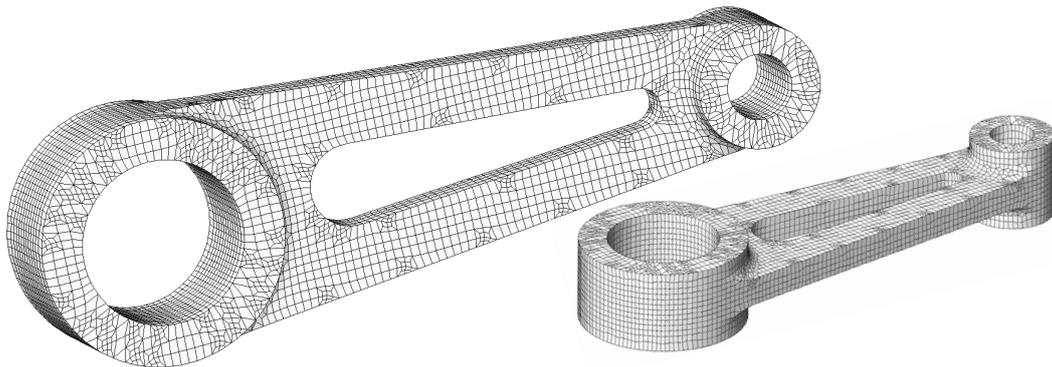


Fig. 5.12. Hexahedral mesh of a connecting rod showing the capabilities of the presented all-hexahedral meshing technique.

5.4 Addition of interface layers

Addition of thin layers of elements on selected surfaces of existing meshes is an important step in setting up resistance welding simulations. Interface layers as described in Section 4.3 (see particularly Fig. 4.5) and also eventual coatings are included in thin layers of elements on the surfaces of object meshes. The layers are most easily added to a surface by extruding the surface mesh in the normal direction with the given thickness. The surface normal for projection of a surface node (node normal) is taken as the normalized average of the normals of the surrounding element faces (face normals). An example is given in Fig. 5.13a in terms of a sheet with a longitudinal em-

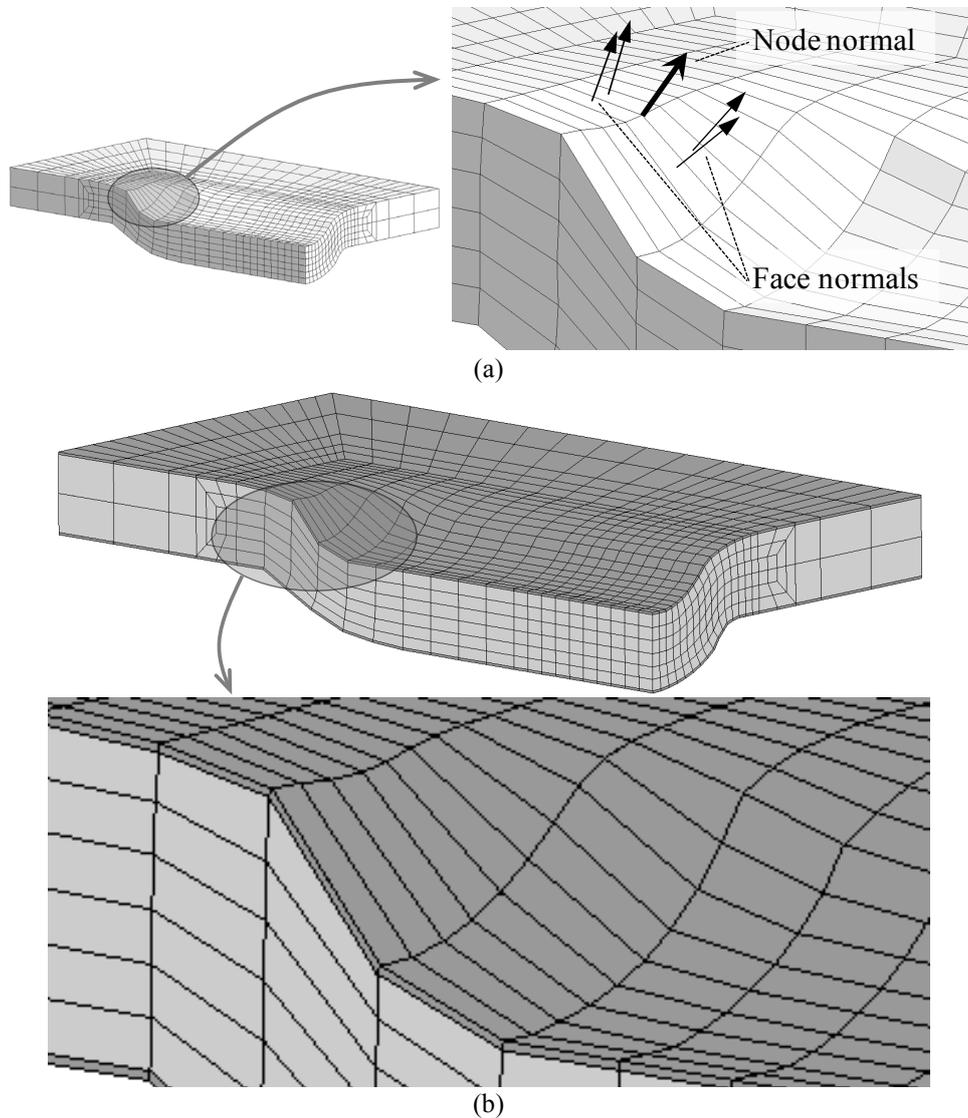


Fig. 5.13. Addition of upper and lower thin layers of elements for supplying coating or interface properties to a sheet with a longitudinal projection. (a) Definition of node normals by normalized averaging of surrounding face normals. (b) Resulting mesh after addition of the two layers.

bossment for projection welding. The right side of the figure shows a node normal in one of the nodes and the relevant face normals for the normalized averaging. Besides projection of the nodes along the node normals with a distance corresponding to the layer thickness, the procedure also involves setup of element topologies for the new layer of elements.

Fig. 5.13b shows the sheet with embossment after adding a thin layer of elements on each side of the sheet, while there is no layer on the sides of the sheet. The two layers are added separately under constraints defining where to add the layer. Two types of constraints are included to limit the surface area to include the new layer. The most

important constraint is only to extrude element faces with face normal within the range of a reference normal and its tolerance. The upper layer in Fig. 5.13b, as an example, is added with an “upward” reference normal and a tolerance of 60° . In practical terms, all element faces with an “upward” normal or within a cone of 60° are extruded to form the upper layer. This excludes clearly the lower side with face normals pointing “downward”, and it also excludes the side of the sheet, which would only be included if the tolerance was 90° or more. The lower layer is added subsequently with a “downward” reference normal and again a tolerance of 60° . The sheet with its interfaces is shown in Fig. 5.13b.

The second constraint for the selection of element faces to be extruded is limitation by coordinates of the centroid of the face. This is useful when adding a layer of elements to only part of a surface.

5.5 Remeshing

The description of the remeshing procedures is based on the example shown in Fig. 5.14, representing resistance welding of a square nut to a sheet. Resistance welding is an extreme case of multi-object simulation involving electro-thermo-mechanical modeling as described in the present work, and in terms of remeshing it presents several complications due to multiple objects and local effects presented by the process.

Fig. 5.14a shows the initial mesh of the case simulated by one quarter due to symmetries. A standard component in the automotive industry in form of an M10 steel square nut (1) is welded to an AISI 1008 steel sheet (2) of 1.4mm thickness. A $50\mu\text{m}$ thin layer of elements (3) on top of the sheet provides the interface properties between the square nut and the sheet. Electrical and thermal resistances stemming from oxide layers, surface films and contaminants are included in this layer. The electrical and thermal contact properties change with temperature and contact pressure due to formation of real contact area break down and squeeze out of the impurities. A 15kA direct current is applied through the copper alloy electrodes (4) and (5) from each side of the square nut and the sheet.

The temperature field after simulating 80ms of the resistance welding process is shown in the deformed mesh of Fig. 5.14b. Fig. 5.14c-d show the temperature of the original and the remeshed cases after additionally 40 simulation steps, corresponding to 100ms total process time. Nearly identical shape and temperature distributions prove the accuracy, reliability and validity of the overall procedure outlined in the following.

At the stage shown in Fig. 5.14b several elements in the bottom of the square nut (1) are flattened such that remeshing becomes necessary to carry on the simulation with high accuracy. However, the meshes in the remaining objects (2), (3), (4) and (5) are practically not distorted and, therefore, do not undergo remeshing. In other words, remeshing is only performed in the selected object (1). In case of remeshing, the sur-

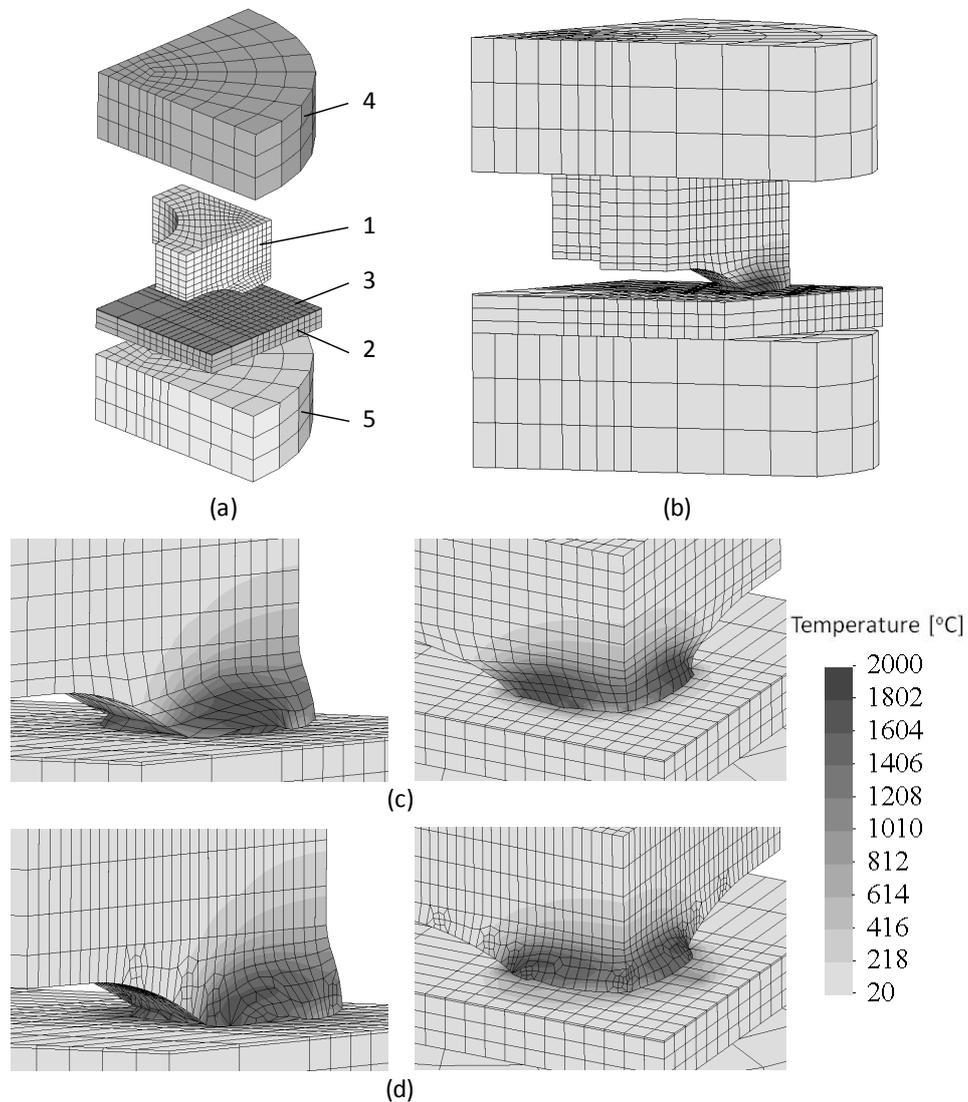


Fig. 5.14. Resistance welding of a square nut to a sheet. (a) Initial mesh of the multi-object finite element model. (b) Temperature distribution in original deformed mesh after $80ms$. (c) Temperature distribution in original deformed mesh after $100ms$. (d) Temperature distribution in the remeshed configuration after $100ms$, where remeshing took place after $80ms$.

face mesh is extracted from the deformed geometry by splitting each of the surface quadrilaterals into two triangular elements. Hereafter follows the all-hexahedral meshing procedures outlined in Section 5.3 for the individual object.

5.5.1 Multi-object procedures and tool contact

The presence of multiple objects poses the necessity of paying special attention to ensure that penetration or gaps are avoided in regions where contact conditions prevail. The majority of nodes after remeshing will in general be located on the element faces

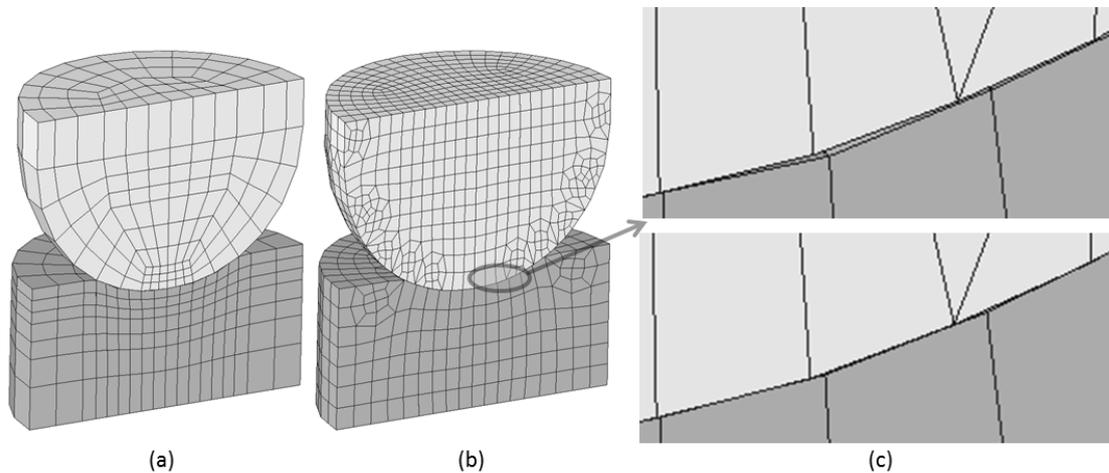


Fig. 5.15. Multi-object remeshing. (a) Mesh before remeshing. (b) Mesh after remeshing. (c) Detail of the contact interface before (upper) and after (lower) reconstruction of contact conditions.

of the previous distorted mesh but without coinciding with previous nodes. As a result, the surface of an object after remeshing will not be identical to the surface before remeshing, and therefore contact conditions are not guaranteed to be maintained unless the interfaces are planar. The solution developed in this work is to reposition nodes of one object by orthogonal projection to an element face of another object if the orthogonal distance between them is less than a certain threshold tolerance. Additionally, all surface nodes of an object are tested for penetration into elements belonging to any other objects even if it exceeds the aforementioned tolerance.

A similar procedure is implemented for maintaining contact conditions between an object and a rigid tool with the constraint that only nodes of the object can be moved.

Fig. 5.15 illustrates the above discussion for a case involving non-planar surfaces in contact. Fig. 5.15a shows the mesh resulting from the simulation of a hemispherical ball pressed into a circular cylinder end face including deformation due to contact. Fig. 5.15b shows the new mesh after individual remeshing of each object has taken place. Fig. 5.15c shows the resulting interface after remeshing (upper) and after orthogonal projection (lower) to close the gap stemming from remeshing.

5.5.2 Transfer of history dependent variables

An additional step is necessary to complete the remeshing. The history dependent field variables, such as strain, current density and temperature, need to be transferred from the old to the new mesh. This requires the evaluation of the nodal values of these quantities in the old mesh.

Averaging by weighted volumes of surrounding Gauss points is frequently applied, but it is possible to compute better nodal values extrapolated from Gauss points by applying a recovery technique based on least square fitting. The application of least square fitting requires the minimization of the following functional, I ,

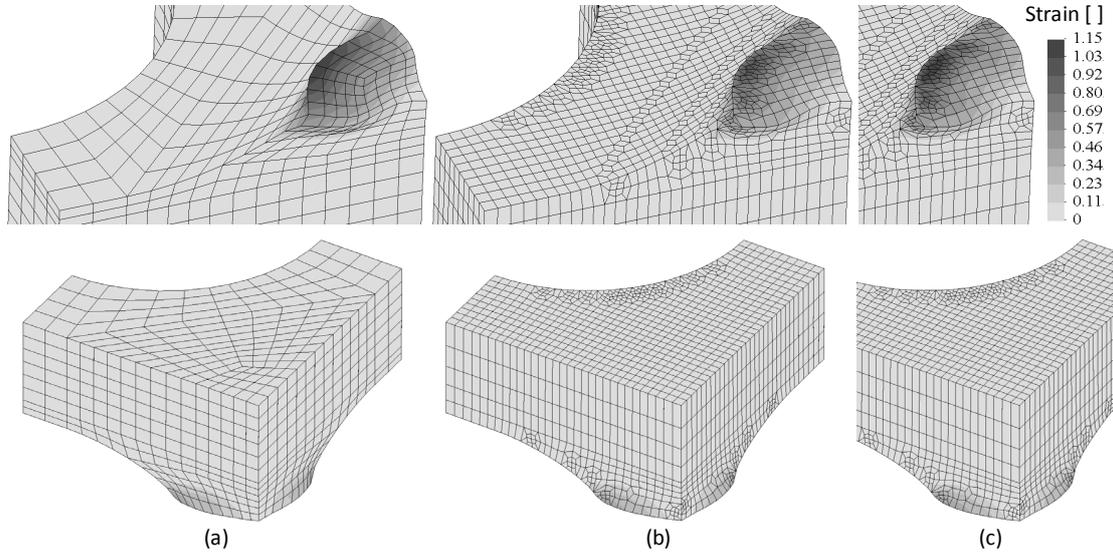


Fig. 5.16. Mesh and remesh details of the square nut after 80ms . Transfer of field variables is performed for the effective strain. (a) Distribution of effective strain in the original mesh. (b) Distribution of effective strain after remeshing based on the averaged values of surrounding Gauss points weighted by element volumes. (c) Distribution of effective strain after remeshing based on least square fitting according to (5.6).

$$I = \int \sum_k \left[\sum_i N_i f_i - c_k \right]^2 dV_k \quad (5.6)$$

where c_k is the known value of the time-integrated field variable at the centre Gauss point of element k , f_i is the nodal quantity to be determined and N_i is the conventional shape function of node i . Details of the procedure are described elsewhere by Martins et al. [20] and Fernandes and Martins [17].

Detailed views of the mesh and transfer of field variables in the example of the square nut to sheet resistance welding case are provided in Fig. 5.16. A large number of elements have been applied to capture the details of the leg of the square nut (see also Fig. 5.14c-d). However, because the remaining part of the square nut has no or little deformation, the mesh density in this region is made lower in order to reduce the overall number of elements. The resulting mesh is shown in Fig. 5.16, where the entire square nut is shown in the lower figures and the details near a leg are shown in the upper figures. The overall number of elements is raised by a factor of 2.5 due to remeshing and resulting refinement in order to capture all the technological relevant local details.

The transfer of field variables (here exemplified by the effective strain) is performed from the original mesh in Fig. 5.16a to the new mesh in Fig. 5.16b-c. In Fig. 5.16b the transfer is accomplished by averaging of neighboring Gauss point values by weighted volumes, whereas the least square method according to (5.6) has been applied in Fig.

5.16c. As observed, the peak values (compare the dark color) are kept better when the transfer is performed by least square fitting than by volume weighted averaging.

References

- [1] Coupez T, Soyris N, Chenot JL (1991) 3-D finite element modelling of the forging process with automatic remeshing. *Journal of Materials Processing Technology* 27(1-3):119-133.
- [2] Tekkaya AE, Martins PAF (2009) Accuracy, reliability and validity of finite element analysis in metal forming: a user's perspective. *Engineering Computations* 26:1026-1055.
- [3] Benzley SE, Perry E, Merkley K, Clark B (1995) A comparison of all hexagonal and all tetrahedral finite element meshes for elastic and elasto-plastic analysis. *Proceedings of the 4th International Meshing Roundtable*, pp. 179-191.
- [4] Kraft P (1999) Automatic remeshing with hexahedral elements: Problems, solutions and applications. *Proceedings of the 8th International Meshing Roundtable, South Lake Tahoe, USA*, pp. 357-367.
- [5] Schneiders R, Buntin R (1995) Automatic generation of hexahedral finite element meshes. *Computer Aided Geometric Design* 12:693-707.
- [6] Santos A, Makinouchi A (1995) Contact strategies to deal with different tool descriptions in static explicit FEM for 3-D sheet metal forming simulations. *Journal of Materials Processing Technology* 50:277-291.
- [7] Boussetta R, Coupez T, Fourment L (2006) Adaptive remeshing based on a posteriori error estimation for forging simulation. *Computer Methods in Applied Mechanics and Engineering* 195:6626-6645.
- [8] Behrens BA, Kerkeling J (2011) Pierced forgings: tool development for a combined single step process. *Production Engineering – Research and Development* 5:201-207.
- [9] Fernandes JLM, Alves LM, Martins PAF (2012) Forming Tubular Hexahedral Screws – Process Development by Means of a Combined Finite Element-Boundary Element Approach. *Engineering Analysis with Boundary Elements* 36:1082-1091.
- [10] Rodrigues JMC, Martins PAF (1998) Coupled thermo-mechanical analysis of metal-forming processes through a combined finite element-boundary element approach. *International Journal for Numerical Methods in Engineering* 42:631-645.
- [11] Zienkiewicz OC, Phillips DV (1971) An automatic mesh generation scheme for plane and curved surfaces by isoparametric coordinates. *International Journal for Numerical Methods in Engineering* 3:519-528.
- [12] Martins PAF, Barata Marques MJM (1992) Model3 – A three-dimensional mesh generator. *Computers & Structures* 42(4):511-529.
- [13] Schneiders R, Oberschelp W, Kopp R, Becker M (1992) New and effective remeshing scheme for the simulation of metal forming processes. *Engineering with Computers* 8:163-176.
- [14] Zhu J, Gotoh M (1999) An automated process for 3D hexahedral mesh regeneration in metal forming. *Computational Mechanics* 2:373-385.
- [15] Karadogan C, Tekkaya AE (2001) Geometry defeaturing and surface relaxation algorithms for all-hexahedral remeshing. *Proceedings of the Seventh NUMIFORM Conference, Toyohashi, Japan*, pp. 161-166.

- [16] Kwak DY, Im YT (2002) Remeshing for metal forming simulations – part II: three-dimensional hexahedral mesh generation. *International Journal for Numerical Methods in Engineering* 53:2501-2528.
- [17] Fernandes JLM, Martins PAF (2007) All-hexahedral remeshing for the finite element analysis of metal forming processes. *Finite Elements in Analysis and Design* 43:666-679.
- [18] Nielsen CV, Fernandes JLM, Martins PAF, All-hexahedral meshing and remeshing for multi-object manufacturing applications. Submitted for publication in international journal.
- [19] O'Rourke J (1998) *Computational Geometry in C*, Cambridge University Press, Cambridge.
- [20] Martins PAF, Marmelo JCP, Rodrigues JMC, Barata Marques MJM (1994) Plarmsh3 – a three dimensional program for remeshing in metal forming. *Computers & Structures* 53:1153-1166.

6. Parallel Equation Solver

When solving large finite element problems, solution time becomes a factor which cannot be ignored, and it is among the concerns when considering modeling in three dimensions instead of two dimensions. Different approaches are available to reduce the computational cost. Decomposition of a finite element domain into subdomains allows naturally for parallel computation of the subdomains to save overall computation time; see e.g. El-Sayed and Hsiung [1]. Interface nodes between substructures couple the substructure solutions, and thus communication between the processors are needed. In order to keep the amount of interface nodes minimal, Farhat [2] and Al-Nasra and Nguyen [3] have proposed algorithms for optimal decompositions. Another way of saving computation time is to apply faster solution techniques to solve the system of equations. This can be done either by solving iteratively, sequentially or in parallel, or by parallelizing the equation solver, such that it remains a direct solver.

This chapter presents the parallelization of a direct solver implemented in skyline storage format. The parallelization is part of the present work and has given the possibility of performing the presented simulations within reasonable time as well as it will save time in future simulations both in I-Form3 and SORPAS 3D. It is furthermore available for other existing finite element codes as far as the developers want to utilize it. This option is made possible through publication of the source code [4-5], which is also included in Appendix A.

6.1 Strategies of solution techniques

In iterative solvers, the solution is found iteratively to satisfy the equation system to within a specified tolerance. This is faster than directly solving the equation system as long as the rate of convergence is fast enough. Lanczos [6] and Hestenes and Stiefel [7] have e.g. proposed conjugate gradient (CG) iterative solvers, which were later improved by preconditioning, see Meijerink and van der Vorst [8], where a matrix is multiplied to each side of the system to precondition the system and thereby improve the convergence behavior.

The drawback of the iterative solving is that accuracy is lost compared to direct solving. The accuracy depends on the threshold value used for accepting the solution, and a compromise between accuracy and computation speed is necessary. The small

inaccuracies accumulate and may result in poor satisfaction of boundary conditions, and symmetries may not be exactly obeyed (for instance, a zero displacement associated with a symmetry condition may be computed as a very small non-zero displacement creating problems in the overall modeling accuracy). In problems involving contact it may, for larger threshold values, also disturb the contact algorithms, eventually leading to penetration. Iterative solvers have also been reported unstable when dealing with ill-conditioned equation systems, whereas direct solvers are more robust, cf. Farhat and Wilson [9]. Due to the highly ill-conditioned systems dealt with in the present finite element implementation (the irreducible flow formulation with penalty contact), direct solvers are preferred and therefore parallelization of the iterative solver will not be considered for the present modeling purposes, while it may be considered for other finite element implementation purposes.

Parallelizing direct solvers is another way of saving computation time. The parallelization itself is considered more tedious, but once it is done, the time savings are easily obtained, and the accuracy is maintained to precision comparable to the sequential direct solver. Applying a parallel direct solver also diminishes the need for decomposition, although they can go together. Diminishing of this necessity entails that the parallel solver can be directly applied to any problem. Parallelization can be applied for local memory processors as well as for shared memory processors, where the first typically is applied to a cluster of multiple computers, whereas the latter typically would be one computer with multiple threads.

6.2 Parallel skyline solver

This section presents the parallelization of a skyline solver that was originally developed in the present work (see also Nielsen and Martins [4]). The skyline format of the system matrices is chosen because of the large sparsity typical for finite element models. Alternative compressed sparse row storage formats would also be relevant and can be an alternative for later improvements. When finite element programs are transferred to the industry, they are increasingly often intended for execution on standard PC's, which nowadays are equipped with several cores and threads with shared memory. It is therefore an obvious request that the programs can utilize all the threads to reduce the computational time.

The proposed parallel skyline solver is easily implemented into existing finite element codes as only the call to the skyline solver has to be replaced by a call to the new solver. The requested inputs are the stiffness matrix in skyline storage format together with the corresponding pointers to the diagonal positions, the right hand side, the number of equations and the number of threads to be utilized. Before this solver, Farhat and Wilson [9] published a parallel skyline solver programmed in Force, and Synn and Fulton [10] have proposed procedures to predict the performance of parallel skyline solving.

6.2.1 Brief overview of equation solving by Gauss elimination in skyline storage format

A review of the pure equation solving in skyline format is given before presenting the procedures for parallelization. A regular system of equations, like $\mathbf{K}\mathbf{v} = \mathbf{f}$, is considered, where \mathbf{K} is a symmetric $n \times n$ matrix and \mathbf{v} and \mathbf{f} are $n \times 1$ vectors containing the unknowns and the right hand side, respectively. Due to symmetry of the system matrix, only half of the matrix needs to be built and stored (slightly more than half due to storage of all diagonal positions). Furthermore, since most finite element systems are sparse and by proper node numbering have many zeros far from the diagonal, a skyline format as depicted in Fig. 6.1 is adopted. Omitting all zeros above the skyline reduces the storage and later the solution time significantly. Zeros may still exist below the skyline as the skyline encloses all non-zero positions.

In skyline format, the system matrix is typically stored in a one-dimensional vector \mathbf{s} with an additional index vector \mathbf{i} pointing to the diagonal positions. This is illustrated in Fig. 6.1c up to the seventh column. The size of the skyline vector is the number of positions under the skyline. The size of the index vector equals the number of rows or columns n . Then, it follows that the size of the skyline vector is i_n , since the last diagonal is the last position in the skyline vector.

The solution of the equation system is performed by Gauss elimination with column reduction, which is composed by the following three steps:

- Factorization of system matrix and reduction of right hand side (this step is performed column by column, thereby being “with column reduction”).
- Division of right hand side by system matrix diagonals.
- Backward substitution.

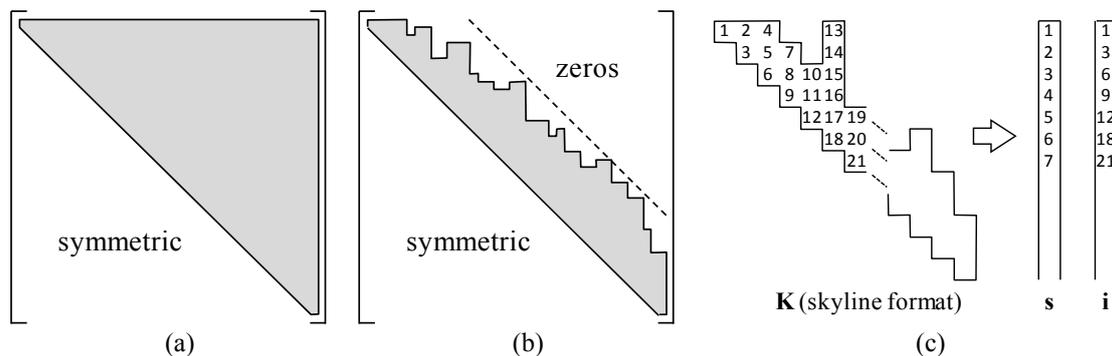


Fig. 6.1. System matrix in skyline storage format. (a) System matrix storage by utilizing symmetry only. (b) Skyline format by omitting zeros. The dashed line by the tallest skyline indicates the storage in banded format. (c) Format of skyline vector \mathbf{s} and index vector \mathbf{i} based on the original system matrix \mathbf{K} . Numbers correspond to position in the skyline vector.

Factorization of system matrix and reduction of right hand side

The first step in the factorization of the system matrix is illustrated by Fig. 6.2a for column j assuming that all columns $< j$ have been processed. The number of operations equals the active column height minus two as the diagonal position and the top-most position are not processed. The illustration in Fig. 6.2a has active column height 7 and five operations are depicted in the subfigures. In each operation, the k 'th position is subtracted the dot product formed by the vectors marked by “ o ”s and “ x ”s, i.e.

$$s_k = s_k - \{s_{\{o\}}\} \cdot \{s_{\{x\}}\} \quad (6.1)$$

Hereafter, all positions above the diagonal are divided by the diagonal position in the same row. In Fig. 6.2b, this corresponds to

$$\{s_{\{x\}}\} = \frac{\{s_{\{x\}}\}}{\{s_{\{o\}}\}} \quad (6.2)$$

where division is element-wise.

This is followed by reduction of the diagonal term by

$$s_k = s_k - \{s_{\{x\}}\} \cdot \{s_{\{x\}}^{old}\} \quad (6.3)$$

which means subtraction of each multiplication of new and old off-diagonal position. New is defined as “after (6.2)” and old is defined as “before (6.2)”.

Finally, the j 'th position in the right hand side is reduced by subtraction of the dot product spanned by the marked “ x ”s and “ v ”s in Fig. 6.2b. Note that the positions marked by “ x ”s are now the latest updated, meaning after (6.3). The reduction of the right hand side can be written as

$$f_j = f_j - \{s_{\{x\}}\} \cdot \{f_{\{v\}}\} \quad (6.4)$$

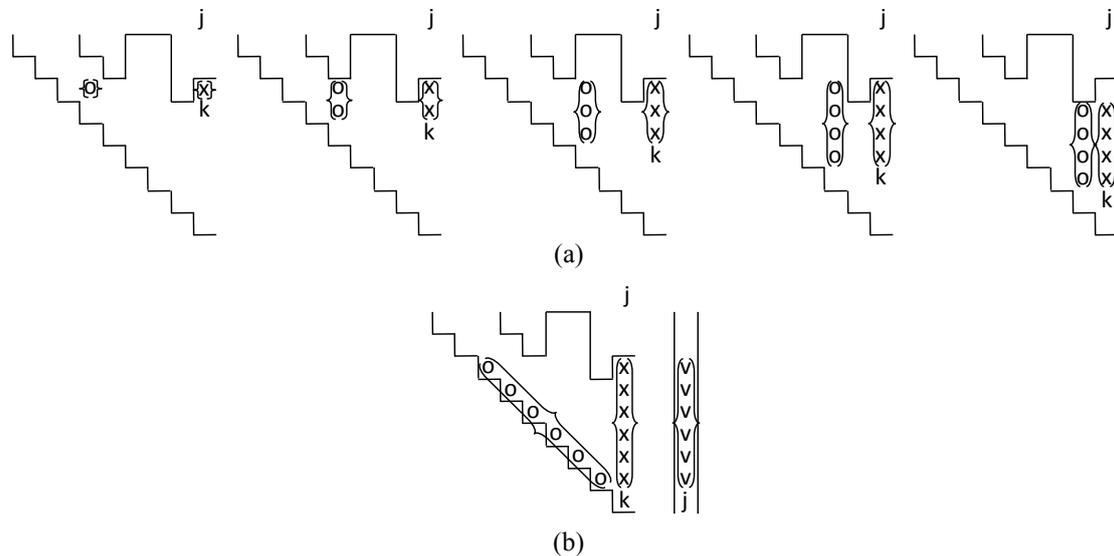


Fig. 6.2. Factorization of system matrix and reduction of right hand side. (a) Reduction of off-diagonal positions in column j by subtraction of dot product formed by “o”s and “x”s. (b) Reduction of diagonal positions in system matrix and reduction of right hand side.

Division of right hand side by system matrix diagonals

This step is straight forward: Each position j in the right hand side vector is divided by the system matrix diagonal term from the j 'th column. This completes the Gaussian elimination, such that all unknowns 1 to $j - 1$ are eliminated from the j 'th row.

Backward substitution

The unknowns are now found by backward substitution and stored in the right hand side vector, \mathbf{f} . The right hand side is processed backwards, such that Fig. 6.3 illustrates the substitution for the j 'th position assuming all positions $> j$ already processed. The positions marked by “v”s in the figure are modified by subtraction of positions in the system matrix marked by “x”s in the j 'th column. Before subtraction, these positions are multiplied by f_j , the j 'th position in the right hand side, that is

$$\{f_{\{v\}}\} = \{f_{\{v\}}\} - f_j \cdot \{s_{\{x\}}\} \quad (6.5)$$

The right hand side vector now includes the unknowns; $\mathbf{v} = \mathbf{f}$.

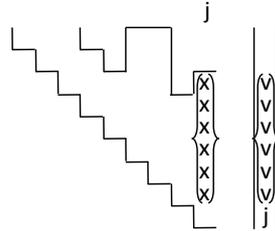


Fig. 6.3. Backward substitution by modification of positions $< j$ in right hand side vector by subtraction of system matrix positions marked by “x”s multiplied by the j 'th position of the right hand side.

6.2.2 Parallelization of skyline solver

The parallelization is column based in the sense that each thread is assigned a column to process, and when finishing one column assigned the next unprocessed column. The columns, however, cannot be processed independently, implying that communication between the threads is necessary. This is accomplished through the shared memory by updating the relevant variables from each thread while making sure that only one thread is updating certain variables at a time. The complete processing of a column requires completion of all preceding columns, but partial processing can be initiated even if this is not fulfilled. Then, while performing the partial processing, more of the preceding columns may have been fully processed in other threads, and in that case the remaining, or yet another partition, of the column can be processed. This procedure may lead to waiting time in each thread while dependent variables are being processed in other threads, especially when the differences in the skyline heights are large, corresponding to increased unevenness of the skyline profile in Fig. 6.1b.

The skyline solver is parallelized by OpenMP instructions in a FORTRAN implementation and is explained in details in the following with variable names matching the source code included in Appendix A. The factorization of the system matrix and reduction of the right hand side is parallelized, whereas the division of the right hand side by the system matrix diagonals as well as the backward substitution are left sequential as the time spent on these tasks are marginal compared to the factorization and reduction. The following is devoted to the description of the parallelized part of the solver.

A few variables are introduced preliminary:

<code>j</code>	Column number
<code>jmax</code>	Latest processed column number
<code>kmax</code>	Latest processed diagonal position
<code>nthreads</code>	Number of threads
<code>ithread</code>	Actual thread number, $ithread \in \{0, 1, \dots, nthreads-1\}$
<code>iloop</code>	Counter for commenced column in each thread
<code>iquit</code>	Flag identifying when all columns have been processed

skmatx	Skyline vector s containing the system matrix in skyline format
maxa	Index vector i pointing to the diagonal positions in s
fmatx	Right hand side f

Parallel region

The factorization and reduction are performed within a parallel region defined by the following listing:

```

jmax=1
kmax=1
!$OMP parallel default (none) &
!$OMP private (...private variables...) &
!$OMP shared (...shared variables...)
  ithread=omp_get_thread_num()
  iloop=0
  iquit=0
  do while (iquit.eq.0)
    ! To be described
  enddo
!$OMP end parallel

```

The first diagonal position in the skyline, corresponding to the entire first column, does not change during the factorization. Therefore, `jmax` and `kmax` are both initialized 1 (as if they were already processed). Hereafter, the parallel region is defined starting from “`!$OMP parallel`” and ending by “`!$OMP end parallel`”. In the beginning of the parallel region, all variables are identified as either private or shared. Typical private variables are actual thread number, counters and intermediate results, which are unique variables on each thread. The shared group of variables are variables that are read or modified on all threads. In case of modifications, it is important to know, which thread is modifying, and then let the remaining threads wait if they are about to modify the same variable. Within the parallel region, each thread works independently, except for the shared variables. The first instructions are to get the actual thread number, `ithread`, and initialize `iloop` and `iquit`. This is followed by the main loop (`do while (iquit.eq.0)`) which will loop through all columns until the factorization and reduction have finished.

Main loop

Each thread is assigned a certain column to process in each cycle. The main loop takes the following form:

```

do while (iquit.eq.0)
  iloop=iloop+1
  j=(iloop-1)*nthreads+ithread+2
  if (j.gt.ntotv) then
    iquit=1
    exit
  endif
  jr=maxa(j-1)

```

```

jd=maxa(j)
jh=jd-jr
is=j-jh+2
ie0=0
!$OMP flush (kmax)
do while (kmax.lt.jd)
  ! To be described
enddo
enddo

```

The thread's counter, `iloop`, is incremented during each cycle. Based on the local counter and the total number of threads, each thread is assigned a column, j , to process. If there are no more columns to process, `iquit=1` to terminate the specific thread. This might happen while other threads still process their last column. Fig. 6.4 illustrates the numbering of columns and the assignment of threads. As long as there are still columns to process, a few characteristic positions and dimensions are specified. The positions are the diagonal positions j_r and j_d of the $(j-1)$ 'th and the j 'th column, respectively, and the dimension is the active height j_h of column j . Finally, i_s is specified as the row number of the second active position in column j . These are all illustrated in Fig. 6.5. Before processing the column, a private variable, `ie0`, is initialized. This variable is later used as indicator of the amount of the column to be processed without waiting. Variable `kmax` is flushed before the core loop, meaning that it is updated in all threads, such that a change in one thread will be visible for all other threads.



Fig. 6.4. Schematic skyline matrix with 243 degrees of freedom and a total skyline vector length 735. The example shows processing on three threads. The numbers in the diagonal correspond to the diagonal positions in the skyline vector. Numbers above the matrix show column number, j , and the thread number, $ithread$, of the thread to process a given column.

before modifications, such that it is later available from both the current and the previous cycles. At this stage it is necessary to know how far the other threads are in their calculations, so k_{\max} and j_{\max} are flushed. An if-statement ($k_{\max} < j_r$) is judging if the entire column can be processed. If the statement is false (meaning $k_{\max} = j_r$), $i_{\text{hesitate}} = 0$ is kept and the entire column is processed in one cycle. Otherwise, if the statement is true, column $j-1$ has not been fully processed and it is necessary to hesitate, $i_{\text{hesitate}} = 1$. Hesitate does not mean wait, but means that only part of the column may be processed. The amount of the column that can be processed is defined by $i_{e0} = j_{\max}$. In the example of Fig. 6.5, only two positions in the column can be processed in this cycle of the loop; namely from i_s to i_{e0} . Then, in the next cycle (or a later cycle) k_{\max} and j_{\max} will have been increased on another thread, and the processing can be continued from $i_{e0} + 1$ to the diagonal position or to a limit provided by the new j_{\max} .

After judgment of i_{hesitate} and i_{e0} , the factorization and reduction can take place. If the active part of column j is one ($j_h = 1$), only the diagonal exists, and no processing is needed. If the active height is two ($j_h = 2$), only the diagonal and one other position exist and $i_s = j$, so only reduction of the diagonal is necessary. For all other heights ($j_h > 2$), reduction takes place in the diagonal and positions above. These two cases, $j_h = 2$ and $j_h > 2$, are detailed subsequently.

The factorization and reduction is followed by an if-statement to allow processing only when $i_{\text{hesitate}} = 0$. When this part of the code is reached without hesitation ($i_{\text{hesitate}} = 0$), the entire column has already been processed, and j_{\max} and k_{\max} have to be updated in order to let other threads proceed with the information that column $j_{\max} = j$ has been processed and the maximum processed diagonal position is $k_{\max} = j_d$. These variables are flushed immediately after the assignment. The assignment itself of the variables is enclosed in a critical region specifying that only one thread at a time can write to the variables. This is to avoid other threads to overtake and assign older values.

Reduction of off-diagonal terms, diagonal term and right hand side (case: $j_h > 2$)

The general case with $j_h > 2$ is explained first. Hereafter, the case with $j_h = 2$ is a special case. The following listing shows the reduction of off-diagonal positions, reduction of diagonal positions and finally reduction of the right hand side.

```

ie=jd-1+(ie0-j+1)*ihesitate
k00=jh-j-1+ie0old
if (k00.lt.0) k00=0
k0=0
do k=max0(jr+2,jd-j+ie0old+1),ie
  ir=maxa(is+k0+k00-1)
  id=maxa(is+k0+k00)
  ih1=min0(id-ir-1,1+k0+k00)
  if (ih1.gt.0) then

```

```

        ih2=min0(id-ir-j+(j-1-k0-k00)*ihesitate,2-j+k0+k00+(j-1-k0-
k00)*ihesitate)
        if (ih2.lt.1) ih2=1
        skmatx(k)=skmatx(k)-dot_product(skmatx(k-ih1:k-ih2),skmatx(id-
ih1:id-ih2))
        endif
        k0=k0+1
    enddo
    if (ihesitate.eq.0) then
        ir=jr+1
        ie=jd-1
        k=j-jd
        do i=ir,ie
            id=maxa(k+i)
            d=skmatx(i)
            skmatx(i)=d/skmatx(id)
            skmatx(jd)=skmatx(jd)-d*skmatx(i)
        enddo
        fmatx(j)=fmatx(j)-dot_product(skmatx(jr+1:jr+jh-1),fmatx(is-
1:is+jh-3))
    endif

```

The reduction of off-diagonal positions takes place in the first loop in which the operations illustrated by Fig. 6.2a are performed. In case of hesitation, $ihesitate=1$, only part of the operations can be performed in the current cycle. Fig. 6.6 illustrates such an example. Reading the figure from left to right, the first cycle starts, and j_{max} and k_{max} have been flushed. The “*” marks the latest processed diagonal position in the example. Then, since the $(j-1)$ 'th column has not yet been processed, the cycle is performed with hesitation. In the example in Fig. 6.6, only the first three operations can be performed. Then a new cycle is started with flushing of j_{max} and k_{max} . Assume that the $(j-1)$ 'th column has now been processed, as illustrated by “*”. This cycle is then performed without hesitation, and the remaining operations can be performed.

After the reduction of off-diagonal positions, the reduction of the diagonal position takes place inside the if-statement, which is only entered without hesitation, $ihesitate=0$. Immediately after, still only without hesitation, the right hand side is reduced. These reductions follow the illustration and corresponding descriptions to Fig. 6.2b. The core loop (paragraph above) will now reach the update of k_{max} and j_{max} without hesitation.

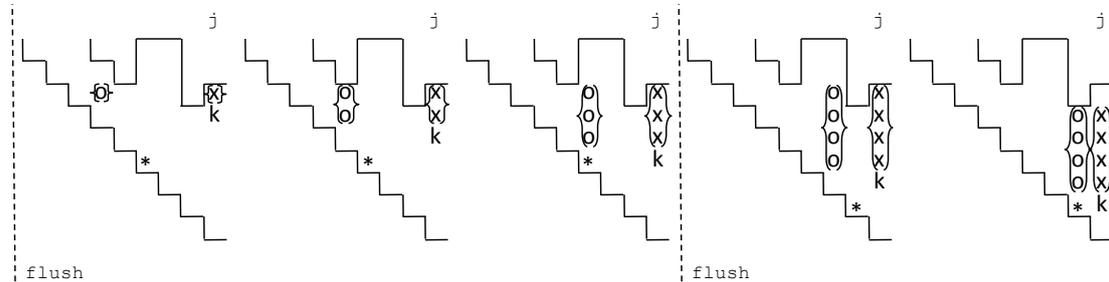


Fig. 6.6. Example of reduction of off-diagonal positions with hesitation. The “*” indicates latest processed diagonal position at latest flush.

Reduction of diagonal term and right hand side (case: $j_h=2$)

The case where the active height of column j is two is a special case of the above. There are no off-diagonals to process before reduction of the diagonal, so the algorithm is simplified to the following:

```

iwait=1
do while (iwait.eq.1)
!$OMP flush (jmax,kmax)
  if (kmax.ge.jr-1) then
    d=skmatx(jr+1)
    skmatx(jr+1)=skmatx(jr+1)/skmatx(jr)
    skmatx(jd)=skmatx(jd)-d*skmatx(jr+1)
    iwait=0
  endif
enddo
fmatx(j)=fmatx(j)-skmatx(jr+1)*fmatx(j-1)
ihesitate=0

```

A parameter, $iwait=1$, is defined and a loop will continue until this is set to $iwait=0$. The loop keeps flushing $kmax$ until column $j-1$ has been processed. Hereafter, the diagonal reduction and reduction of the right hand side take place as above. The flags are set to $iwait=0$ and $ihesitate=0$ to proceed without hesitation to update $kmax$ and $jmax$ in the end of the core loop.

Source code

The parallel skyline solver is listed in Appendix A with the above parallel factorization of the system matrix and reduction of the right hand side. The remaining tasks are sequential as the time spent on these are limited. These tasks are the division of the right hand side by system matrix diagonals and the backward substitution.

6.3 Comparison of skyline solver with other solvers

The parallel skyline solver is compared with a band solver and an iterative solver. The band solver is a direct solver as the skyline solver, but it works on a system matrix stored in band form shown in Fig. 6.1b by the dashed line and many zeros are therefore stored and processed compared to the skyline storage format decreasing the overall efficiency. The iterative solver included in the comparison is based on the conjugate gradient method with preconditioning; see more details provided by Fernandes and Martins [11]. All simulations are performed on a Dell Optiplex 980 desktop with an Intel(R) Core(TM) i7-860 processor with four cores and eight threads. It has 8GB RAM, 8MB cache and a clock frequency of 2.8GHz. The system is 64-bit, but the program is running in 32-bit. The operating system is Windows 7. In order to keep the computer under the same global workload when testing the solution speed, all eight threads have been active during all simulations. When testing solution time using N threads, the remaining $8 - N$ threads have been running similar dummy simulations.

Fig. 6.7a shows the test case used in the comparison of different solvers. The test is simple upsetting of a cube between two flat parallel platens. Two of the cube faces have prescribed symmetry, and contact between the cube and the tools is frictionless. The cube with dimensions $10 \times 10 \times 10 \text{mm}^3$ is compressed to half height through 100 simulation steps of $\Delta t = 0.05 \text{s}$ with a velocity $v = 1 \text{mm/s}$. The cube with material described by the flow stress curve $\sigma = 180.65 \varepsilon^{0.183} \text{MPa}$ is discretized by e^3 8-node isoparametric elements of equal initial size. This discretization implies $(e + 1)^3$ nodes and $3(e + 1)^3$ degrees of freedom with three unknown velocity components per node.

Fig. 6.7b shows the solution time as function of the number of degrees of freedom. The solution time is normalized by the solution time of the parallel skyline solver using eight threads. As expected, the band solver is much slower than the other solvers, and having the other solvers available, the band solver becomes outdated. Among the skyline solvers, the solution time is ideally halved when going from sequential (one thread) to two threads, from two to four threads, and from four to eight threads. The solution time is not completely halved since the program is not 100% parallel due to the waiting time described in the end of Section 6.2, other tasks than equation solving, and due to overhead. The iterative solver has solution time comparable with the skyline solver. Comparing to the parallel skyline solver using eight threads, the iterative solver is slower below approximately 20,000 degrees of freedom, and above it is faster. When using fewer threads in the skyline solver, this separation number of degrees of freedom is smaller. On the other hand, if more threads were available, the parallel skyline solver would be faster than the iterative solver at even larger numbers of degrees of freedom.

When the solution times of the iterative solver and the parallel skyline solver are in the same range, the iterative solver has the benefit that other threads are still available for other computations. However, the skyline solver has the benefit of being direct, implying better accuracy than the iterative. Fig. 6.8 shows an example of the accuracy

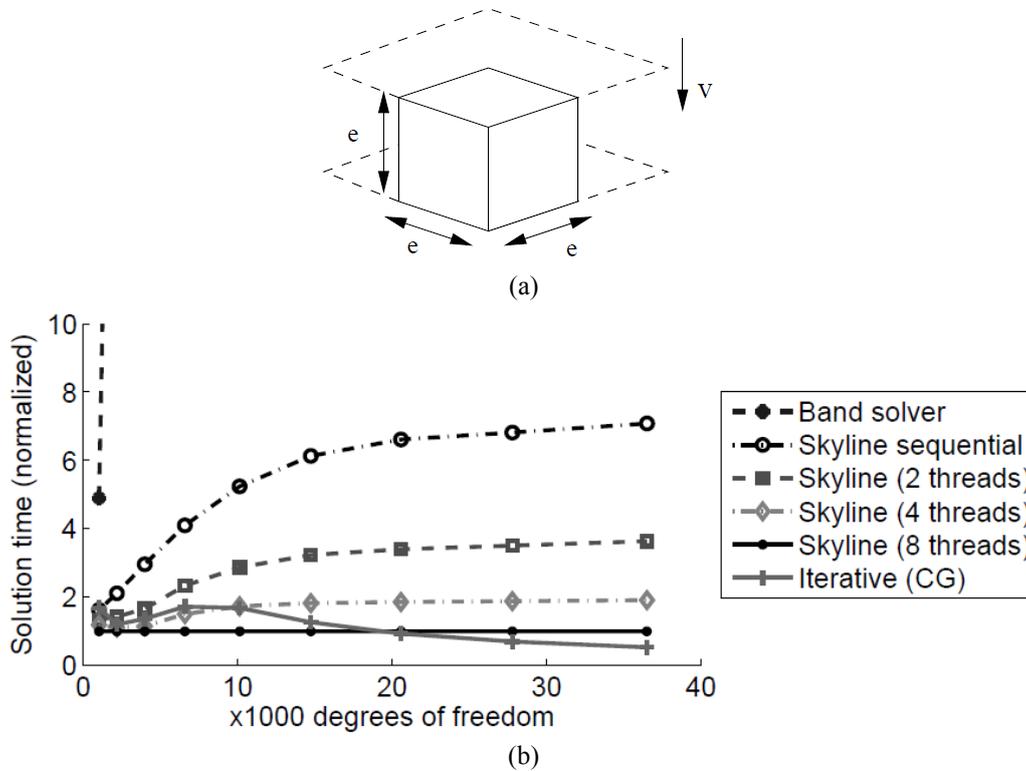


Fig. 6.7. Comparison of band solver, iterative solver and parallel skyline solver. (a) Simple upsetting test case. (b) Normalized solution time as function of degrees of freedom. The solution time is normalized by the solution time of the skyline solver using eight threads.

differences between the solvers. The vertical stress component is shown on the cube after compression to half height. The resulting stress distribution when applying the iterative solver varies as shown in Fig. 6.8a between $-166.3MPa$ and $-171.4MPa$, whereas the distribution when applying the direct skyline solver is uniform with a value of $-168.8MPa$. The deviations in the result from the iterative solver are -1.48% and 1.54% relative to the result from the direct solver. Observations of the iterative solver showed that the solution did not converge in any of the 100 steps. Decreasing the tolerance for convergence would therefore not have any effect. Instead, the limit on the number of iterations was removed to ensure convergence (if possible). The average number of iterations was increased about 1.20 times, and the solution of the equation system converged in all steps. The above deviations reduced to -0.178% and 0.237% . However, due to the increased number of iterations, the solution time increased about 13%, and the longer solution time for the iterative solver makes it less attractive than it appears in Fig. 6.7b. For more ill-conditioned systems of equations (e.g. due to rigid zones or contact between deformable bodies), the increase of iterations and solution time would be even more.

The differences between the results in Fig. 6.8, even when the iterative solver converges, become crucial when analyzing more complex geometries including contact between deformable bodies. On top of accuracy problems, the iterative solver may be-

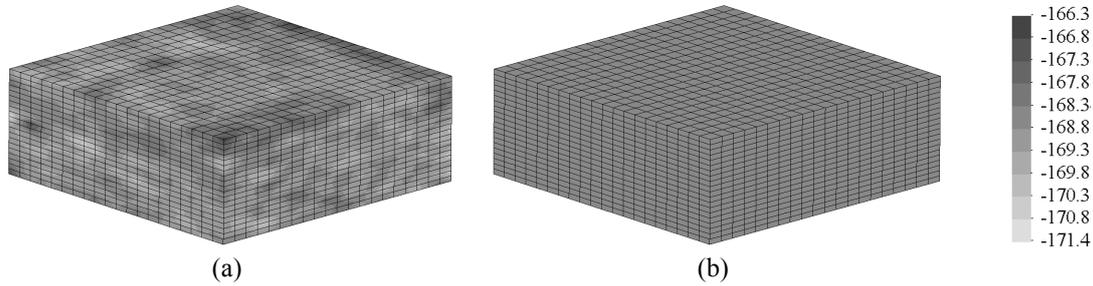


Fig. 6.8. Vertical component of the stress field in the cube compression example with 20 elements along each side. (a) Solution by iterative solver. (b) Solution by direct skyline solver (any number of threads). Common scale bar takes the minimum and maximum values according to (a).

come unstable when dealing with ill-conditioned equation systems; see Farhat and Wilson [9] and Fernandes and Martins [11]. Ill-conditioned equation systems are likely to appear when penalty methods are applied as in the present computer program. The skyline solver has therefore been adopted as the standard solver, and after the parallelization, the solution time is not further minimized by an iterative solver for the majority of system sizes dealt with.

6.4 Performance evaluation of parallel skyline solver

Speed-up, efficiency and parallel fraction are evaluated based on the compression of a cube to half height presented in Section 6.3. The speed-up (ratio of the solution time on one thread T_1 to the solution time on N threads, T_N),

$$\sigma = \frac{T_1}{T_N} \quad (6.6)$$

is ideally N , but because the finite element program is not entirely parallel, the actual speed-up shown in Fig. 6.9 as function of degrees of freedom is less than ideal. Part of the program is still sequential, since only the equation solver of the main system of equations has been parallelized, and in addition heading (physical communication to and between the threads) takes time. As the system size (degrees of freedom) increases, relatively more time is necessary to solve the equation system, which means that the fraction of the code running in parallel becomes relatively larger. This results in the larger speed-up seen in Fig. 6.9 at increasing number of degrees of freedom. It is also seen in the figure that the speed-up is largest for the smaller number of threads. This is a result of increased heading time and increased waiting time between threads when more threads are used, but it is also a result of a relatively smaller time fraction being parallel, simply because the amount of solution time with more threads is less compared to the overall time.

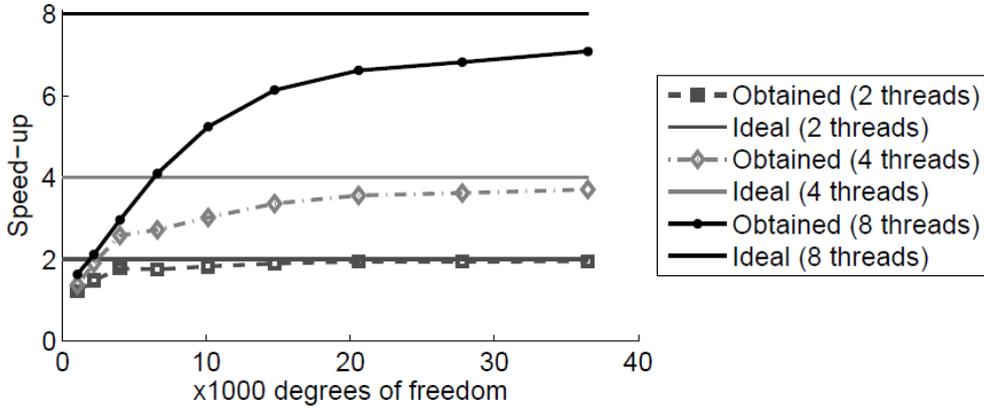


Fig. 6.9. Speed-up for 2, 4 and 8 threads as function of degrees of freedom.

Similar conclusions can be drawn from Fig. 6.10a, which shows the efficiency,

$$\eta = \frac{\sigma}{N} \quad (6.7)$$

defined as the ratio of the actual speed-up to the ideal speed-up. The efficiency in Fig. 6.10a is shown as function of degrees of freedom for the different number of applied threads. Fig. 6.10b shows the parallel fraction, which is defined through Amdahl's law (originating from Amdahl [12]),

$$\tilde{\sigma} = \frac{1}{(1-P) + \frac{P}{N}} \quad (6.8)$$

in which P is the fraction of the program being parallel. Amdahl's law (6.8) estimates the speed-up $\tilde{\sigma}$ for a certain number of threads N based on the sequential contribution $1-P$ and the parallel contribution $\frac{P}{N}$. Rearranging allows the estimation of the parallel fraction from the actual speed-up as follows based on actual speed-up σ ,

$$P = \frac{1 - \frac{1}{\sigma}}{1 - \frac{1}{N}} \quad (6.9)$$

The parallel fraction depicted in Fig. 6.10b shows a steep increase in the beginning and then flattens out as the sequential part becomes small. In the end of the curve, the parallel fraction has reached 97-98% for all number of threads. Note that this is of the entire program, not only the skyline solver, which means that the CPU time taken in the rest of the program is negligible. Insertion of this parallel fraction into (6.8) shows that this corresponds to potential speed-up of 33-50 times if sufficient threads were available. This is found by letting the number of threads N go to infinity as this will

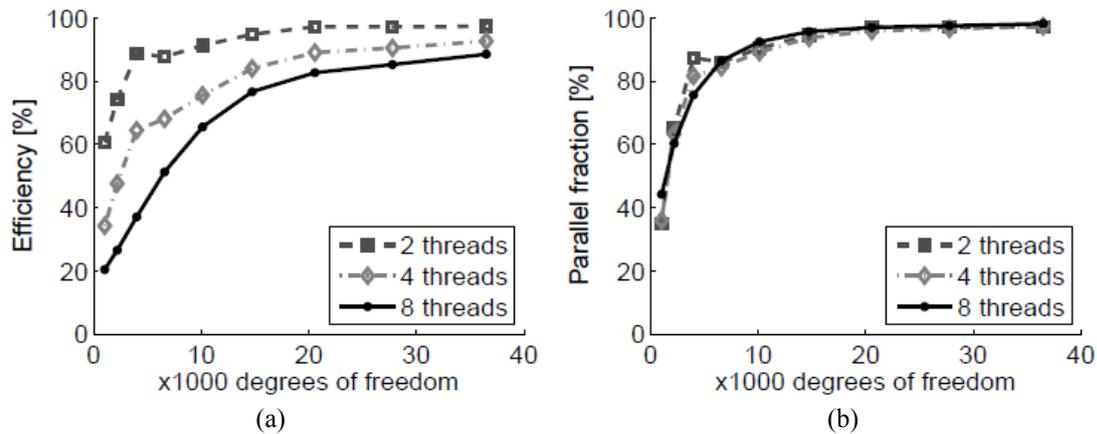


Fig. 6.10. Efficiency (a) and parallel fraction (b) for 2, 4 and 8 threads as function of degrees of freedom.

predict an upper limit of the speed-up. Assuming a parallel fraction of 97.5%, the maximum speed-up is 40 and the speed-up as function of applied threads is given by Amdahl's law, (6.8). These are plotted in Fig. 6.11 together with the achieved speed-up using one (trivial), two, four and eight threads on the cube example with 36,501 degrees of freedom. This figure shows the potential of the parallel skyline implementation as it must be expected that more threads will be available on standard PCs in future. The computer used in the present work has eight threads. This corresponds to an estimated speed-up of 6.8 (the achieved was 7.1).

A standard PC with 16 threads may not be far away, and in this case the estimated speed-up is 11.6. Dreaming further to reach e.g. 32 and 64 threads in standard PC's, the estimated speed-up is 18.0 and 24.9, respectively. Since the curve flattens out, it is also clear that if e.g. two simulations are to be run, it is more efficient to run the two simulations simultaneously sharing the available threads rather than running one after the other using all threads.

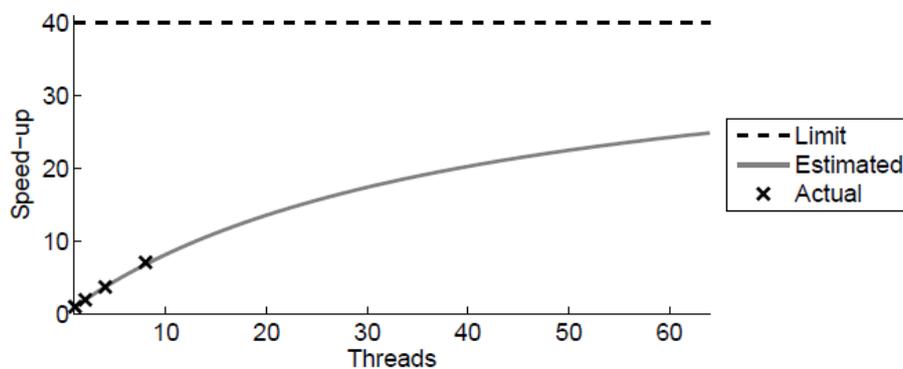


Fig. 6.11. Estimated speed-up as function of applied threads shown together with the theoretical limit and the actual speed-ups achieved for the cube compression example with 36,501 degrees of freedom.

6.4.1 Evaluation by a resistance welding case

The parallel skyline solver is also tested for an industrial case by evaluating the solver in simulation of resistance welding with different number of threads. The welding case is shown in Fig. 6.12a and consists of two AISI 1008 steel alloy sheets of 1mm thickness that are spot welded between two copper alloy electrodes with tip diameter $\varnothing 6\text{mm}$.

The electrode center axes are placed in a distance 13mm to three of the sheet edges, but only 4mm from the fourth edge. Total simulated process time is 340ms . The electrode force is raised linearly to 3kN within 20ms and kept constant hereafter. AC current is applied after 40ms , lasting 200ms at a level of 8kA RMS with a conduction angle of 80% . After the current is turned off, the electrode forces are kept for additionally 100ms while the weld nugget solidifies. The mesh shown in Fig. 6.12b consists of $7,666$ nodes giving rise to $22,998$ degrees of freedom in the mechanical model and $7,666$ degrees of freedom in the electrical and thermal models. Fig. 6.12c shows

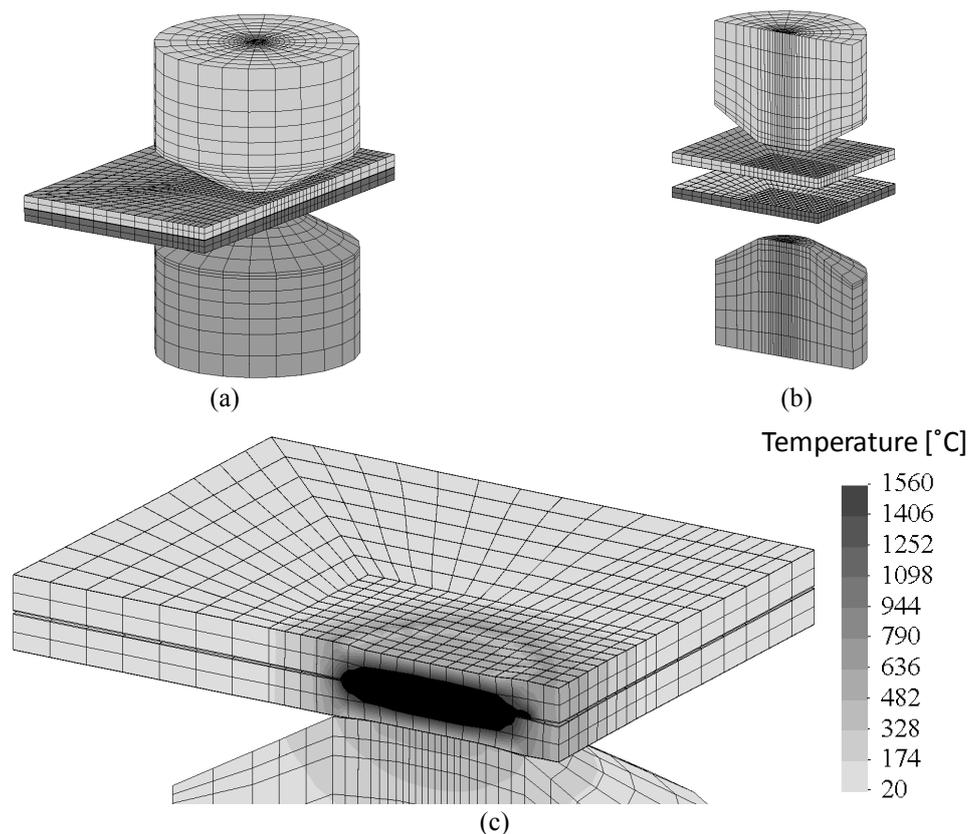


Fig. 6.12. Resistance spot welding test case. (a) Arrangement of electrodes and sheets for testing spot welding near an edge. (b) Applied mesh using symmetry. Total number of nodes is $7,666$. (c) Temperature field (shown without upper electrode) in the end of the weld time.

the resulting temperature field after the applied welding time. The spot seems almost axisymmetric showing that the chosen distance to the edge may not be a problem. However, this is without analysis of splash, which may be determining. Due to less material on the edge side, the temperature decreases slower near the edge and this asymmetric cooling may result in a microstructure and residual stress distribution that the welding engineer has to be aware of.

The effect of node numbering optimization in the overall sparsity of the resulting finite element stiffness matrix is shown in Fig. 6.13, where the skyline height as function of the column number is shown before and after contact between the objects in Fig. 6.12b. The skyline heights are shown for three different cases. Fig. 6.13a-b show the heights without node numbering optimization, Fig. 6.13c-d show the resulting heights with node numbering optimization without information of contact between objects, and Fig. 6.13e-f show the resulting heights with node numbering optimization with initial contact. The straight lines in the figures show the height corresponding to a full matrix, i.e. to a full upper triangular matrix due to symmetry. This shows, for all the cases in Fig. 6.13, the importance of an efficient storage format, as the skyline format applied here.

Without considering contact, i.e. in the configuration shown in Fig. 6.12b, the immediate benefit of node numbering optimization is seen by comparing Fig. 6.13a and Fig. 6.13c. Fig. 6.13a shows the skyline height of the original mesh, and Fig. 6.13c shows the skyline height of the optimized mesh, where the number of matrix positions below the skyline is more than halved. Of more interest for the calculation time is the skyline height after the objects have been brought into contact. When the optimization is made without considering contact, peaks will typically appear as in Fig. 6.13d when the objects get in contact due to expansion of the skyline as presented in relation to Fig. 4.4. The differences in node numbers between the contacting nodes are large, and hence the skyline heights peak due to contact. Comparing Fig. 6.13d to Fig. 6.13b, it is seen that the number of matrix positions below the skyline is still reduced to about 72%. Fig. 6.13e-f show the resulting skyline profile when optimization of node numbering is made with inclusion of the initial contact information. The number of matrix positions is not improved as much as before (in fact the number increases before contact and only decreases slightly with contact). However, the peaks due to contact are reduced significantly, which will contribute to better speed-up in case of parallel solving of the equation system.

In case of contact, e.g. Fig. 6.13d, the advantage of the skyline storage format is huge compared to the banded storage format. The amount of stored positions in a banded format would be the positions enclosed by the straight line in the figure and the horizontal line that would cover the tallest skyline height.

The solution times and obtained speed-ups in the welding case are shown in Fig. 6.14 for the two approaches to node numbering optimization. Solution times and speed-ups are shown as function of applied threads, where, as in the above analyses, when applying N threads, the remaining $8 - N$ threads have been applied to a similar dummy simulation. The solution times in Fig. 6.14a and 6.14c are shown for the total running time of the entire simulation as well as for the pure solution time of the equa-

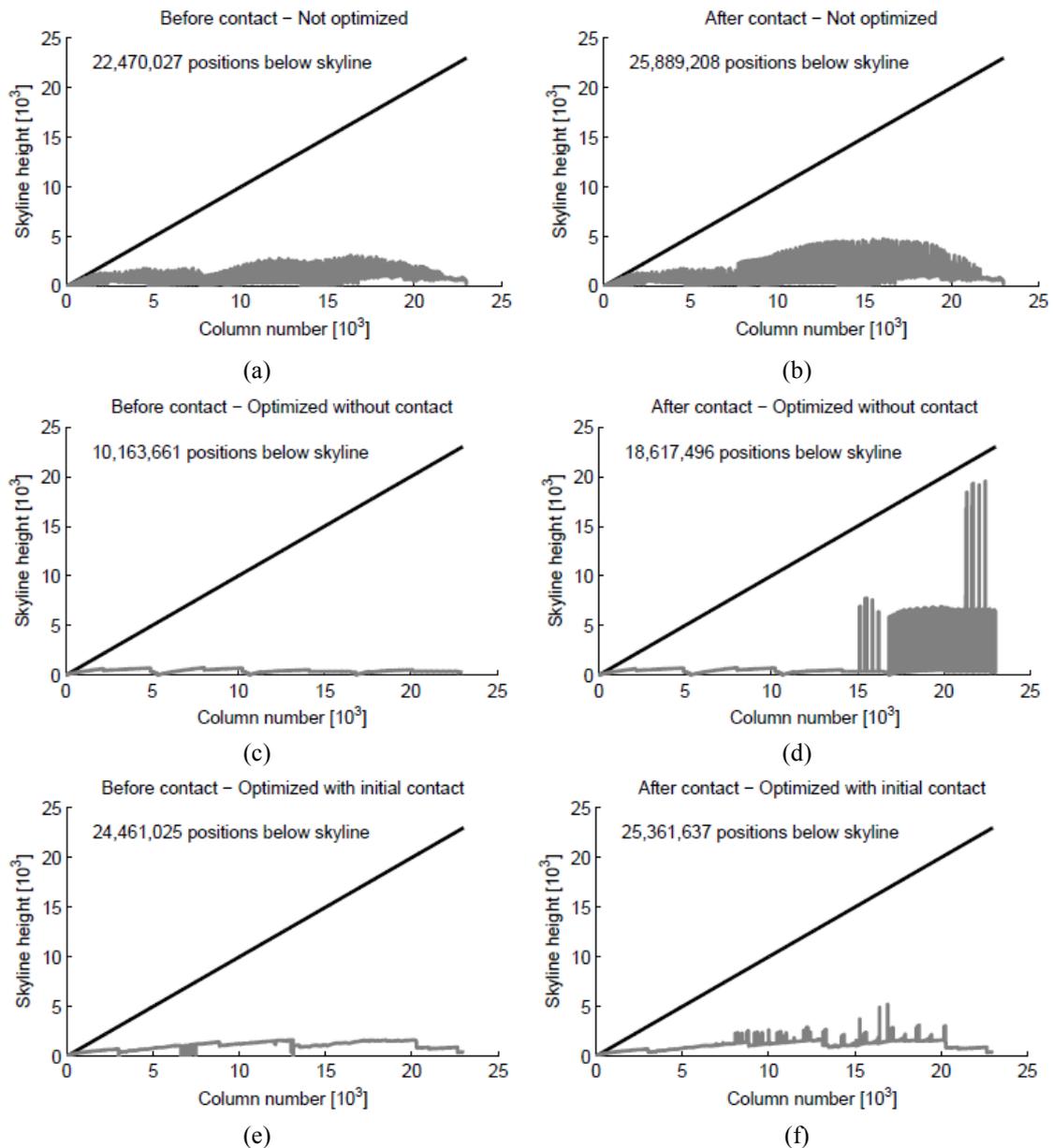


Fig. 6.13. Comparison of skyline topology before (a,c,e) and after (b,d,f) contact depending on optimization of node numbering. (a-b) Without optimization. (c-d) With optimization independent of contact. (d-e) With optimization including initial contact. The straight lines show the height corresponding to a full matrix (halved due to symmetry). Number of matrix positions below the skyline is noted in each figure.

tion system in the mechanical model and in the electrical and thermal models. These pure solution times are accumulated over the entire simulation. It is clear from the figure that the equation solving in the mechanical model is the main contributor to the total solution time.

The combined solution time in the electrical and thermal models is much less, partly due to fewer iterations and in particular due to the smaller system size (7,666 degrees of freedom compared to 22,998 degrees of freedom in the mechanical model). The figures also include the remaining time spent in the simulation, i.e. the total time subtracted the pure solution time in the main equation systems. Thus, the remaining solution time is a sum of setting up the equation systems, searching for and evaluating contact, updating variables before time stepping, etc.

The overall solution time decreases with increasing number of applied threads. This is mainly accommodated by the shorter time spent in the pure solution of the mechanical equation system. The time spent in pure solution of the electrical and thermal equation systems decreases only little, and the time spent on remaining tasks should be unchanged, since it is not parallelized. An interesting difference is observed between the two approaches to the node numbering optimization. When the optimization is performed without information of the contact, the solution time does not decrease noticeable when applying more than three threads (Fig. 6.14a), and correspondingly the speed-up does not increase noticeable when applying more than three threads (Fig. 6.14b). On the other hand, the solution time decreases and the speed-up increases remarkably over the whole range of applied threads when node numbering optimization includes information of initial contact. The reason for this difference is explained by peaks in the skyline height due to contact, see Fig. 6.13d and 6.13f. In the case where node numbering optimization is performed without initial contact, the peaks are high and separated by shorter columns. This results in waiting time in the threads processing the shorter columns, and thereby poorer speed-up. The column heights are more equal in case of optimization with initial contact, and therefore speed-up is better preserved.

In both cases, the speed-up of the pure solution time in the mechanical model is the highest, which is a result of the larger system size (22,998 degrees of freedom) compared to the electrical and thermal system sizes (7,666 degrees of freedom). The speed-up of the total solution time is closer to that of the mechanical, since the mechanical model is the main contributor to the time. The total speed-up is slightly lower than the mechanical speed-up due to the lower speed-up in the electrical and thermal models and due to no speed-up in the remaining non-parallelized code.

When it comes to speed-up, the above comparison shows that the node numbering optimization taking initial contact into account is clearly better than the optimization without contact information. Fig. 6.14e compares the actual solution time of the two approaches for different numbers of applied threads. The solution times are normalized by the solution time of the optimization without contact information. The figure shows that the approach including initial contact (which has the better speed-up) is slower by a factor of 1.4 when using one thread. However, due to the better speed-up, it becomes faster when applying six or more threads. The reason for the slower solution when using few threads is that the initial contact is much more than the contact after separation of the sheets outside the weld zone, and therefore the optimized skyline according to the initial contact is not optimal throughout the entire solution.

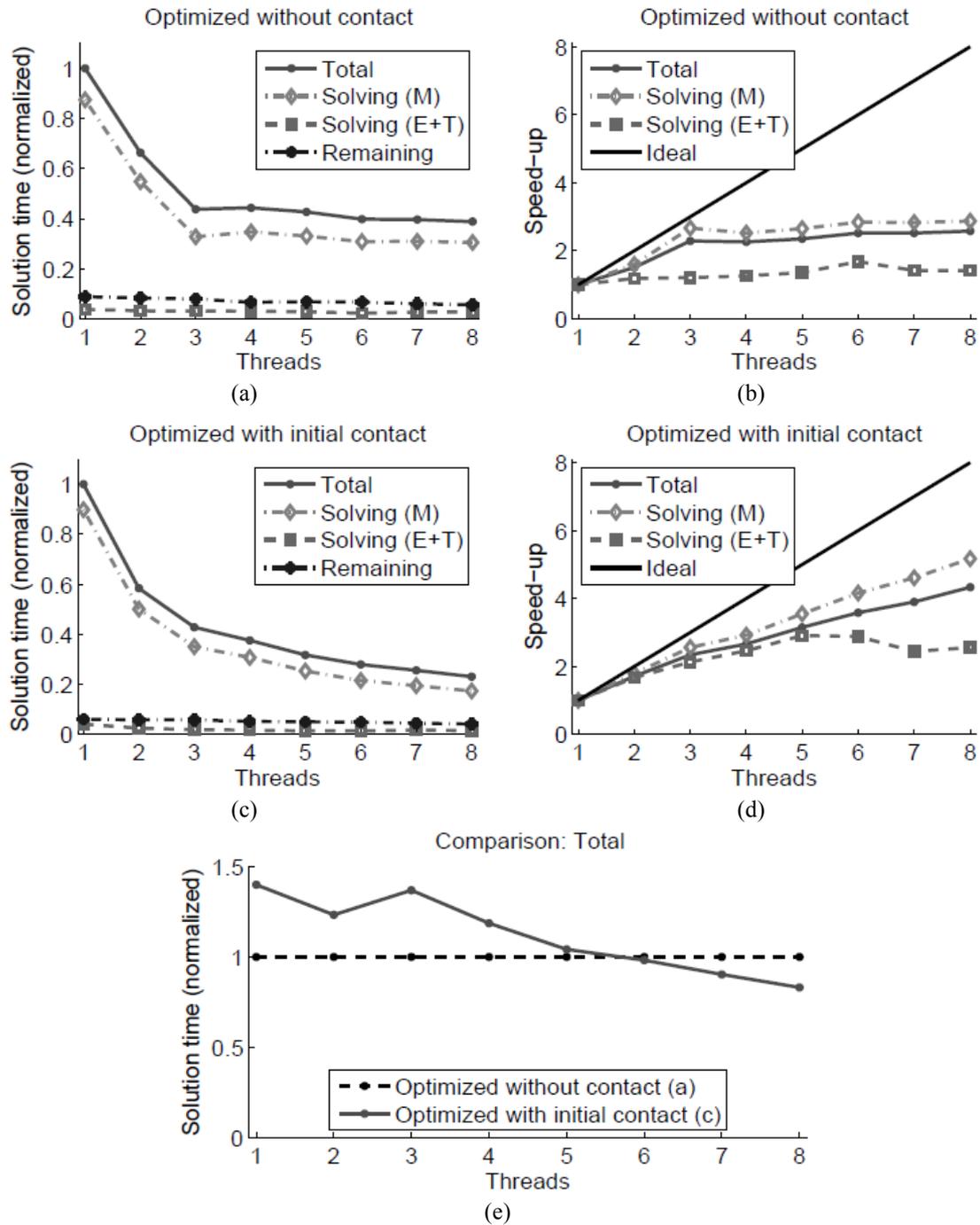


Fig. 6.14. Normalized solution time (a,c) and speed-up (b,d) as function of applied threads. (a-b) Optimization of node numbering independent of contact. (c-d) Optimization of node numbering with initial contact. (a-d) include the total solution time of the entire solution, the pure equation solving in the mechanical model (M), and the combined pure equation solving in the electrical and thermal models (E+T). The solution time of the “remaining” is also included, which is the total subtracted the pure equation solving. A comparison of the total solution times is given in (e).

Fig. 6.14c-d, as well as the figures related to the cube compression example, prove the presented parallel skyline solver. Hereafter, it is up to a correct approach for the node numbering optimization to get the best use of it. In the specific welding case, an improved strategy would be to start out with an optimized node numbering based on the initial contact, and then reoptimize the node numbering when the sheets have separated outside the weld zone.

References

- [1] El-Sayed MEM, Hsiung C-K (1990) Parallel finite element computation with separate substructures. *Computers & Structures* 36(2):261-265.
- [2] Farhat C (1988) A simple and efficient automatic fem domain decomposer. *Computers & Structures* 28(5):579-602.
- [3] Al-Nasra M, Nguyen DT (1992) An algorithm for domain decomposition in finite element analysis. *Computers & Structures* 39(3/4):277-289.
- [4] Nielsen CV, Martins PAF, Parallel skyline solver: Implementation for shared memory on standard personal computers. Submitted for publication in international journal.
- [5] Nielsen CV, Zhang W, Alves LM, Bay N, Martins PAF (2012) Modeling of Thermo-Electro-Mechanical Manufacturing Processes with Applications in Metal Forming and Resistance Welding. Springer (in press).
- [6] Lanczos C (1952) Solution of systems of linear equations by minimized iterations. *Journal of Research of the National Bureau of Standards* 49(1):33-53.
- [7] Hestenes MR, Stiefel E (1952) Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards* 49(6):409-436.
- [8] Meijerink JA, van der Vorst HA (1977) An iterative solution method for linear systems of which the coefficient matrix is a symmetric m-matrix. *Mathematics of Computation* 31(137):148-162.
- [9] Farhat C, Wilson E (1987) A parallel active column equation solver. *Computers & Structures* 28(2):289-304.
- [10] Synn SY, Fulton RE (1995) The performance prediction of a parallel skyline solver and its implementation for large scale structure analysis. *Computing Systems in Engineering* 6(3):275-284.
- [11] Fernandes JLM, Martins PAF (2009) Robust and effective numerical strategies for the simulation of metal forming processes. *Journal of Manufacturing Technology Research* 1(1/2):21-36.
- [12] Amdahl GM (1967) Validity of the single processor approach to achieving large scale computing capabilities. *AFIPS Joint Computer Conferences*, pp. 483-485.

7. Material Characterization

Awareness and understanding of the basic procedures to determine the flow stress, the frictional response and the electric and thermal contact resistances under different conditions of strain-rate and temperature are fundamental for improving the quality of data to be inserted in finite element computer programs. Because accuracy and reliability of numerical simulations are critically dependent on input data, the following sections will provide a brief overview of the most widespread experimental techniques that are utilized for material, friction and contact characterization.

7.1 Mechanical properties at room temperature

From a metal forming point of view, the most important data for modeling material behavior is the flow curve because it characterizes strain-hardening and determines the force and work requirements of a process as well as the relative material flow. In case of cold forming, the flow curve should be available to strain levels above “1” for bulk metal forming, and up to “1” for sheet metal forming processes.

The compression test performed on solid cylinder specimens is one of the most widespread mechanical testing methods for determining the flow curve in the field of metal forming. The capability of evaluating material response to much larger strains than in tensile tests, due to the absence of necking, in conjunction with the aptitude to better emulate the operative conditions of real forming processes, such as forging, rolling and extrusion, which are carried out under high compressive loads, are seen as the main reasons for its extensive utilization.

The compression test is performed by axially pressing a solid cylinder specimen between two flat polished, well lubricated, parallel platens and the flow curve is determined by combining the experimental values of force and displacement. A variant of the compression test is utilizing Rastegaev specimens, see Lange [1] and illustration in Fig. 7.1a, to reduce friction towards the platens by having a reservoir for the lubricant. This reduces barreling effectively, but leads to errors in measuring the height of the specimens due to bending of the surrounding walls and end faces not remaining plane. Fig. 7.1b shows an example of a flow stress curve by tabulated data giving a best fit of the measured data, in this case best fit of six repetitions. Molykote DX paste was utilized as lubricant in the specific example.

The flow curve is in many cases approximated by fitting curves for easy description of the material, e.g. in finite element programs. Two typical approximations are shown in Fig. 7.1b by the Hollomon (7.1) and Swift (7.2) equations,

$$\sigma = C\varepsilon^n = 777\varepsilon^{0.243} [MPa] \quad (7.1)$$

$$\sigma = C(B + \varepsilon)^n = 775(0.012 + \varepsilon)^{0.243} [MPa] \quad (7.2)$$

where C is the flow stress at strain $\varepsilon = 1$ or $B + \varepsilon = 1$, n is the strain-hardening exponent and B corresponds to a pre-straining. Both fitted curves are representing the overall behavior, but details like the yield point phenomenon existing in low-carbon steels, see detail in Fig. 7.1b, cannot be captured by such approximations. Due to the additional parameter in terms of the pre-strain, the Swift equation provides a better approximation, but only taking the details near the yielding point in an average sense. To overcome this problem, computer programs not only include more sophisticated flow stress models (e.g. Johnson-Cook and Preston–Tonks–Wallace, among others) as they have the option of including tabulated data, such that the actual material response can be modeled.

The solid cylinder specimens utilized in the compression test are limited within the aspect ratio range $1 \leq h_0 / d_0 \leq 3$ of the height h_0 to the diameter d_0 , Gunasekera et al. [2] and Czichos et al. [3], though practically not exceeding $h_0 / d_0 = 1.5$. The upper limit on the aspect ratio prevents failure by buckling or bending while the lower limit is commonly justified by the increased sensitivity to friction along the contact interface with compression platens (Alves et al. [4]), by technical difficulties to operate exten-

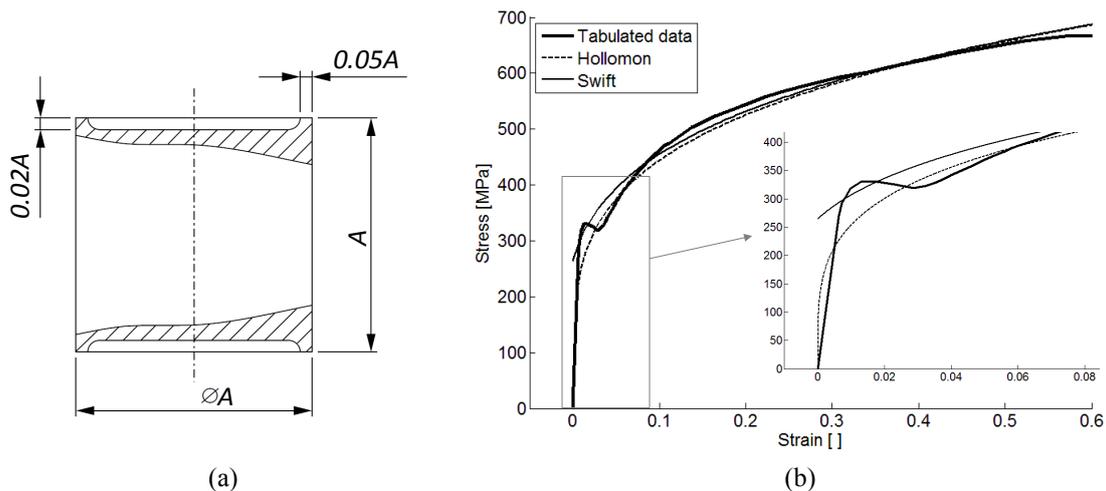


Fig. 7.1. Material testing by upsetting of Rastegaev specimens. (a) Geometry of the Rastegaev's compression test specimen. (b) Experimental stress-strain curve for a structural steel S235JR+AR and approximations by Hollomon and Swift curves. The size of the test specimens is defined by $A = 20mm$.



Fig. 7.2. Conventional and stack compression tests of Aluminum AA2011-O.

someters directly on the specimens, House [5], or not having enough displacement at all compared to the uncertainty of the measurement. This inhibits the utilization of the compression test for constructing the flow curve of materials available in form of sheets and plates.

As discussed by Alves et al. [4], the stack compression test proposed by Pawelski [6] is the best alternative experimental procedure for evaluating the flow curve of raw materials supplied in form of sheets and plates. The test makes use of circular discs that are cut out of the blanks and stacked to form a cylindrical specimen with an aspect ratio in the range of solid cylinders employed in the conventional compression test (Fig. 7.2).

As shown in Fig. 7.2 the stack compression test can be utilized for the construction of flow curves, although the procedure is not standardized. The resulting flow stress is nearly identical to that obtained by means of conventional compression tests.

However, it is worth noting that compression (as well as tensile) tests are performed under proportional loading while metalworking processes often involve non-proportional or cyclic loading. During non-proportional loading, the strain path influences the flow stress behavior as discussed by Huml and Lindegren [7] for cyclic loading and shown by Tekkaya and Martins [8] in finite element modeling of fullering with intermediate 90° turning of the specimen in-between two blows. The simulation was able to model the load-displacement response accurately in the first blow, but not as accurate in the second blow due to induced anisotropy. This is important when analyzing multi-stage processes with different loading paths in each stage because uniaxial material testing (under proportional loading) can be insufficient for accurate modeling of such cases.

Flow curves for a large number of materials can be found in Doege et al. [9].

7.2 Friction characterization

Part of the characterization of frictional behavior is the recognition of levels of normal pressure and corresponding selection of friction model. Amonton-Coulomb's law,

$$\tau_f = \mu p \quad (7.3)$$

is prone to overestimate the friction in metal forming because of the high normal pressures typically involved. On the other hand, the constant friction law,

$$\tau_f = mk \quad (7.4)$$

may also overestimate the friction in regions of low normal pressure because it does not take into account the actual stress state. Wanheim and Bay [10] have proposed a general friction model resembling the two laws at low and high normal pressures and providing a smooth transition in-between, see Fig. 7.3. The model,

$$\tau_f = f\alpha k \quad (7.5)$$

is based on slipline analysis calculating the ratio between real and apparent area of contact α between a rough workpiece surface and a smooth tool surface assuming the friction stress in the real area of contact τ_r to be constant and a fraction f of the material shear flow stress k ,

$$\tau_r = fk \quad (7.6)$$

where $0 \leq f \leq 1$. The curves are determined by discrete points but later put on formula [11]. It should be pointed out that although the model in principle solves the problem of describing friction in the entire interval from low to high normal pressures, it does not account for bulk plastic deformation of the subsurface when calculating the real

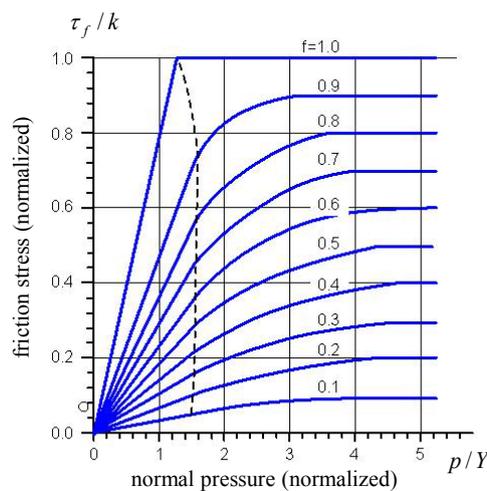


Fig. 7.3. Normalized friction stress versus normalized normal pressure with friction factor as a parameter.

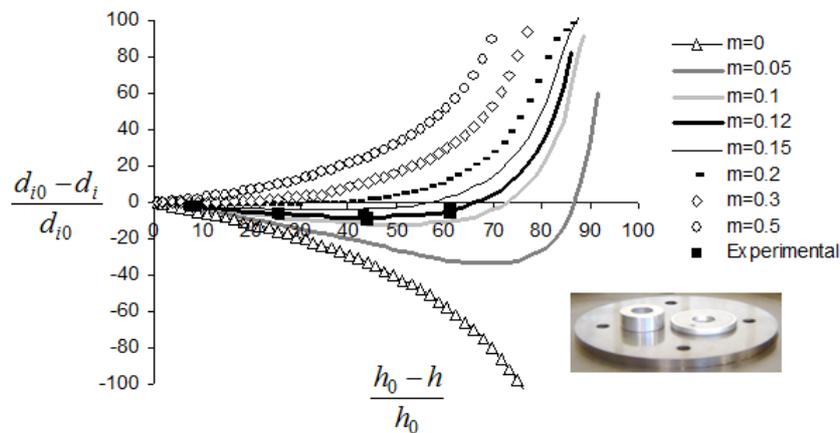


Fig. 7.4. Friction factor calibration curves obtained by finite element simulations under assumption of constant friction law and flow curve obtained from Aluminum AA1100-O. Experiments correspond to testing with lubricant Castrol Iloform PNW 124 mineral oil.

contact area. This simplification implies underestimation of the contact area and thus also friction.

As regards determination of friction data, μ , m or f , one of the well-known standard tests is the ring compression test. If calibration curves are not available, they may be constructed by finite element simulation as shown in Fig. 7.4. The ring test, however, only supplies friction data for the given contact pressure and surface expansion valid for this test. It should furthermore be emphasized that the interface temperature during testing should be correctly emulated since viscosity of many metal forming lubricants is very sensitive to temperature.

Since modeling and quantification of friction by means of simple models such as Amontou-Coulomb's law and the law of constant friction is questionable, friction coefficients or factors are in many cases tuned by the users during the numerical simulation in order to provide good estimates of the forming loads and of the deformed shape of the workpiece.

7.3 Mechanical properties at elevated temperatures

At elevated temperatures, e.g. in warm and hot forming processes or resistance welding, the flow curve is not only a function of the strain but also of the strain rate and temperature.

Fig. 7.5 presents a set of flow curves obtained experimentally by upsetting $\varnothing 8\text{mm} \times 10\text{mm}$ specimens between two flat parallel anvils at different temperatures and deformation rates. In this testing procedure, performed on Gleeble 1500 equipment, the temperatures in the specimens are controlled by sending high current pulses through the specimens to increase temperature. The temperature on the specimen surface is measured by a mounted thermocouple. The compression is performed in three inter-

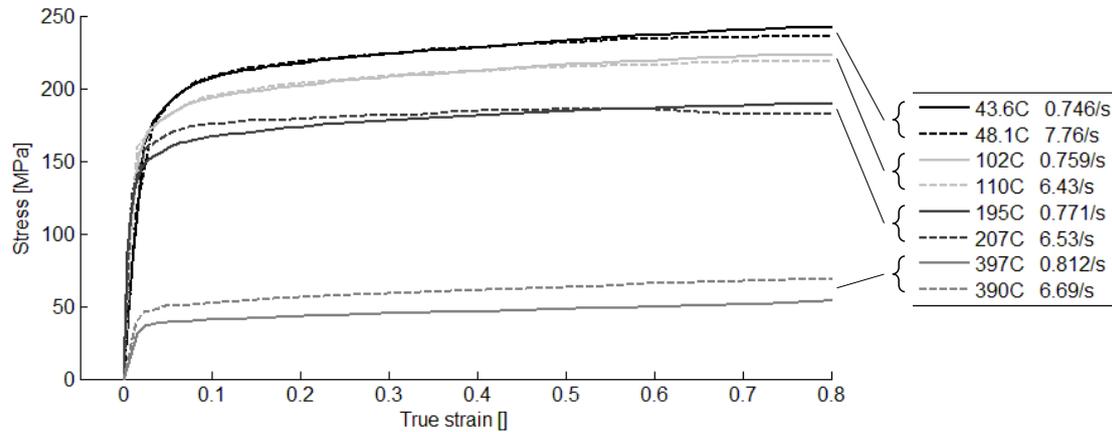


Fig. 7.5. Experimentally obtained flow curves for Aluminum AA6060-T6 at different temperatures and strain rates.

vals; an acceleration interval, a compression interval and an overtravel interval in order to obtain a strain rate during the compression interval as constant as possible. The specimen end faces are flat in contrast to the Rastegaev specimens in order to ensure proper contact to the anvils for the resistance heating. The friction is lowered by inserted graphite foils to minimize barreling. Additional corrections in the establishment of the flow curves are due to machine compliance and thermal expansion of the test specimens. This follows earlier work by Song et al. [12].

The material response represented by Fig. 7.5 is representative for many metals in terms of the lowered strength with increasing temperature. However, other responses can be identified by the testing procedure as e.g. blue brittleness in some steels, where the strength increases from room temperature to a level, say 400°C, after which the strength decreases. The effect of strain rate is furthermore available from Fig. 7.5, showing little or no influence at lower temperatures, while at higher temperatures, the material has higher strength with increasing strain rate. The range of strain rates in the example is limited and sparse. Flow curves for a large number of materials at different temperatures and strain rates can be found in Doege et al. [9].

The need to perform material characterization for higher strain and strain rates than those currently attained requires the utilization of torsion testing machines, drop hammers, Hopkinson bar apparatus and inverse analysis. Viscous effects, such as viscoplastic behavior, are usually handled by a simplified approach of specifying the flow curves as a function of the equivalent plastic strain rate. However, it is important to notice that the associated constitutive equations are time independent.

7.4 Electrical contact properties

The electrical contact resistance across an interface between metals, as described in Section 4.3, is difficult to predict and can vary significantly between batches or even

from one weld to another. Fig. 7.6 shows a test setup employed for characterization of the electrical contact resistance. Two cylindrical specimens are placed between the anvils in Gleeble 1500 equipment for the characterization of the interface between the two cylinders. The temperature is controlled by a thermocouple mounted close to the interface and high current pulses as described in Section 7.3. The contact pressure is controlled by movement of the anvils. A Rogowski coil is introduced as shown in Fig. 7.6a to measure the current of the applied pulses for heating the specimens, and the corresponding voltage drop over the interface is measured by mounted wires shown in Fig. 7.6b. Based on corresponding values of current and voltage drop, the resistance is given from Ohm's law. This follows earlier work by Song et al. [12].

The data pairs of current and voltage are selected at the time instants where the current peaks. This is to avoid the influence of induced electromotive force (emf) in the voltage measurement, which would otherwise lead to errors in the calculated resistance. The electromotive force is proportional to the first derivative of the current, and therefore supposed to vanish when the current peaks. It is furthermore proportional to the spanned area of the wires measuring the voltage drop, and the twisting of the wires seen in Fig. 7.6b is in order to minimize the spanned area.

Electrical bulk resistivity can be measured in a similar way by using only one specimen and typically increasing the length of the measured voltage drop. Obtained bulk resistivities are used to improve the calculation of contact resistance by subtraction of the resistance of the bulk material between the wires for measurement of the voltage drop. Further corrections are due to the changed cross-sectional area and distance between the wires for voltage drop measurement stemming from the compression and thermal expansion.

An example of obtained contact resistance between two specimens of stainless steel AISI 316L with end faces prepared by turning is presented in Fig. 7.7 as function of contact pressure at different temperatures. The figure shows the typical behavior of de-

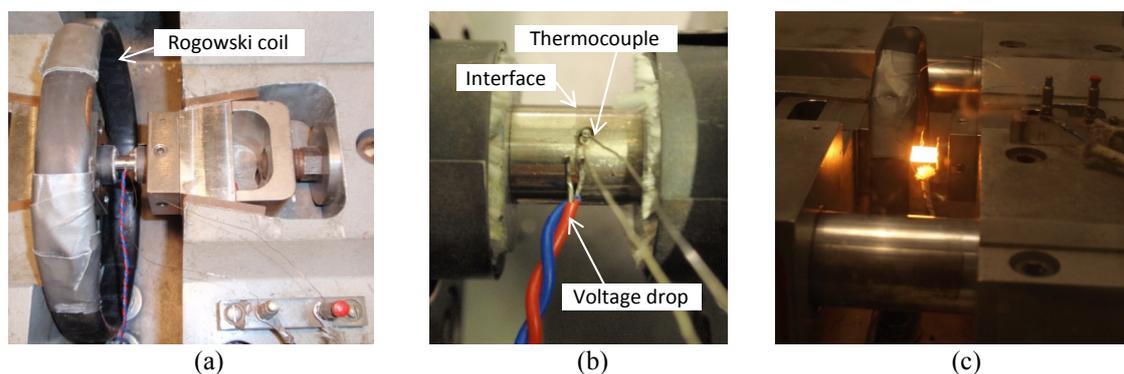


Fig. 7.6. Measurement of electrical contact resistance. (a) Test setup in Gleeble with a Rogowski coil to measure the applied current pulses. The test specimens are placed between two anvils applying a certain compression. (b) Close-up of the test specimens with mounted thermocouple for temperature measurement near the contact interface and mounted wires for measuring the voltage drop across the interface. (c) Example of testing at high temperature.

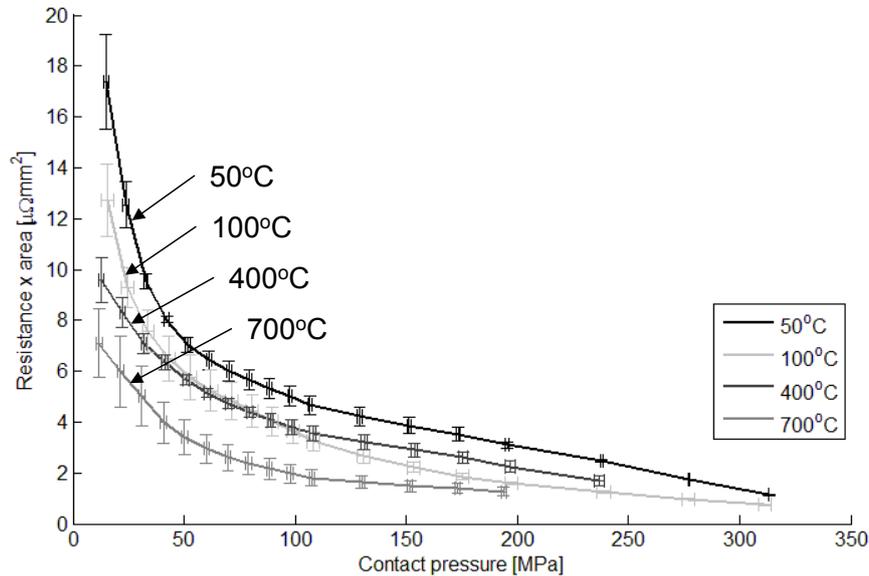


Fig. 7.7. Experimentally obtained contact resistance times contact area as function of contact pressure at different temperature levels for stainless steel AISI 316L.

creasing contact resistance with increasing contact pressure and temperature. The contact resistance is represented by the product of the contact resistance and the contact area in order to present the data independent of the contact area, which changes during testing. The relation to the finite element modeling is obtained through

$$\rho_c l_c = R_c A_c \quad (7.7)$$

where $R_c A_c$ (the product of contact resistance and contact area) is directly the presented curves and l_c is the thickness of the contact layer of elements introduced in the simulations as interface layers as described in Section 4.3. Once the thickness of the layer has been decided, the contact resistivity ρ_c is available for input to the simulation.

The contact resistivity is modeled by (4.34), where the term $\rho_{contaminants}$ stemming from the actual surface condition is used to scale the model according to the experimental curves (7.7) and Fig. 7.7.

References

- [1] Lange K (1984) Umformtechnik I. Handbuch für Industrie und Wissenschaft. Springer-Verlag, 2nd edition. (In German).
- [2] Gunasekera J, Chitty E, Kiridena V (1989) Analytical and physical modelling of the buckling behavior of high aspect ratio billets. *Annals of CIRP* 38:249-252.

- [3] Czichos H, Saito T, Smith L (2006) Handbook of materials measurements methods. Springer, Berlin.
- [4] Alves LM, Nielsen CV, Martins PAF (2011) Revisiting the fundamentals and capabilities of the stack compression test. *Experimental Mechanics* 51:1565-1572.
- [5] House JW (2000) Testing machines and strain sensors in mechanical testing and evaluation. ASM International, Materials Park, pp. 225-227.
- [6] Pawelski O (1967) Über das stauchen von holzylindern und seine eignung zur bestimmung der formänderungsfestigkeit dünner bleche. *Arch Eisenhüttenwes* 38:437-442. (In German).
- [7] Huml P, Lindegren M (1992) Properties of cold-formed metal products. *Annals of CIRP* 41(1):267-270.
- [8] Tekkaya AE, Martins PAF (2009) Accuracy, reliability and validity of finite element analysis in metal forming: a user's perspective. *Engineering Computations* 26(8):1026-1055.
- [9] Doege E, Meyer-Nolkemper H, Saeed I (1986) *Fliesskurvenatlas metallischer Werkstoffe*. Hanser Verlag München Wien, ISBN 3-446-14427-7. (In German).
- [10] Wanheim T, Bay N (1978) A model for friction in metal forming processes. *Annals of CIRP*, 27(1):189-194.
- [11] Bay N (1987) Friction stress and normal stress in bulk metal-forming processes. *Journal of Mechanical Working Technology* 14:203-223.
- [12] Song Q, Zhang W, Bay N (2005) An experimental study determines the electrical contact resistance in resistance welding. *Welding Journal* 84(5):73s-76s.

8. Verification and Applicability of Mechanical Contact Model

This chapter is focused on verification and applicability of the mechanical aspects of contact, which from a numerical point of view is more convoluted than the electrical and thermal contact models due to the vector representation of the velocity field compared to scalar representations of potential and temperature fields. From an experimental point of view, mechanical contact is more easily tested and observed than electrical and thermal contact. Mechanical contact is therefore suited for testing the numerical implementation. The chapter is divided into three sections, where Section 8.1 presents contact experiments, which by comparison to numerical simulations serve as verification of the numerical implementation in terms of geometric comparison of cross-sections along with comparisons of force-displacement curves. Applications of the mechanical contact algorithm are shown in Sections 8.2 and 8.3 in relation to real components for further verification and usability. All cases in this chapter are at room temperature.

8.1 Verification by contact experiments

A selection of contact experiments are presented in the following while comparing with numerical simulations. The examples consist of compression of pairs of two specimens of similar as well as dissimilar material combinations with geometries designed for dynamic development of the contact area as function of applied force.

8.1.1 Experimental setup

The experimental setups are schematically shown in Fig. 8.1. Both setups are designed for an existing subpress operated in a 60t universal, hydraulic Mohr&Federhaff press. The tools in Fig. 8.1 are made of PM high speed steel hardened and tempered to HRC 64 and polished to roughness $R_a \leq 0.025mm$. A $\varnothing 4mm$ hole is made in the tools to center the test specimens, which have a corresponding machined pin giving a combined slide fit (H7/g6). The tools in Fig. 8.1b have an additional groove to align the specimens perpendicular to each other. Interfaces between tools and specimens and be-

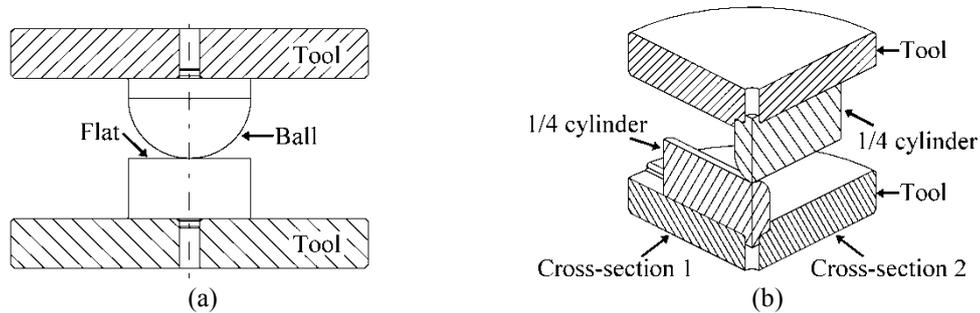


Fig. 8.1. Schematic test setups (a) for ball-flat geometry by pressing a hemisphere into a cylinder end face, and (b) for cross-cylinder similarity test.

tween specimens are in all cases lubricated with Molykote DX paste to keep friction minimal.

The presented specimen geometries include a hemisphere pressed towards a cylinder end face (Fig. 8.1a) and a similarity test of a cross-cylinder compression (Fig. 8.1b). The hemisphere specimen in Fig. 8.1a has ball radius 12mm and total height 16mm . The cylinder in Fig. 8.1a has radius 12mm and total height 12mm . In the cross-cylinder similarity test shown in Fig. 8.1b, the two specimens have width 10mm and a semi-circular edge of radius 5mm . Length and height of these specimens are $50\text{mm} \times 13\text{mm}$. All heights are excluding the centering pin.

The ball-flat geometry can be seen as a similarity test for contact development involving e.g. electrodes or projections. This setup is ideally axisymmetric and can thus be handled two-dimensionally. However, it serves as an important test for the three-dimensional code as well, where it is tested with a non-axisymmetric mesh. Hence, the capability of the program to represent symmetry is tested.

The cross-cylinder geometry is a similarity test for cross-wire welding, where the geometry represents a natural projection. The contact development for this geometry is three-dimensional, thus testing the program in the range, where a two-dimensional implementation is inapplicable.

8.1.2 Materials

Three materials are used for the test specimens. Aluminum AA6060-T6 is the softest, structural steel S235JR+AR is used as a medium material and stainless steel AISI 316L is the hardest material utilized. The materials were received in squared bars with cross-sectional dimensions $25\text{mm} \times 25\text{mm}$. Preceding processing of the material into these profiles may have introduced anisotropy due to uneven strain-hardening, which will influence the deformation path during the tests. The stress-strain behaviors of the materials are tested by standard upsetting applying the Rastegaev test illustrated in relation to Fig. 7.1. Besides obtaining the stress-strain curves, the upsetting tests were used to examine the existence of eventual induced anisotropy.

The stainless steel and the structural steel showed no effect of induced anisotropy, while the aluminum showed clear effects hereof. Fig. 8.2a shows the initial specimens which are machined from the bars with the cylinder axis along the length axis of the

bar (Fig. 8.2b) or across the length axis of the bar (Fig. 8.2c). When upsetting the aluminum specimens oriented along the bar, the cross-section changes from circular to a more squared shape as exemplified in Fig. 8.2b. When upsetting across the bars, the cross-section changes from circular to more elliptical as exemplified in Fig. 8.2c. Both tests show that induced anisotropy influences the material flow of the aluminum. Similar tests with the stainless steel and the structural steel revealed no influence as the cross-sections in both cases remain circular.

With reference to Fig. 7.1a, the dimensions of the Rastegaev upsetting specimens were $A = 24\text{mm}$ for the aluminum, $A = 20\text{mm}$ for structural steel and $A = 15\text{mm}$ for stainless steel, where the different sizes reflect their relative strengths when utilizing the full range of the press. The obtained flow stress curves are shown in Fig. 8.3 in terms of smoothed tabulated data, which directly serve as input for the numerical simulations with linear interpolation between data points.

The reason for using tabulated data prior to fitted equations, such as Hollomon or Swift equations, are the ability of tabulated data to actually follow the experimentally obtained curves, while the fitted equations can only represent the behavior in an average sense. An example is the structural steel of these tests, which is the actual material behind the curves presented in Fig. 7.1b.

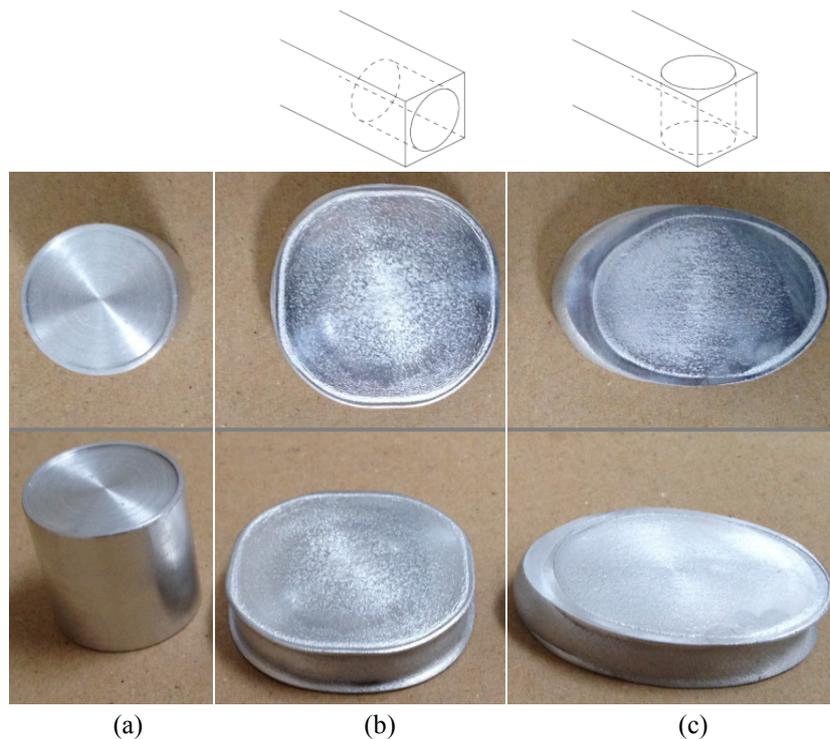


Fig. 8.2. Anisotropy in aluminum test material induced by preceding manufacturing into squared bars. Initial Rastegaev upsetting specimens (a) and deformed specimens after compression test along the length axis of the squared bars (b) and after compression across the length axis of the bars (c).

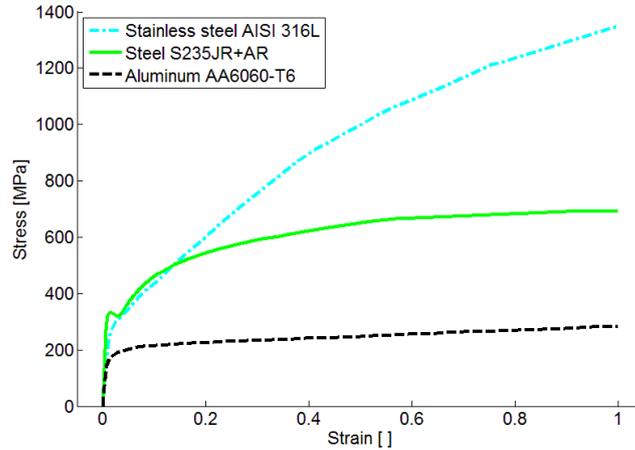


Fig. 8.3. Stress-strain curves of tested materials. Each curve represents best fit of six repetitions.

8.1.3 Comparison between simulations and experiments

All 15 combinations of test geometries and specimen materials (nine with ball-flat contact and six with cross-cylinder contact) are tested and a selection is taken out for analysis of cross-sections. Pairs of specimens compressed to certain reductions are molded in a polymer resin. They are cut and ground to make the cross-sections visible for comparison with simulated compressions. As regards the experiments, a possible error source is misalignment during compression, and on top of that is possible misalignment during molding into the resin as well as the cutting and grinding may not show the exact cross-sections.

Besides the visual comparison of geometries, force-displacement curves are compared and examined. During experiments, the force was measured by a load transducer placed in the subpress below the lower tool. The displacement was measured by an extensometer between the tools.

All simulated results are based on contact algorithm *I* (Section 4.2). Using contact algorithm *II* did not show significant changes in the presented cases.

Ball-flat contact

The geometrical comparisons for the ball-flat geometries are shown in Fig. 8.4 for three material combinations, each at two reductions. Generally, good agreement between experiments and simulations are observed. Only minor differences can be found. An example is the contact interface at high reduction in the case with two aluminum specimens (Fig. 8.4b), where the curvature of the contact interface changes direction in the experiment. The reason for this is related to the induced anisotropy proven by Fig. 8.2. The strength of the aluminum varies in the cross-section as a result of the preceding manufacturing of the aluminum into the received square profiled bars from which the specimens were produced.

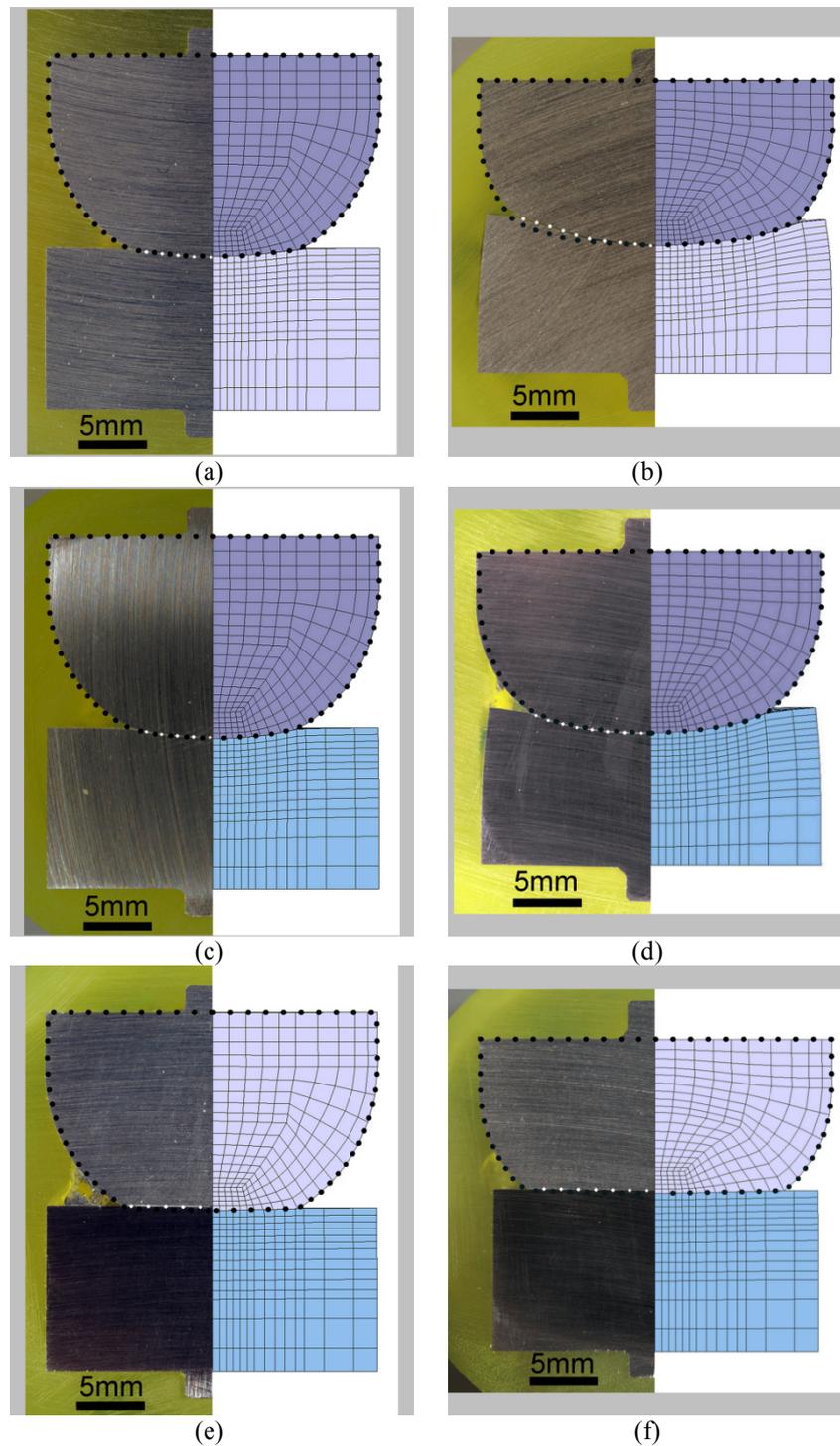


Fig. 8.4. Comparisons between experiments and simulations with ball-flat geometry for (a - b) aluminum with 6.7% - 23% reduction, (c - d) stainless steel with 7.0% - 17% reduction, and (e - f) aluminum ball and stainless steel cylinder with 5.8% - 17% reduction. Black dots follow the simulations and are mirrored to the experiments. White dots follow experimental contact interfaces.

This effect is not captured in the simulation, since the material is modeled as isotropic. The effect is shown by a photograph of a similar pair of contacting specimens of aluminum in Fig 8.5a. A couple of test specimens were heat treated prior to compression to show that the effect is absent after homogenization of the aluminum. Fig. 8.5b shows the resulting contact surfaces being smooth with monotonic curvature, which shows that the effect seen in Fig. 8.4b and Fig. 8.5a has been removed. The heat treatment of the specimens in Fig. 8.5b corresponds to annealing by heating in an oven at 470°C for one hour after the temperature has been linearly raised from room temperature during two hours. The specimens were subsequently cooled slowly inside the oven.

Force-displacement curves for the ball-flat geometry are shown in Fig. 8.6 for the three selected material combinations. Good agreement between experiments and simulations is again observed. The main difference between simulated curves (solid lines) and experimental curves (dotted lines) is the stepwise increases in the simulated curves, which is a result of the finite element discretization. The contact area increases stepwise with resolution corresponding to the size of the elements. Each time new groups of nodes get in contact, the area of contact abruptly increases with corresponding abrupt increase in the force. This phenomenon can be minimized by increasing number of elements.

Fig. 8.7 includes the remaining force-displacement curves of the experiments and simulations, which are not geometrically compared by cross-sections. These curves also show agreement between simulations and experiments. The largest deviations are found at high compression of the stainless steel and the structural steel balls into aluminum cylinder end faces (bottom curve of Fig. 8.7a and 8.7b). At large compression, the relatively harder steel balls cause the softer aluminum cylinder to deform heavily implying a large sliding area between the aluminum and the tool. The frictional contribution to the applied force was not included in the simulation, where frictionless conditions along tools were assumed. This explains the slightly larger force observed in experiments compared to the simulated force. This difference is therefore not caused by the contact implementation between the deformable objects, but by the assumption of frictionless tool contact.

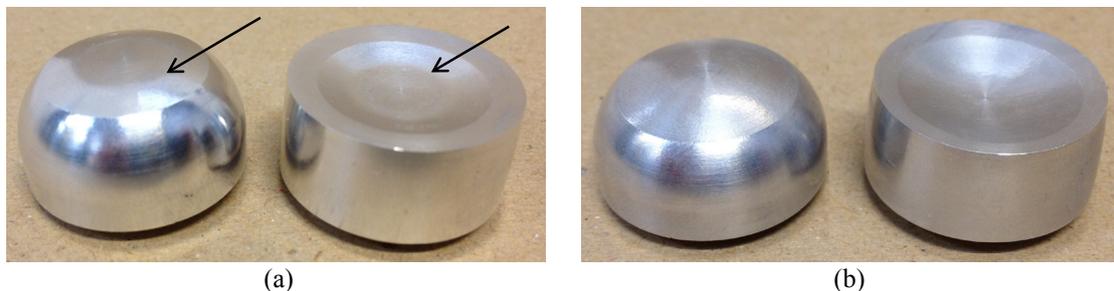


Fig. 8.5. Photographed surfaces of compressed ball-flat contact pairs in aluminum. (a) Compression of as-received aluminum corresponding to the cross-section in Fig. 8.4b. Arrows indicate the discussed effect. (b) Compressed annealed aluminum with homogeneous material properties.

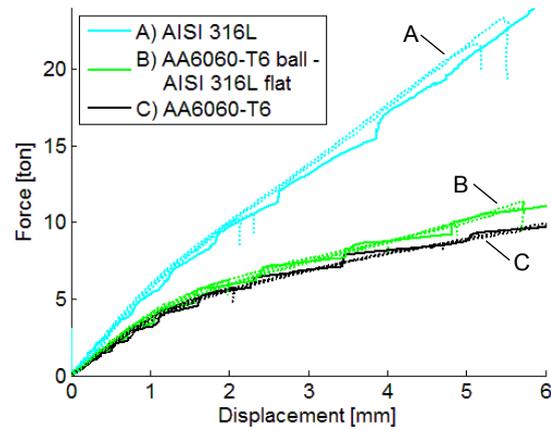


Fig. 8.6. Force-displacement curves for ball-flat geometry of the three material combinations examined by cross-sections. Solid curves correspond to simulations, while dotted curves correspond to experiments.

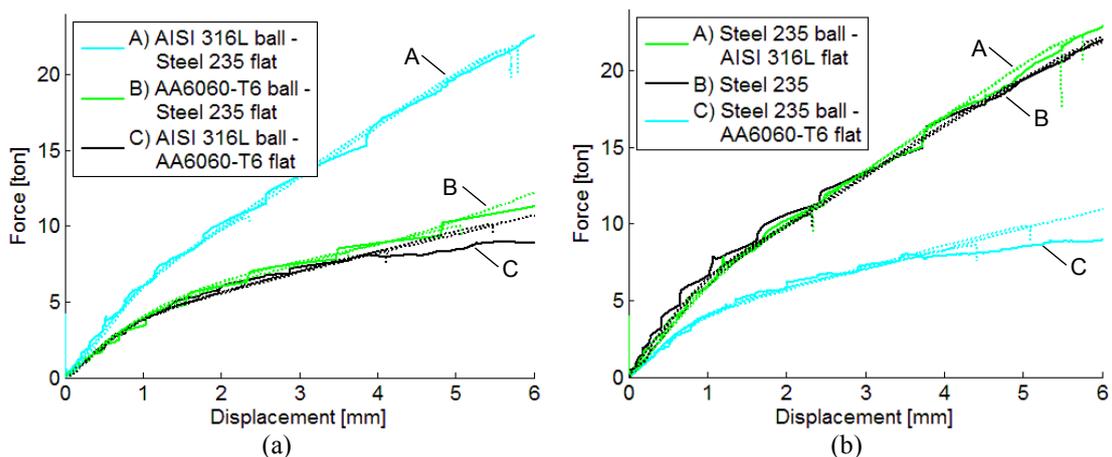


Fig. 8.7. Force-displacement curves for ball-flat geometry of the six material combinations not examined by cross-sections. Solid curves correspond to simulations, while dotted curves correspond to experiments.

Cross-cylinder contact

Three material combinations are examined by comparing cross-sections of the experiments and the simulations. Fig. 8.8 and Fig. 8.9 include the comparisons with similar materials, respectively with aluminum and structural steel S235. Both are compared at two different reductions. Good agreement between experiments and simulations are again observed. The largest difference is between the experimental and simulated position of the contact interface at the largest reduction with the aluminum specimens (Fig. 8.8c-d). The simulated result is symmetrical in the sense that the two aluminum specimens are compressed equally, whereas one specimen is compressed more than the other in the depicted experiment. This effect was seen in some experiments, but not in all, and when it appeared, there was no tendency of always being towards the upper or to-

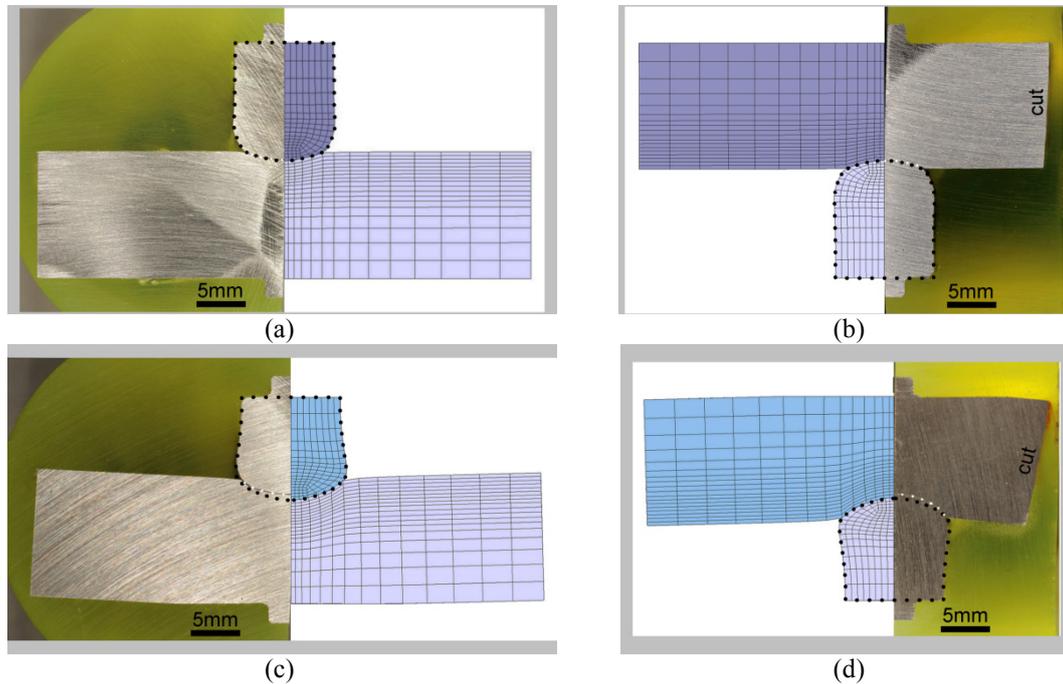


Fig. 8.8. Comparisons between experiments and simulations with cross-cylinder geometry in aluminum shown in cross-section 1-2 with (a-b) 6.9% reduction and (c-d) 19% reduction. Black dots follow the simulations and are mirrored to the experiments. White dots follow experimental contact interfaces.

wards the lower specimen. It does therefore not reveal problems of the numerical implementation, but rather uncertainty in the experiment stemming from the material.

The compression of two steel specimens (Fig. 8.9) show very good agreement in the contact interface as well as in lifting of the bulk material in regions outside the contact formation at large reduction (Fig. 8.9c).

Compression of specimens of dissimilar materials is shown in Fig. 8.10, where a structural steel specimen (upper) is pressed into an aluminum specimen (lower). The contact development is still simulated in agreement with the experiment. Lifting of the aluminum ends, on the contrary, is not accurately captured by the simulation at the large reduction (Fig. 8.10c). Lifting does appear in the simulation, but not to the same degree as in the experiment. Since the simulated lifting of the steel (Fig. 8.9) is in agreement with the experiment, it may be inaccurate material description of the aluminum causing accuracy problems, cf. discussion in relation to Fig. 8.2, Fig. 8.4b and Fig. 8.5a, where it becomes evident that the aluminum has induced anisotropy and is softer in the center of the original rod than in the outer regions.

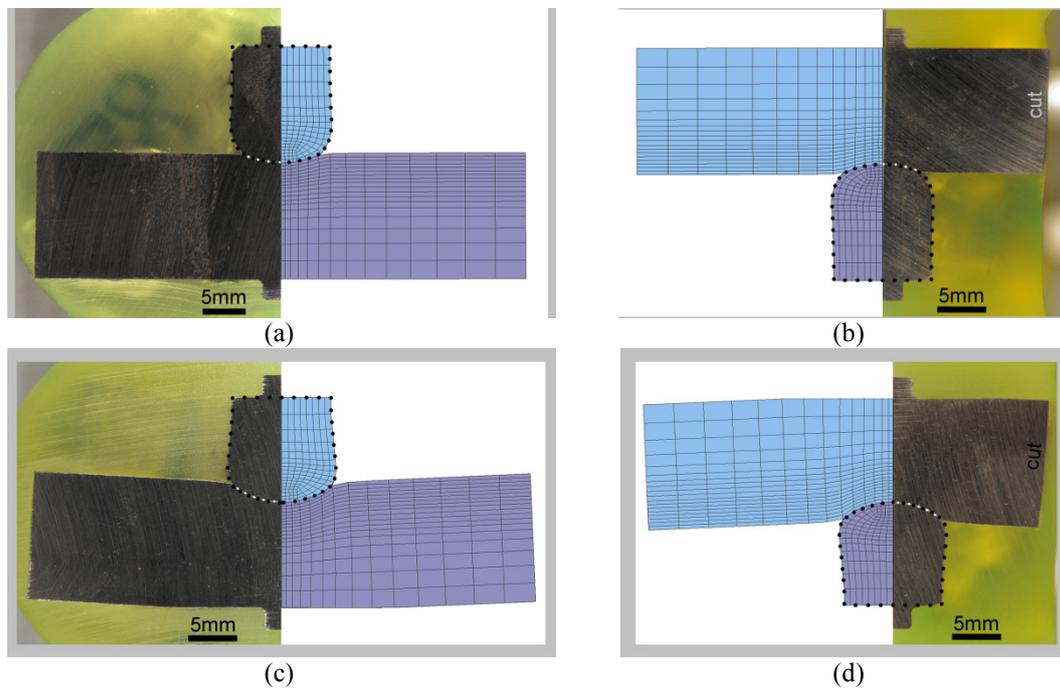


Fig. 8.9. Comparisons between experiments and simulations with cross-cylinder geometry in steel 235 shown in cross-section 1-2 with (a-b) 7.9% reduction and (c-d) 18% reduction. Black dots follow the simulations and are mirrored to the experiments. White dots follow experimental contact interfaces.

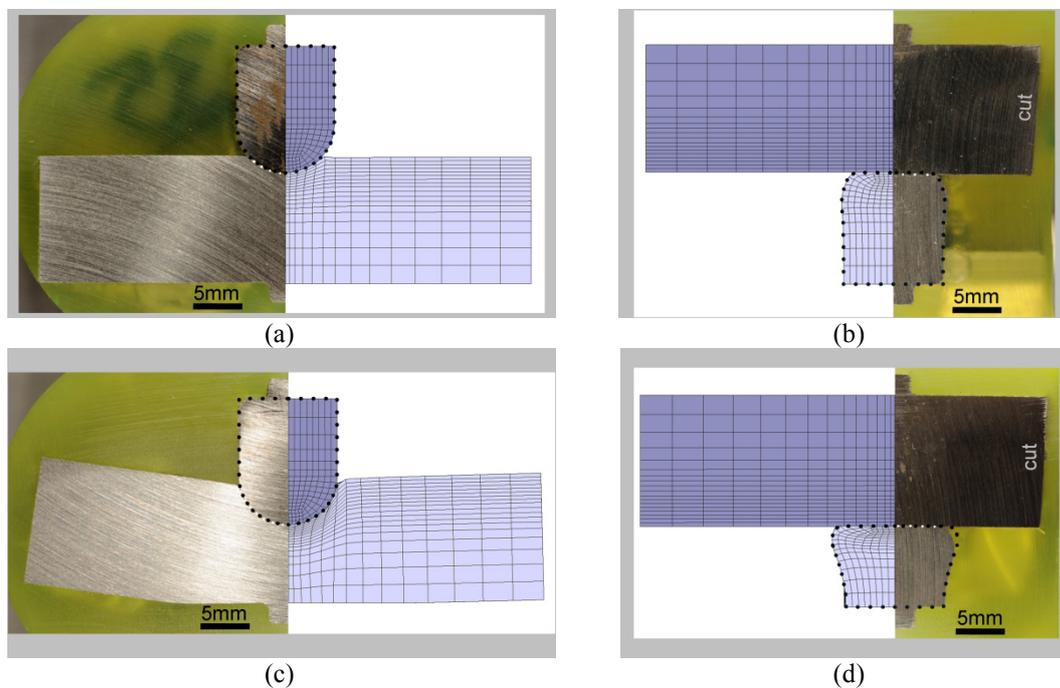


Fig. 8.10. Comparisons between experiments and simulations with cross-cylinder geometry in aluminum (lowest) and steel 235 (topmost) shown in cross-section 1-2 with (a-b) 6.2% reduction and (c-d) 19% reduction. Black dots follow the simulations and are mirrored to the experiments. White dots follow experimental contact interfaces.

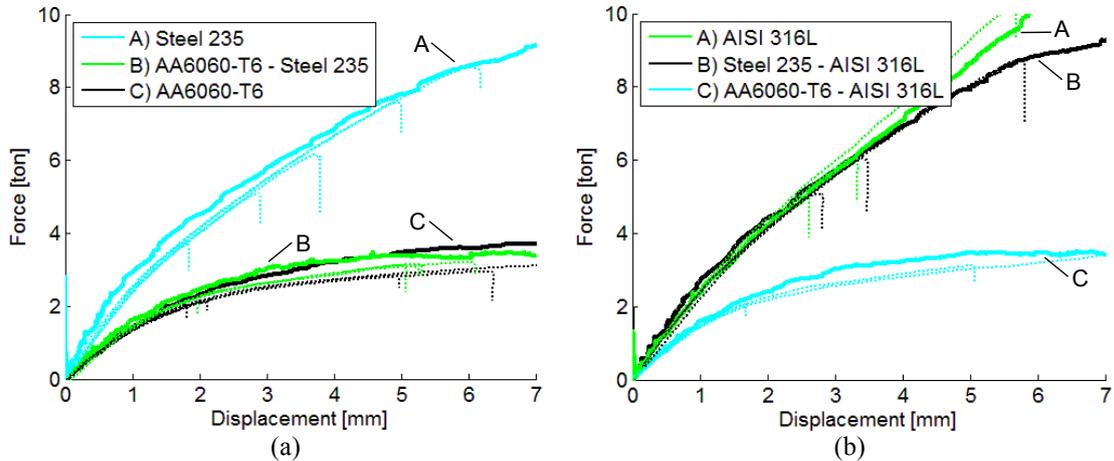


Fig. 8.11. Force-displacement curves for cross-cylinder geometry of (a) the three material combinations examined by cross-sections and (b) the three material combinations not examined by cross-sections. Solid curves correspond to simulations, while dotted curves correspond to experiments.

Fig. 8.11a shows the corresponding force-displacement curves for the cross-cylinder tests. The simulated curves are again increasing step-wise, but compared to Fig. 8.6 they increase more smoothly. This is a result of more gradual increase of the contact area, because fewer nodes get in contact at the same time when compressing these geometries. Satisfactory agreement between numerical and experimental curves is obtained also for these tests.

Force-displacement curves of the three remaining material combinations are shown in Fig. 8.11b, which shows agreement between experimental and simulated behavior to a degree similar to the curves shown in Fig. 8.11a.

Verification of mechanical contact model

The comparisons of experimental and simulated cross-sections as well as presented force-displacement curves have shown that the mechanical contact development can be simulated by the implemented numerical algorithms. The maximum volume loss is for all cases less than 0.15%, showing that the irreducible flow formulation works well with the contact penalty model.

8.2 Forming of seamless reservoirs

Alves et al. [1-2] have utilized the contact implementation developed in the present work to simulate the contact between tubes and deformable mandrels in the development of a metal forming technology for manufacturing small-size, seamless cylindrical reservoirs or metallic liners for composite overwrapped pressure vessels (COPVs). The forming process is schematically shown in Fig. 8.12 and consists of: (i) upper and lower semi-ellipsoidal shaped dies, (ii) a container, (iii) a recyclable, deformable mandrel and (iv) a tubular preform [1-2].

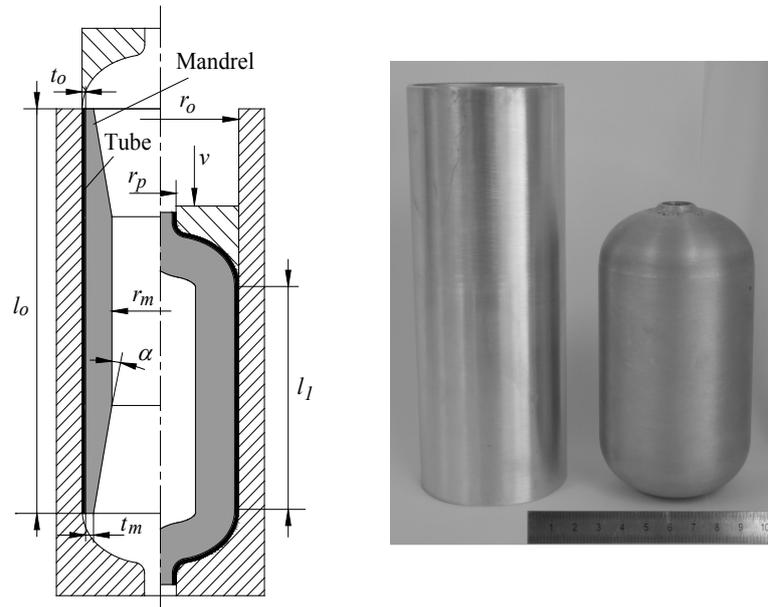


Fig. 8.12. Shaping a tubular preform into a small-size cylindrical reservoir with semi-ellipsoidal ends by cold forming. The enclosed photograph shows the preform and the final reservoir made from Aluminum AA6063-T0 with 60mm diameter.

The forming operation is accomplished by axially pressing the open ends of a tubular preform between the semi-ellipsoidal shaped dies until achieving the desired geometry. The container constrains material from outward flow in order to avoid the occurrence of buckling and helps minimizing the errors due to misalignment between the tubular preforms and the individual dies. The mandrel provides internal support to the tubular preform during plastic deformation in order to avoid collapse by wrinkling and local instability at the equatorial region. The mandrel is made from a low melting point alloy that is capable of continuously adapting its shape to that of the formed tube and is easily removed by melting (recyclable), while leaving the reservoir intact, at the end of the process.

Fig. 8.12 includes an example of a formed reservoir fabricated from commercial tubes of aluminum AA6063-T0 with the utilization of internal mandrels made from a commercial low melting point alloy MCP137 ($T_{melt} = 137^{\circ}C$) comprising bismuth, lead, tin and cadmium. This forming technology is a consequence of four basic mechanisms that compete with each other; plastic work, friction, local buckling and wrinkling. Plastic work is caused by compression along the circumferential direction which gradually deforms the tube against the dies. Friction develops gradually as the tube deforms against the semi-ellipsoidal shaped dies. Local buckling and wrinkling are associated with compressive instability in the axial and circumferential directions and limit the overall formability of the process by giving rise to non-admissible modes of deformation.

In the finite element implementation I-Form, including the contact algorithms developed in this work for dealing with the contact between the tube and the deformable

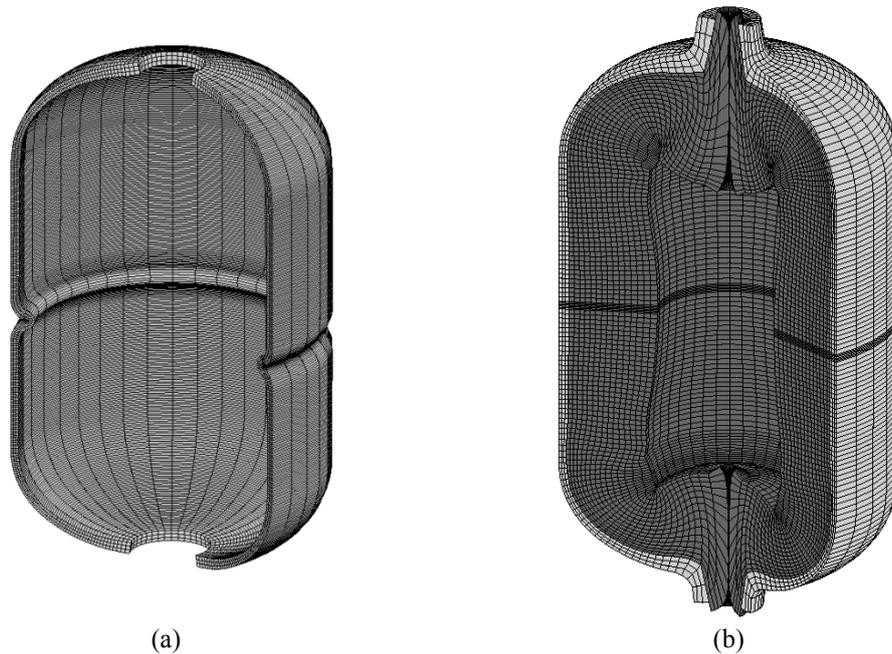


Fig. 8.13. Finite element predicted geometry at the end of the process (a) without and (b) with internal deformable mandrel.

mandrel, the forming process is simulated as shown in Fig. 8.13 in terms of the final geometry [1-2]. Fig. 8.13a shows the final geometry if the mandrel is left out, while Fig. 8.13b shows the final geometry obtained by utilization of a mandrel. The simulations show the important role taken by the mandrel in order to avoid instability to cause inwards flow. Subsequent removal of the internal mandrel by melting results in the reservoir shown in Fig. 8.12. This example puts into evidence the critical role played by contact algorithms in ensuring adequate estimates of plastic flow and validates the implementation in relation to forming of real components. The mechanical contact algorithms are responsible for the contact between the reservoir and the mandrel as well as the later self-contact of the mandrel near the ends of the reservoir.

8.3 Mechanical joining of tubes

Alves and Martins [3] have developed a tube branching method by mechanical joining and presented simulations of the process including contact between deformable objects based on the contact implementation presented in Section 4.2.

The mechanical joining technology makes use of out-of-plane local buckling by means of asymmetric compression beads as shown schematically in Fig. 8.14a and by an example in Fig. 8.14b. Asymmetric compression beading works at room temperature and is accomplished by axially compressing the tube while leaving a gap opening

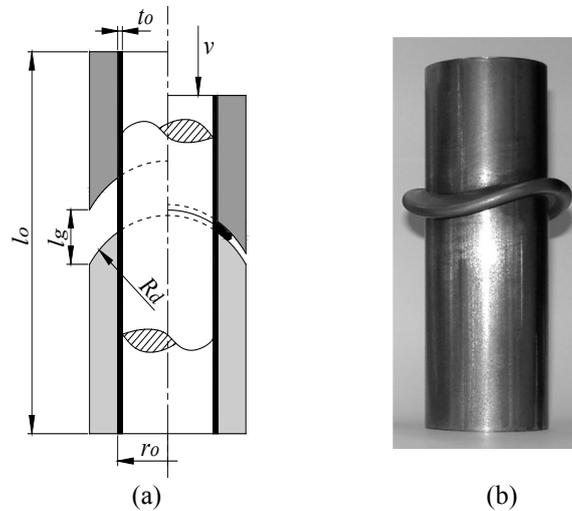


Fig. 8.14. Asymmetric compression beading of thin-walled tubes. (a) Schematic representation of the process and (b) commercial S460MC carbon steel tube showing an asymmetric bead.

in-between the dies that support and hold the tubes. The tube collapses at the gap opening creating the required asymmetric bead.

On the contrary to axisymmetric beads, which are naturally formed by local buckling (during successive in-plane instability waves) in tubes subjected to axial loading between parallel flat dies, asymmetric beads require the development of out-of-plane instability waves between contoured dies; see Gouveia et al. [4]. A tool set-up for producing out-of-plane instability waves in tubes consists of two (upper and lower) contoured dies and an inner mandrel (Fig. 8.14a).

The asymmetric compression bead shown in Fig. 8.14b was performed in a commercial S460MC carbon steel tube that was formed in the as-received condition. The stress-strain curve of the S460MC tubes was determined by means of tensile and stack compression tests performed at room temperature on a universal testing machine with a cross-head speed equal to $100\text{mm}/\text{min}$ (refer to Section 7.1),

$$\sigma = 616\varepsilon^{0.06}[\text{MPa}] \quad (8.1)$$

The numerical simulation of asymmetric compression beading involved approximately 7500 hexahedral elements. The finite element predicted evolution of the out-of-plane instability wave for a test case performed with a mandrel inside the tube (Fig. 8.15) illustrates the key role played by the upper and lower contoured dies in establishing the final shape of the instability wave and the limits of its propagation path. The utilization of a mandrel inside the tube not only avoids defects along the surface of the tube but also guarantees the dimension of the inner diameter (which in many applications is a critical dimension) to stay within tolerances. Fig. 8.15a shows the initial configuration, Fig. 8.15b shows the forming of the bead by instability and Fig. 8.15c shows the final deformation of the bead when the tube material from the two sides of

the bead comes into contact. This is the moment where the contact model takes effect when creating a single compression bead.

Fig. 8.16 shows the finite element predicted and experimental evolution of the load-displacement curve for the component shown in Fig. 8.14b. The two curves compare

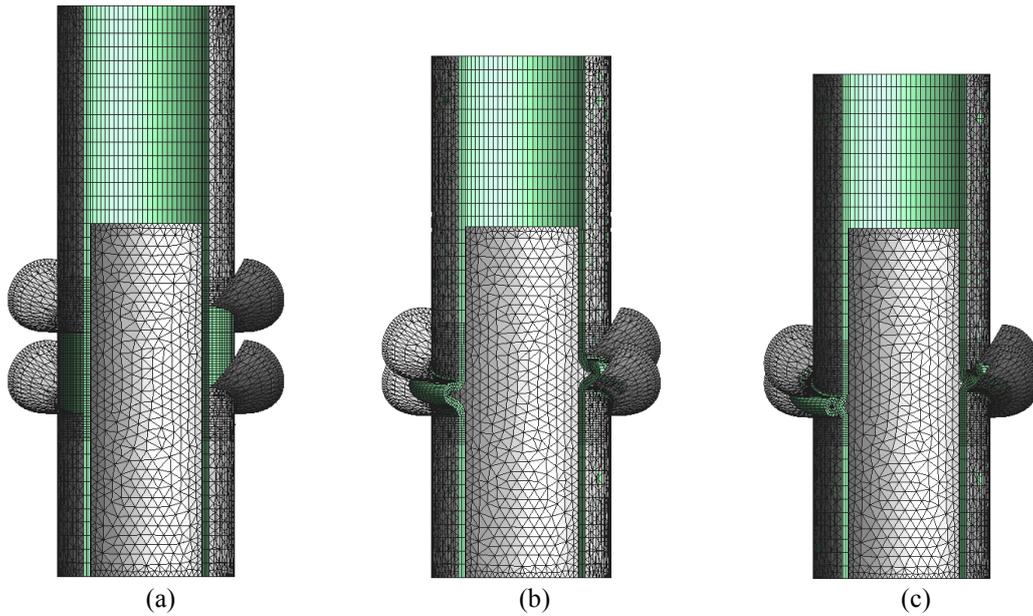


Fig. 8.15. Finite element predicted evolution of the out-of-plane instability wave for a test case performed with a mandrel. (a) Initial configuration. (b) Forming of asymmetric compression bead by instability. (c) Final forming with the two sides of the compression bead being in contact.

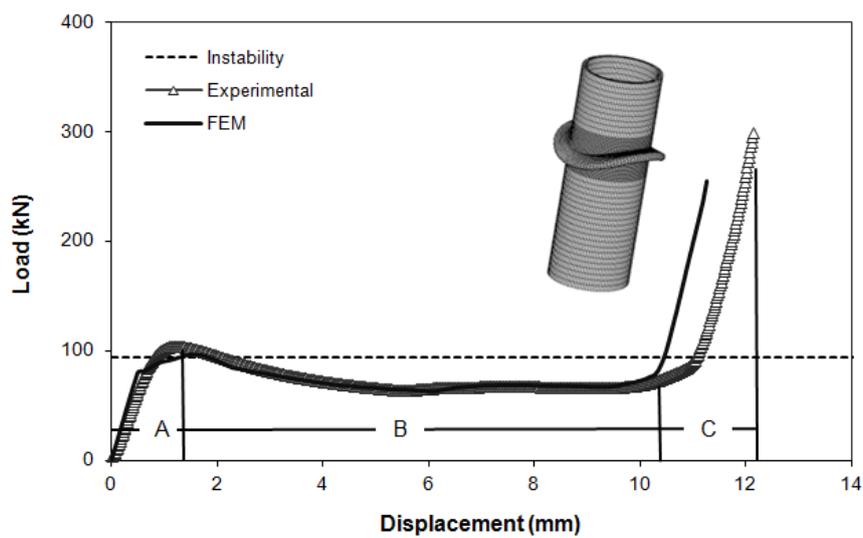


Fig. 8.16. Experimental and finite element predicted evolution of the load-displacement curve.

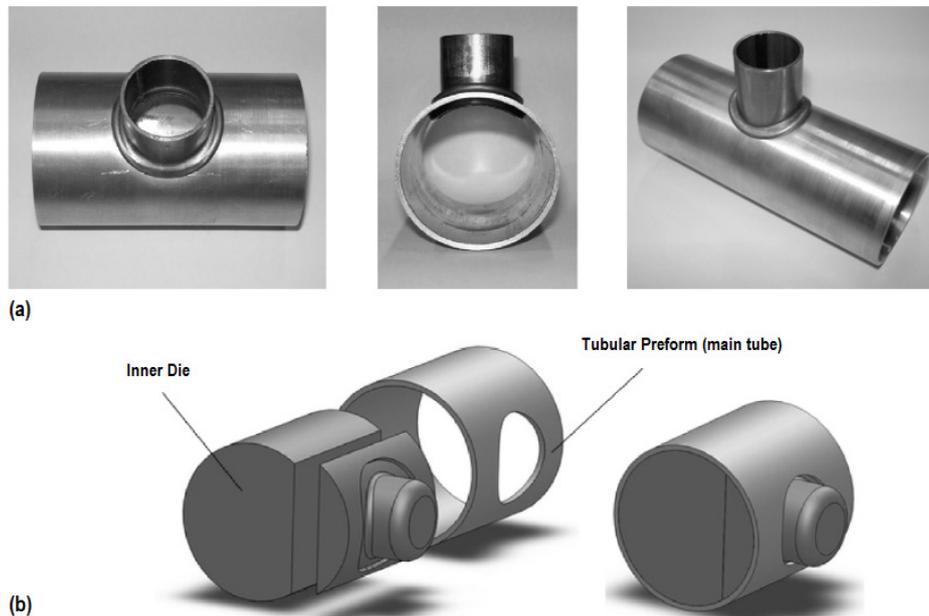


Fig. 8.17. Tube branching by means of asymmetric compression beading. (a) Typical tee fitting produced by the new joining process and (b) schematic representation of the inner sectioned die and pre-form of the main tube.

well in the three stages defined by “A” triggering the out-of-plane instability wave, “B” shaping the asymmetric compression bead from the out-of-plane instability wave and “C” contacting of opposite sides of the asymmetric compression bead.

In the first stage, the load increases steeply as the tube starts being axially compressed. A peak load of approximately $100kN$ is obtained after which the load drops and the out-of-plane instability wave progressively begins to create the asymmetric compression bead in the free gap opening between the contoured dies. The drop in load that is registered throughout the second stage is justified by the fact that forms of equilibrium resulting from local buckling necessitate small values of the axial compressive load as the degree of instability increases. The final sudden increase in the forming load during the third stage is triggered when the opposite sides of the compression bead get in contact as simulated in Fig. 8.15c.

This example shows another utilization of the implemented contact algorithm. The above presented asymmetric compression bead finds application in joining of tubes by the utilization of two opposite asymmetric compression beads to lock tubes by plastic deformation at room temperature. Fig. 8.17 shows an example of two joined tubes.

References

- [1] Alves LM, Martins PAF, Pardal TC, Almeida PJ, Valverde NM (2009) Plastic deformation technological process for production of thin-wall revolution shells from tubular billets. Patent request no. PCT/PT2009/000007, European Patent Office.

- [2] Alves LM, Silva MB, Martins PAF (2011) Fabrication of small size seamless reservoirs by tube forming. *International Journal of Pressure Vessels and Piping* 88:239-247.
- [3] Alves LM, Martins PAF (2012) Tube branching by asymmetric compression beading. *Journal of Materials Processing Technology* 212:1200-1208.
- [4] Gouveia BPP, Alves ML, Rosa PAR, Martins PAF (2006) Compression beading and nosing of thin-walled tubes using a die: experimental and theoretical investigation. *International Journal of Mechanics and Materials in Design* 3:7-16.

9. Challenges in Modeling Projection Welding

Three-dimensional modeling of the resistance welding process is relevant to both spot welding and projection welding. Even so, it is projection welding that will require three-dimensional analysis in most cases, and it is therefore appropriate to put focus on some of the more challenging projection welds from a modeling point of view in this first chapter with applications. Cross-wire welding is widely applied in the electronic industry, for various applications including wire meshes and in the preparation of reinforcements for concrete. At the same time, cross-wire welding presents some of the most difficult aspects of simulation and is therefore well suited for a discussion of challenges and limitations of the numerical implementation. Projection welding of square nuts to sheets is a common application in the automotive industry, and it also presents challenges from a modeling point of view. The two processes differ generally speaking by the type of projection. The cross-wire geometry presents a natural projection, while the square nut has specially designed projections to accomplish the welding.

9.1 Projection welding of cross-wire by natural projection

Cross-wire welding is among the most common non-automotive applications of resistance welding, see e.g. Scotchmer [1]. The primary use of cross-wire welding is in the electronic industry and in the fabrication of wire meshes. In electronics, wires are welded to each other in light bulbs as presented by Goodman [2], who also discusses the loads that the joints will encounter during the life of the light bulb together with eventual failure modes. Knowlson [3] gives other examples of cross-wire welding in electronics; namely welding of a variety of components (resistors, capacitors, diodes and transistors) into small closely packed devices. Wire meshes are used in various products in kitchen wear, shopping carts and for reinforcing concrete structures. Resistance welding poses an alternative to woven structures of wire meshes that are also commonly used. In large scale production of wire meshes, entire rows of cross-wire welds are performed simultaneous as e.g. presented by Jordan [4].

The above mentioned applications of cross-wire welding involve various wire diameters and materials. Wire diameters of $10mm$ are chosen for the present experiments matching the weld setting window with the available welding machines and allowing

easy visual inspection of weld appearance. This diameter lies within the range of diameters used for wire meshes used in reinforcing concrete. The materials utilized in Section 8.1, aluminum AA6060-T6, structural steel S235JR+AR and stainless steel AISI 316L, are used in the experiments.

9.1.1 Experimental analysis

Cross-wire welding of $\varnothing 10\text{mm} \times 100\text{mm}$ wires is performed in the setup shown in Fig. 9.1 between $\varnothing 30\text{mm}$ flat C0-type electrodes. The wires are aligned in the guiding system (Fig. 9.1b) such that they are perpendicular to each other with the intersection point centered relative to the electrodes. The wires are free to move vertically, such that the guides do not provide any stiffness disturbing the weld. The lower wire rests on the lower electrode, while the upper wire is balancing on top of the lower wire by means of e.g. a piece of paper squeezed in between the wire and the guiding system (Fig. 9.1c), which is considered not to add any significant vertical constraint during the welding. The horizontal alignment is naturally solved as soon as the upper electrode presses towards the wires.

During welding (as photographed in Fig. 9.1c), the electrode force and the welding current are measured by a load transducer and a Rogowski coil positioned as shown in Fig. 9.1a.

Choice of welding machine

The choice of welding machine has been taken among two available machines. The major differences between the two machines are the source of applied weld force and type of welding current. One is a 8105 Tecna 250kVA with TE 180 control unit (here-

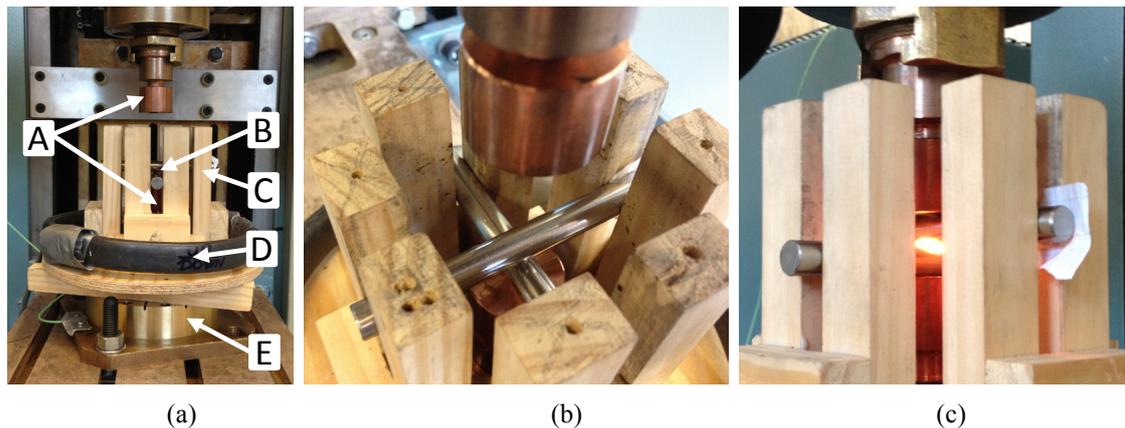


Fig. 9.1. Resistance projection welding of two crossed wires. (a) Setup between two flat C0-type electrodes A. The two crossed wires B are aligned perpendicular to each other and centered between the electrodes by the guidance C holding a Rogowski coil D for measuring the welding current. Below the lower electrode is placed a force transducer E for measuring the welding force. (b) Close-up showing the perpendicular aligned wires resting on the lower electrode in the guidance. (c) Photograph of the cross-wire welding.

after called Tecna). The other is an Expert 170kVA with Harms & Wende HWI 2000 control unit (hereafter called Expert).

The Tecna has a pneumatic force system and supplying AC welding current, while the Expert has a hydraulic force system with additional disc springs to stabilize the force and supply continuous force during eventual collapse of material. The Expert supplies a middle frequency, $1kHz$, welding current, practically acting as a DC current.

Preliminary welding tests were made on both of the machines showing significant differences in the behavior for projection welding. The most severe differences were observed when welding aluminum wires compared to the two steels. Aluminum presents the major problems in resistance welding, both spot and projection welding, due to the narrow window of process parameters resulting in acceptable welds. Because aluminum has relatively high thermal and electrical conductance, only modest heat generation takes place in the aluminum and the heat is quickly transferred through the aluminum. Thus, welding of aluminum generally requires short welding time to avoid the majority of the heat to diffuse away from the welding zone and consequently requires high welding currents to produce enough heat. All in all this results in a narrow process window because the minimum current required to produce a weld is already close to the splash limit.

Cross-wire welding of aluminum was impossible by the Tecna with sufficient set-down and strength. Setdown is defined as in [5],

$$S = \frac{A - B}{A} \quad (9.1)$$

where A is the initial wire diameter and B is the final height of the joint subtracted one diameter A ; see Fig. 9.2. The setdown is a direct measure of the compression and for a given force level an indirect measure of the heat input. It also relates to the weld strength to some degree.

The typical problem observed when welding the aluminum wires by the Tecna is the sudden jump from low setdown to heavy expulsion due to a slight increase in welding current. The visually obvious difference on the two welds is exemplified in Fig. 9.3. The sudden jump in the welding behavior observed with the Tecna is shown in Fig. 9.3a-b, while an example of a weld obtained by the Expert is given in Fig. 9.3c. The aluminum wires used in Fig. 9.3c are glass blasted for increased consistency in the ex-

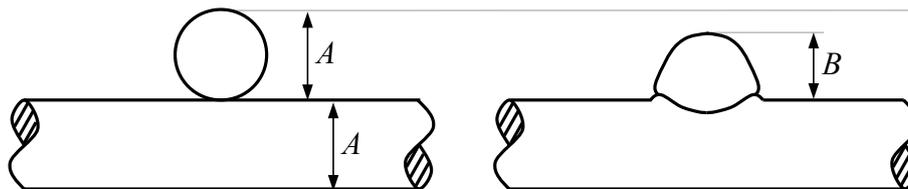


Fig. 9.2. Parameters A and B for definition of setdown.

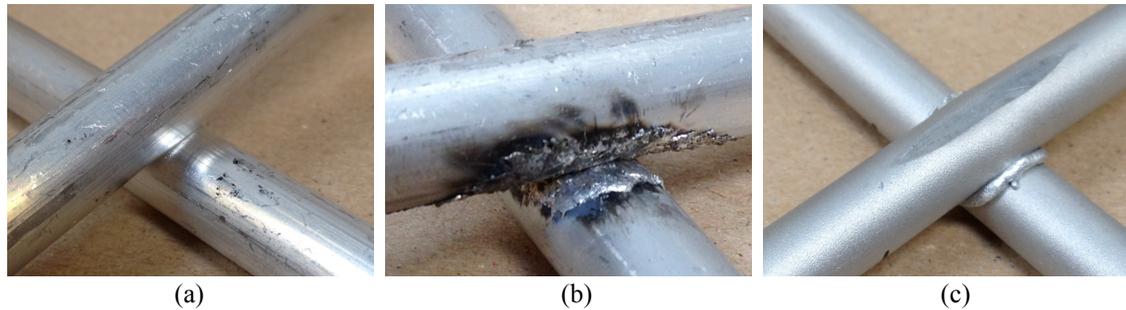


Fig. 9.3. Examples of cross-wire welded aluminum. (a) Low weld setting in Tecna. (b) Slight increase of current in Tecna with resulting heavy expulsion. (c) High weld setting and glass blasted surfaces in Expert.

periments conducted in the Expert.

The major reason for this sudden change in welding behavior in the Tecna is attributed to the pneumatic force system, which does not supply a constant force and which cannot follow up the force and movement when the material softens and eventually melts. The contact area is therefore not properly developed and the high current produces very high temperatures locally resulting in heavy splash on the free surfaces. Another issue relevant to consider in aluminum welding is the type of supplied welding current. When using AC current, the time in-between the peaks with sufficient current will act as cooling time because the low current is not sufficient to heat or even to maintain the temperature in the aluminum near the weld region because of the high thermal conductivity of aluminum. The periods of time without current due to conduction angles less than 100% act in the same way. This increases the need for increased current which leads to higher peaks with large risk of splash.

Batten [6] also reported narrow process window for cross-wire welding of aluminum and discussed similar problems when welding on a pneumatic force controlled machine using AC current. Batten [6] improved the follow-up of the electrode movement when the cross-wire collapses by a sudden release of the pressure on one side of the piston. This allowed reasonable welding of aluminum cross-wires, but cannot be considered a proper control of the welding force or electrode movement. Batten [6] also reported splash in some cases and proved by use of a photo-cell that it always happened near the current peaks in agreement with the above discussion following the experience gained by the Tecna.

Welding by the Expert on the other hand gives the possibility of joining aluminum cross-wires with stable development of setdown as function of welding current. The major benefit gained with the Expert is the hydraulic force control with inserted disc springs between the piston and the weld electrode. The disc springs are designed to follow up the movement of the electrode while keeping the force constant. When the material collapses in projection welding, the load on the springs decreases, which results in expansion that will push the electrode down to follow the specimens being welded. In addition, the middle frequency, DC-like, welding current avoids the above concerns of the current peaks.

The cross-wire welding of aluminum requires the use of the Expert welding machine. The stainless steel can be welded easily on both of the machines. On the contrary to the expectation, it appears that the structural steel is easier welded by the Tecna. The as-received structural steel was welded without problems by the Tecna, while resulting in splash already at low current settings in the Expert. This is not well understood, but it is proven to be related to the contact interface. After grinding the black surface away to arrive at a blank metal surface, the structural steel was welded trouble-free by the Expert. An unconfirmed hypothesis is that when using the AC current in the Tecna, the surface layer is softened and squeezed out during the initial stage of the first quarter cycle of the current. This will avoid sending the peak current from the AC through this layer before having metal to metal contact, and therefore splash is avoided. In the Expert, the current starts almost instantaneously (compared to the AC-form in the Tecna), and hence this will result in the high current sent through this surface layer with high electrical resistivity leading to high and quick heat generation and splash. This, as mentioned, is only a hypothesis and it can be questioned if the contact resistivity between the two wires can be changed so quickly as within the first quarter of a cycle, corresponding to within $10ms$. The support to the hypothesis as regards this question is found in the understanding of the dynamic contact resistance in resistance spot welding, where it plays an even more significant role. Thornton et al. [7] (including some of their references) report that the major change in the electrical contact resistivity happens within the first quarter cycle, which supports the above hypothesis.

To conclude the discussion of the differences between the two welding machines, it is noted that all following welds were made by the Expert welding machine. Due to the above discussion, the structural steel wires are ground before welding. The aluminum wires were glass blasted at least three hours before welding in order to obtain consistent results. All aluminum wires were welded in the same day as they were glass blasted. The glass blasting removes the existing oxide layer and a new even layer can develop before welding. All wires were cleaned before welding.

An example of measured welding current and force from welding on the Expert is shown in Fig. 9.4. The example originates from cross-wire welding of stainless steel wires through $600ms$ weld time with specified weld current $12kA$ under electrode force set by a pressure corresponding to approximately $10kN$. The measured current (Fig. 9.4a) oscillates around a mean value of $11.5kA$. The force is stable at $10.1kN$ before applying the current. When the current is applied, the force is shortly increased to around $10.5kN$ due to thermal expansion, while softening of the material is giving rise to the immediate drop in force level to around $8.4kN$. The force stabilizes hereafter at $9.0kN$ for the remaining process time. The average electrode force during the welding time is $9.0kN$. These process curves are representative for the welds presented in the following, where the process parameters are presented by the current setting (here $12kA$), the weld time (here $600ms$) and the average measured electrode force (here $9.0kN$).

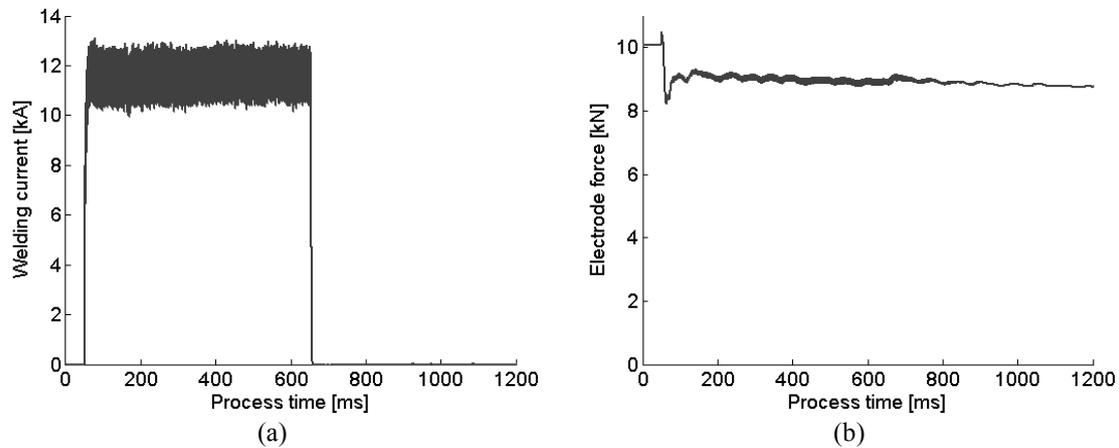


Fig. 9.4. Example of measured welding current (a) and force (b) as function of process time.

Resulting cross-wire welds

Fig. 9.5 presents all the cross-wire welding experiments conducted by the Expert in terms of setdown as defined by (9.1) and corresponding Fig. 9.2. The stainless steel AISI 316L and structural steel S235JR+AR are welded with the same settings of electrode force and weld time and presented in Fig. 9.5a for varying current. The aluminum, AA6060-T6, is welded at lower electrode force and shorter weld time because of the lower strength and the high electrical and thermal conductance properties of aluminum, which at the same time give rise to a higher current level to facilitate welding. Fig. 9.5b presents the setdown as function of current for the aluminum cross-wire welds. The figure also includes the results obtained with the two steels. Despite the different force levels and weld times, the comparison is relevant for showing the differences in applied currents. The comparison also shows that similar setdowns are obtainable in the Expert welding machine.

Between the two steels, the stainless steel experiences a larger setdown under similar welding conditions. This is mainly because of larger heat generation due to larger electrical resistivity and poorer heat conduction in the stainless steel.

The figures include indication of splash by the squares surrounding the relevant points. Before concluding too much about splash versus no splash, it is noted that the splashing behavior and the resulting appearance is different from splash in spot welding. In spot welding the weld nugget is enclosed between the sheets and splash is a result of escape of an amount of molten material between the sheets facilitated by the pressure in the enclosed molten volume great enough to penetrate the closing area of the sheets. It is visible by the related flash when it occurs, and it is visible afterwards as the escaped material typically solidifies on the sheets outside the spot weld. In cross wire welding, splash typically occurs from the material that is already squeezed out (see material piled up around the weld in Fig. 9.3c). As a result of this, the splash occurs from a volume under much less pressure and the appearance is a small drop jumping from the free surface. This may produce a small cone point from the place it left,

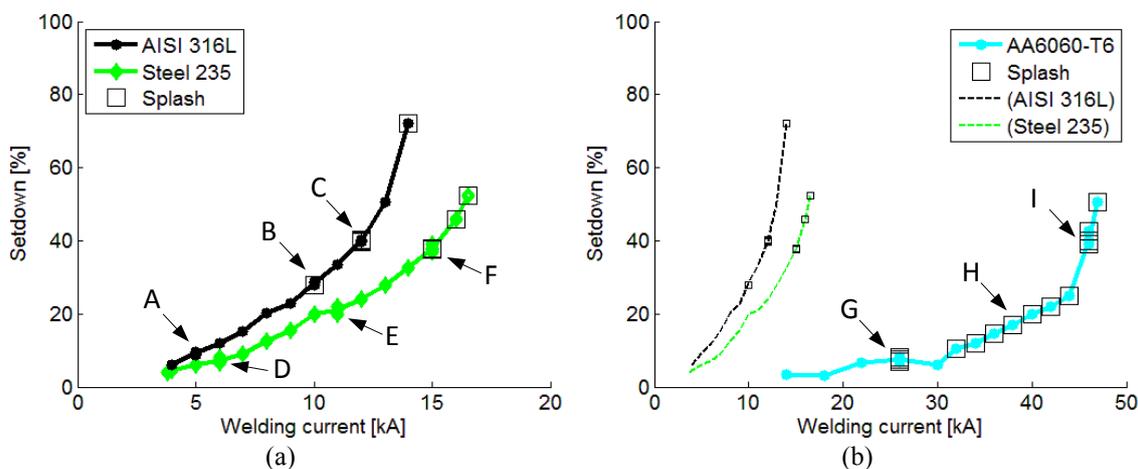


Fig. 9.5. Setdown as function of current for (a) stainless steel AISI 316L (force: 8.9-9.4kN) and structural steel S235JR+AR (force: 8.5-9.8kN) with weld time 600ms, and (c) aluminum AA6060-T6 with force 4.2-5.5kN and welding time 140ms, where also the two steels are included for comparison of current level, though with different electrode force and weld time. Three points on each curve (marked A-I) include five repetitions and are selected for further analysis.

but depending on the remaining process, it may also even out. Furthermore, the drop leaving the weld zone has large possibility of hitting elsewhere than the wires. Summing up, it is more difficult to record all splashes and it is less damaging for the weld when it occurs.

Three points on each curve in Fig. 9.5 are selected for further analysis with capital letters A-I for identification. Each of these points includes five repetitions, which practically coincide. One of the five repetitions are selected for cross-sectional analysis to be presented in the following.

Stainless steel AISI 316L

The three selected stainless steel welds, A-C, are shown in Fig. 9.6a-c. The development of the weld can be extracted from the overall geometry resulting from different current levels. The wires initially deform such that material is starting to flow out from the contact interface (Fig. 9.6a). When the generated heat is more, the material softens with larger amount of material squeezed out between the two wires (Fig. 9.6b-c).

The type of material flow changes with heat input. At low heat input, the magnified cross-section in Fig. 9.6d shows that the deformed material near the interface kept substantial strength and stayed in the solid phase because of the sharp edges remaining. This is seen in the figure despite the many scratches, which originate from particles falling off from the material during grinding. The particles, mainly originating from the left side in the figure, scratch the cross-section during further grinding. This has happened in many of the cross-sections presented in the following, though in less degree than in Fig. 9.6d.

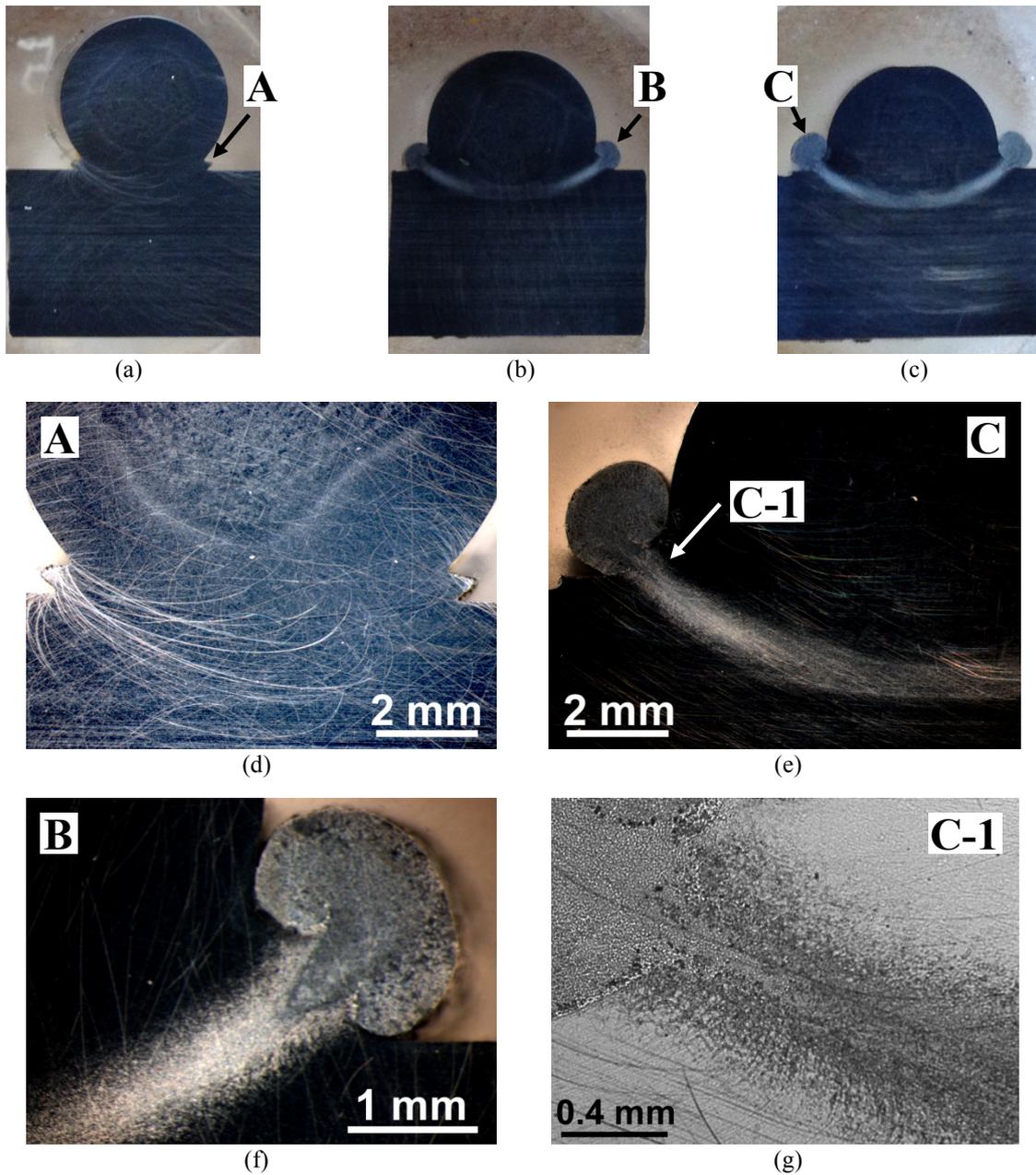


Fig. 9.6. Cross-sections of selected stainless steel AISI 316L cross wire welds. (a) Weld A with 9.5% setdown. (b) Weld B with 29% setdown. (c) Weld C with 40% setdown. (d-f) Magnifications with positions indicated in (a-c). (g) Further magnification of position marked in (e).

When the material is softened more due to larger heat input (Fig. 9.6b-c), the flow changes to produce more rounded geometries. The magnifications in Fig. 9.6e-g show a resulting mushroom shaped squeeze-out of material. The magnifications show that this flow originates from a narrow opening between the wires, which has happened in a mushy or melted state of the material.

The microstructure suggests that there has been no other melting, and hence there is no weld nugget and the welding is not characterized by fusion. It can be compared to friction welding where the joining is facilitated by elevated temperatures (without melting) and high pressure.

The observation and corresponding conclusion is supported by Pan and Watt [8], who also characterize cross-wire welding by having no or little melting, but rather sinking of the two wires into each others' softened zones. They further support the link to friction welding. Fukumoto et al. [9] consider cross-wire welding to be a solid-phase joining process because any liquid phase is expelled together with oxide layers and will appear outside the joining area. Khan et al. [10] examined cross-wire welding of stainless steel and supported the process description by Fukumoto et al [9]. Khan et al. [10] investigated the transition between solid state bonding and fusion welding concluding that fusion welding needs lower electrode force and/or higher weld current in agreement with the nature of heat development.

Structural steel S235JR+AR

The cross-sections, D-F, for the structural steel welds are shown in Fig. 9.7. This steel type shows clear phase and structure changes in the microstructure. The heat affected zone (most clear in Fig. 9.7g) consists of two regions. The central volume has been above the austenitization temperature and the surrounding heat affected zone has been in the transition zone. The zone around the interface seems to have been kept below the melting temperature during all three welds.

The magnifications at low heat input (Fig. 9.7d-e) show a clear identification of the original interface. The material is only joined by local welds across the interface. At larger heat input levels (Fig. 9.7f-g), an interesting difference appears when comparing to the stainless steel welds. The material is squeezed out to produce a sharp edge and in the case with more setdown (Fig. 9.7g) producing two sharp edges because the material from each side curls to keep the circumferential tension low by minimizing the radius. Since the material curls both ways, it has not been sufficiently joined before being squeezed out.

The explanation of the differences between the two steel types in the shape of the squeezed out material cannot be explained by the stainless steel being liquid and the structural steel being solid when forming the expelled shape, as it may look like at first. Returning to the comparison to friction welding, Fig. 9.8 shows a photograph of an earlier friction weld that joins a mild steel and a stainless steel, where none of the materials were melted. This joint shows the same shape of the produced flashes in the two materials as appear in the material squeezed out during cross wire welding.

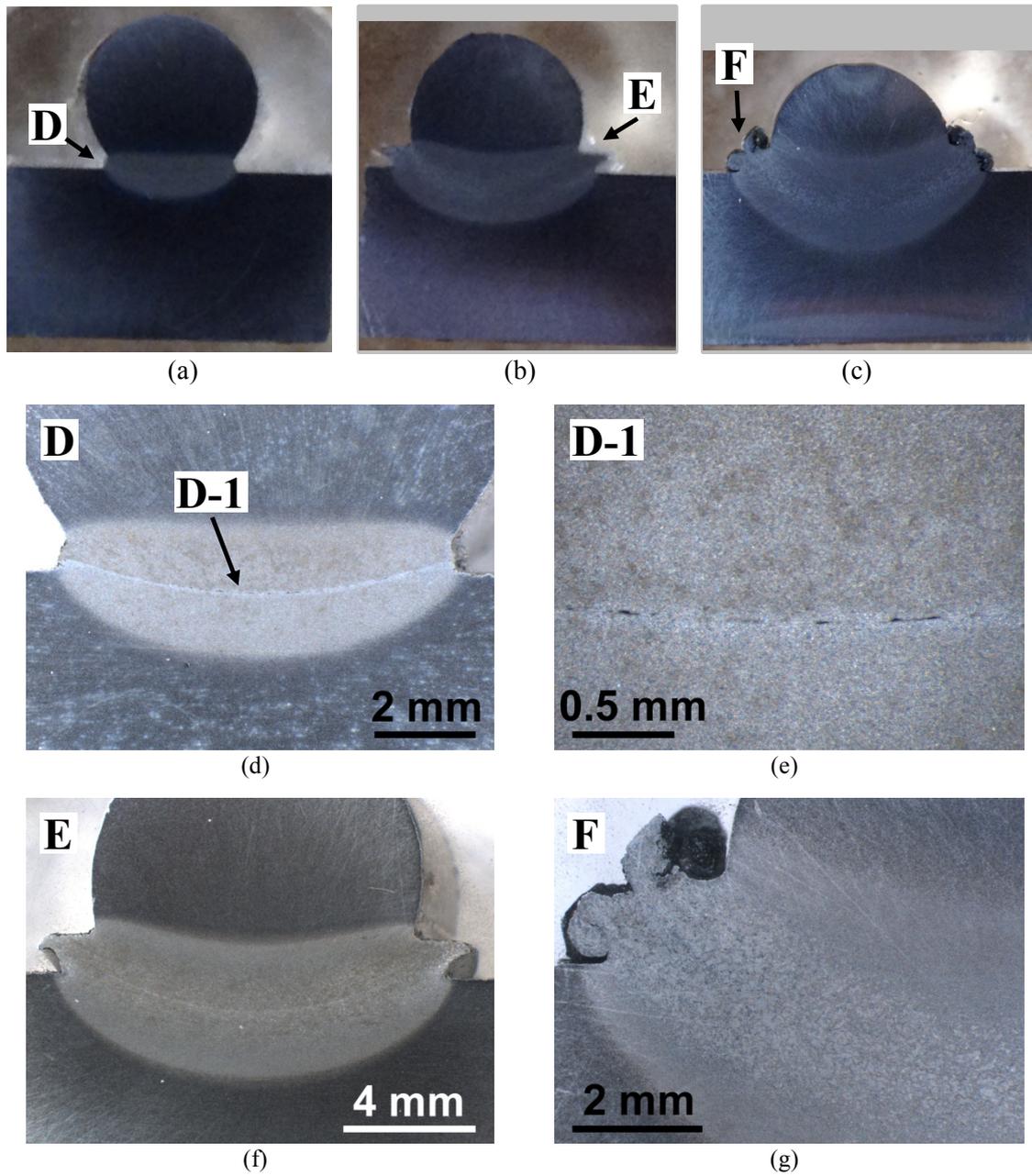


Fig. 9.7. Cross-sections of selected structural steel S235JR+AR cross-wire welds. (a) Weld D with 7.5% setdown. (b) Weld E with 20% setdown. (c) Weld F with 38% setdown. (d,f-g) Magnifications with positions indicated in (a-c). (e) Further magnification of position marked in (d).



Fig. 9.8. Example of friction welded mild steel to stainless steel. The geometries of the formed flashes of the two materials are comparable to the shape of the material squeezed out during cross-wire welding.

Aluminum AA6060-T6

Finally, the aluminum cross-wire welds, G-I, are shown in Fig. 9.9. The tendency of squeezing out material seems less than in the above cases including steel wires. From a heat generation point of view, this is explained by the fast heat conduction that softens the material more evenly than in the cases with stainless steel and structural steel. When the material is overall softer, it experiences more global deformation when enlarging the area supporting the electrode force. It is supported by an explanation from the electrode force itself, which has been selected too high for optimal welding. The larger force results in a larger initial contact area with lower heat concentration as result.

This is supported by the appearance of the interface between the joined wires. At low heat input and setdown (Fig. 9.9a,d), a clear interface is observed between the two wires. The largest setdown (Fig. 9.9c,f-g) associated with high heat input has a more mature bond, while the medium setdown and heat input (Fig. 9.9b,e) lies somewhere in between. Weld H has the central part of the interface intact and only a solid joint on the outer ring. This is to some extent visible from Fig. 9.9b and by the regions *i* (center) and *ii* (outer ring) in Fig. 9.9e. It is shown with further magnifications in Fig. 9.10, where the original interface is clear in the center of the weld while the outer ring is joining the two wires. Another weld from point H (i.e. another of the five repetitions) is shown in Fig. 9.11 after separation of the two wires. This figure shows the central area of the former weld to have a blank surface without sign of fracture. The surface even stays close to the original glass blasted condition. The joint has only been facilitated by the welded outer region, which shows sign of fracture and partial fracture.

Improvement of the aluminum welds would require lower electrode force to allow heat development before increasing the contact area too much.

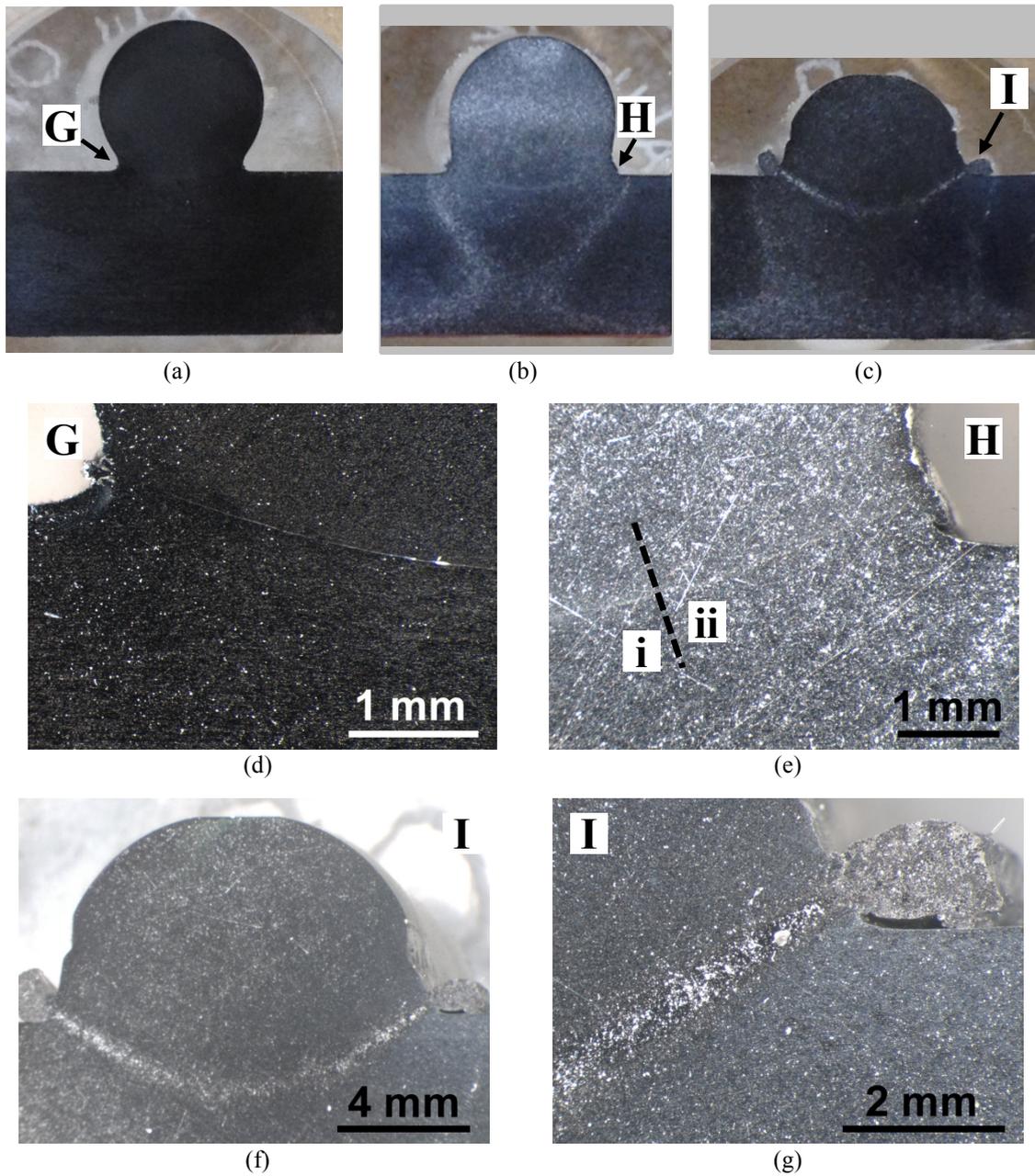


Fig. 9.9. Cross-sections of selected aluminum AA6060-T6 cross-wire welds. (a) Weld G with 7.0% setdown. (b) Weld H with 17% setdown. (c) Weld I with 43% setdown. (d-g) Magnifications with positions marked in (a-c).

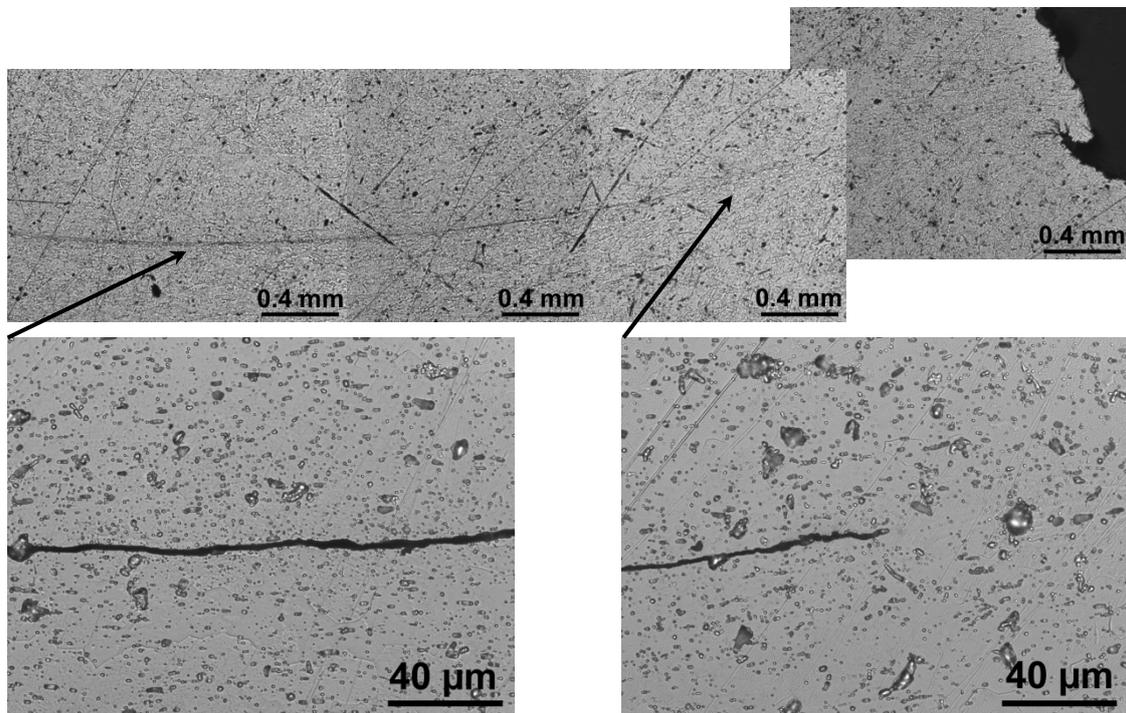


Fig. 9.10. Magnified cross-section of aluminum weld H around the interface in a similar view as Fig. 9.9e. The upper row of four photos are assembled to reproduce half of the interface and the two lower photographs are further magnifications.

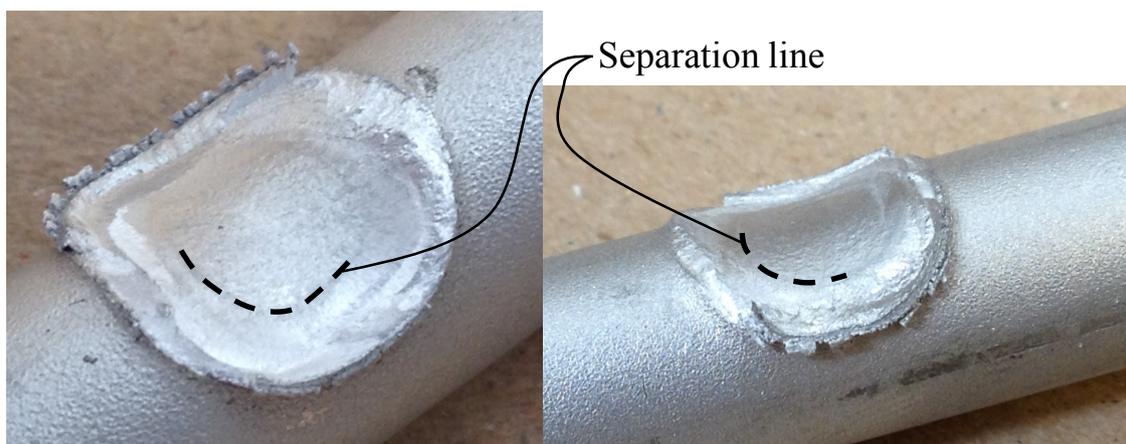


Fig. 9.11. Separated aluminum wires of weld H corresponding to Fig. 9.9b,e and Fig. 9.10 (another of the five repetitions). The partially indicated separation line shows the transition from no bonding in the center to bonding on the outside.

9.1.2 Numerical modeling

Numerical modeling of cross-wire welding is rarely seen in literature. One contribution was given by Scotchmer [1], who presented numerical simulations by SORPAS in two dimensions with good comparison to experimental cross-wire welds. This has required trial-and-error calibration of the third dimension in terms of the thickness of the elements varying through the cross-section. A compromise must have been necessary for obtaining enough mechanical stiffness by thick elements, while having elements thin enough to concentrate and generate heat.

This section presents 3D simulations of cross-wire welding and comparisons to the above experiments when possible. The finite element mesh utilized for the simulations is shown in Fig. 9.12. One quarter of the physical geometry is simulated by utilizing the two vertical symmetry planes that go along the wires. The mesh consists of 7996 8-node elements with mesh density increased around the contact point between the two wires, where the weld takes place.

The following three figures, Fig. 9.13-9.15, show comparisons of simulations and experiments of the cross-wire welds carried out at low weld settings. For the stainless steel, this corresponds to weld A among the experiments; see Fig. 9.5a for identification of the weld and Fig. 9.6a for the experimental cross-section. The weld settings include $5kA$ weld current applied during $600ms$ under an applied electrode force of approximately $10kN$. Fig. 9.13 shows the comparison for this case in terms of the resulting geometry in the cross-section. The simulation shows the process peak tem-

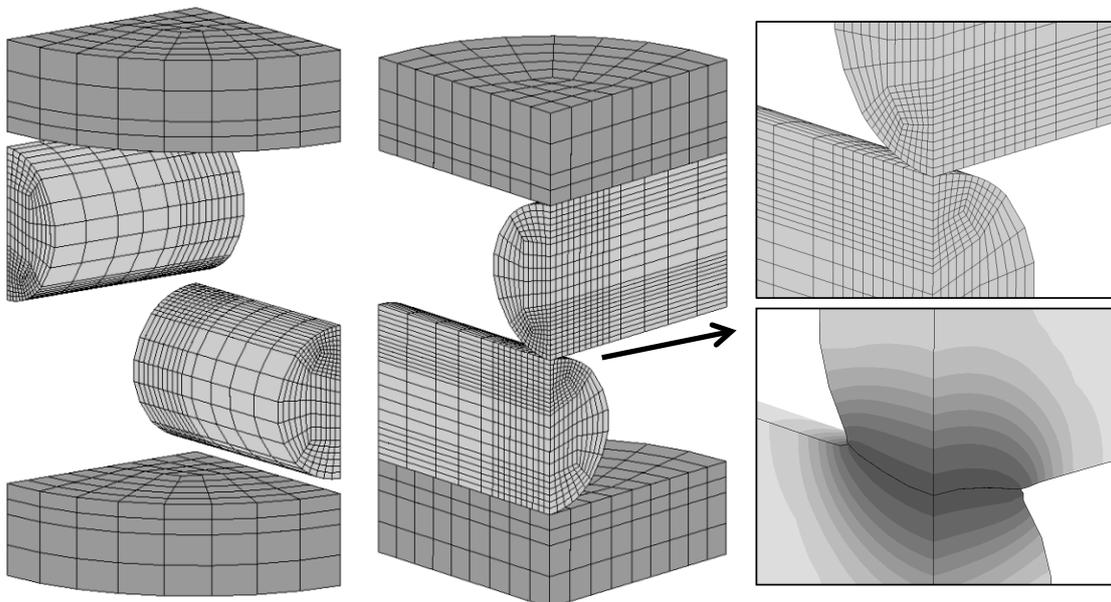


Fig. 9.12. Finite element mesh for simulation of cross-wire welding by application of two symmetry planes. The right figures show the local mesh refinement and an example of the formed contact between the two wires.

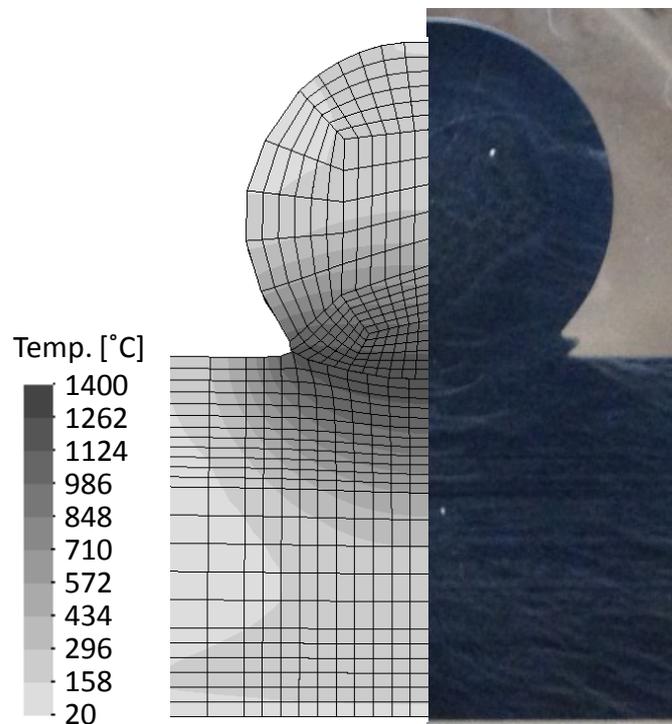


Fig. 9.13. Comparison of simulation and experiment for weld A (Fig. 9.5a and Fig. 9.6a) consisting of two stainless steel rods welded at the low settings. The simulation shows the process peak temperature reached at the end of the welding time.

perature reached until the end of the weld time with overall peak temperature 1224°C , which is below the melting point (1400°C for AISI 316L) as anticipated in the discussion of the experiment. The overall deformation is evaluated by the setdown, which for the experiment is 9.5%. The simulation shows 14% setdown (and 15% setdown after the hold time), corresponding to a final height difference of 0.45mm (and 0.55mm after the hold time) out of 20mm total height before welding.

Besides the larger setdown obtained in the simulation, a noticeable difference is that the experiment shows initiation of flash formed between the two wires and the simulation does not catch this. Part of the explanation can (as always with simulations) be attributed to uncertainties in material properties, but the mesh itself is considered to play a bigger role in the specific example because of being too coarse to reproduce the actual deformation. With too few elements that can only reproduce the geometry changes trilinearly, it is not possible to simulate the local softening and deformation associated with the initiated flash. Another issue that complicates the simulation of the local deformation is that the mesh does not conform to the formed contact area and the flashing, and hence even more elements would be required.

The lacking flash formation in the simulation causes the simulated contact area to remain smaller than in the experiment during the flash formation. This results in more

heat generation and material softening over a larger volume, leading to bulk deformation and finally larger setdown than in the experiment.

This is a limitation of the combination of the current implementation and the utilized mesh. Increasing the overall number of elements does not appear practical on the computers presently used for the simulations. Instead, as a future area of development, the ability of further concentrating the elements can be enhanced and adaptive remeshing procedures can be applied. As regards remeshing, it is important that the remeshing is capable of concentrating elements around the local details in the weld zone. The meshing and remeshing procedures presented in Chapter 5 do not include these possibilities at the moment because the remeshing procedure is, in its nature, aiming at a mesh with uniform distribution of elements. This points out another future area to improve; namely the possibility of using octree based core meshes instead of the grid based core meshes in the remeshing procedures, such that elements can be concentrated in the regions of interest.

A similar comparison is shown in Fig. 9.14 for the structural steel, where better agreement between simulation and experiment is obtained because this weld does not include initiation of flash formation. The actual weld considered is the above weld D, which can be identified in Fig. 9.5a and found by cross-sections in Fig. 9.7a and 9.7d. The weld settings include $6kA$ weld current applied during $600ms$ under an applied electrode force of approximately $10kN$.

Again the overall deformation is evaluated by setdown, which for the experiment is 7.5%. The simulation shows 9.4% setdown (and 9.5% setdown after the hold time), corresponding to a final height difference of $0.19mm$ (and $0.20mm$ after the hold time) out of $20mm$ total height before welding.

Fig. 9.14a shows the overall comparison of simulated geometry to the experiment by its cross-section. It also includes comparison of the temperatures in terms of isothermal contour lines at $700^{\circ}C$ and $900^{\circ}C$ which are mirrored on to the experiment, where the microstructural changes reveal the temperatures obtained during the actual welding. Fig. 9.14b shows the comparison in a magnified view, where the $900^{\circ}C$ isothermal line, which roughly equals the austenitization temperature, practically coincides with the border line between two different microstructures. The volume within the lines is transformed into austenite during welding and is formed into a fine recrystallized grain structure upon cooling. This confirms that both the deformation and heat generation are simulated with high accuracy in this example. It is further justified by a simulated maximum temperature of $1221^{\circ}C$, in agreement with the experiment showing no sign of melting (assumed $1560^{\circ}C$ for S235JR+AR).

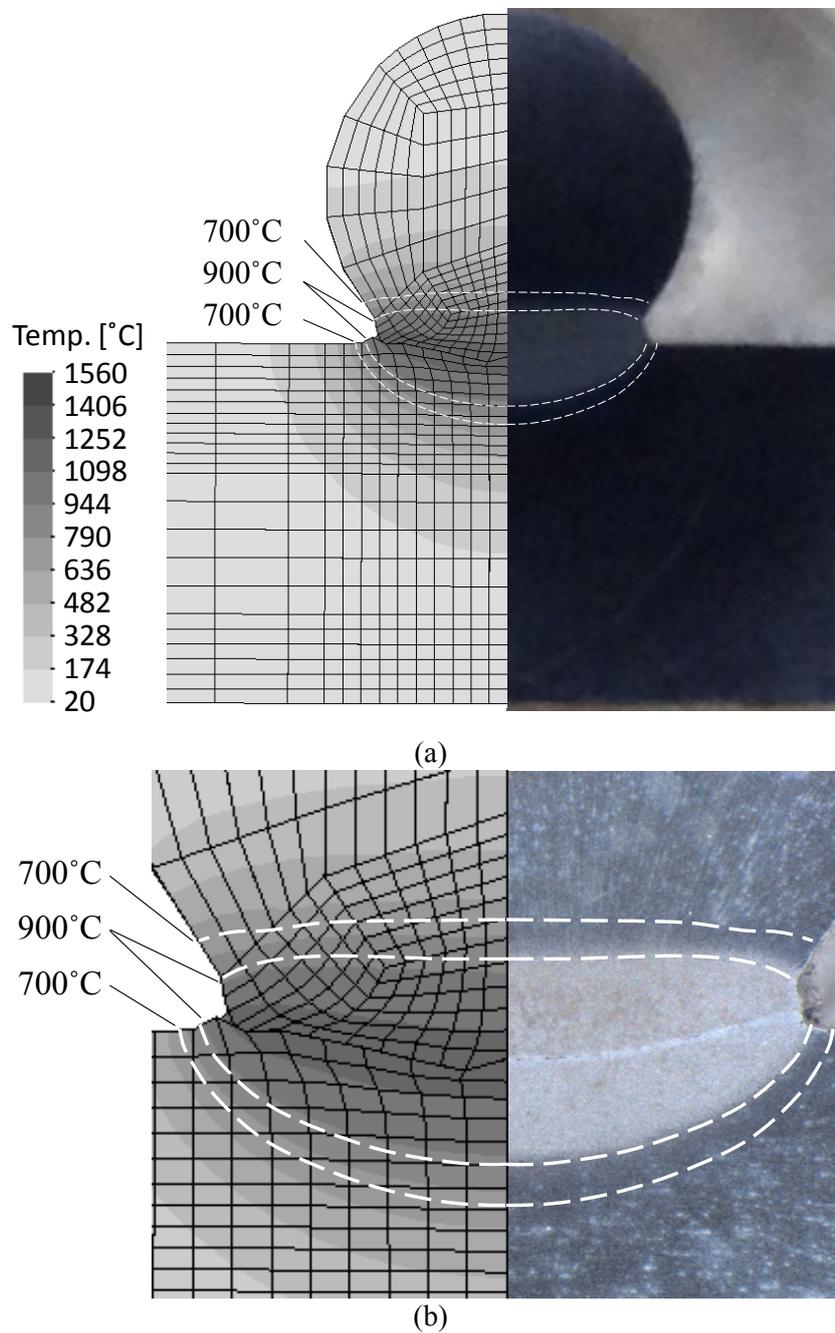


Fig. 9.14. Comparison of simulation and experiment for weld D (Fig. 9.5a, Fig. 9.7a for subfigure (a) and Fig. 9.7d for subfigure (b)) consisting of two structural steel rods welded at the low settings. The simulation shows the process peak temperature reached at the end of the welding time together with corresponding, selected iso-thermal lines mirrored on top of the experiment.

Finally, Fig. 9.15 shows the comparison with aluminum wires as the above weld G, which is identified in Fig. 9.5b and Fig 9.9a. The weld settings include $26kA$ weld current applied during $140ms$ under an applied electrode force of approximately $5.5kN$. The setdown in the experiment is 7.0%, while the simulated setdown is 13% (still 13% setdown after the hold time), corresponding to a final height difference of $0.60mm$ out of $20mm$ total height before welding. The final geometry of the experiment does not show any flash formation, so the explanation for the difference is not the same as for the stainless steel cross-wire in Fig. 9.13. The larger setdown in the simulation of the aluminum cross-wire is rather attributed the indentation in the bulk material, which is visible in the simulation as an amount of material piling up outside the contact area. The reason can either be overestimation of the heat generation, too soft material properties in the simulation or a combination of the two.

It is likely that the mechanical material properties are not accurate since also the pure mechanical contact experiments of Section 8.1 showed inaccuracies when simulating aluminum. The mechanical properties of the squared bars utilized in Section 8.1 are also used for the simulation of the cross-wire welding although the wires were received in $\varnothing 10mm$ rods. The glass blasting may further influence the properties and the formation of contact area.

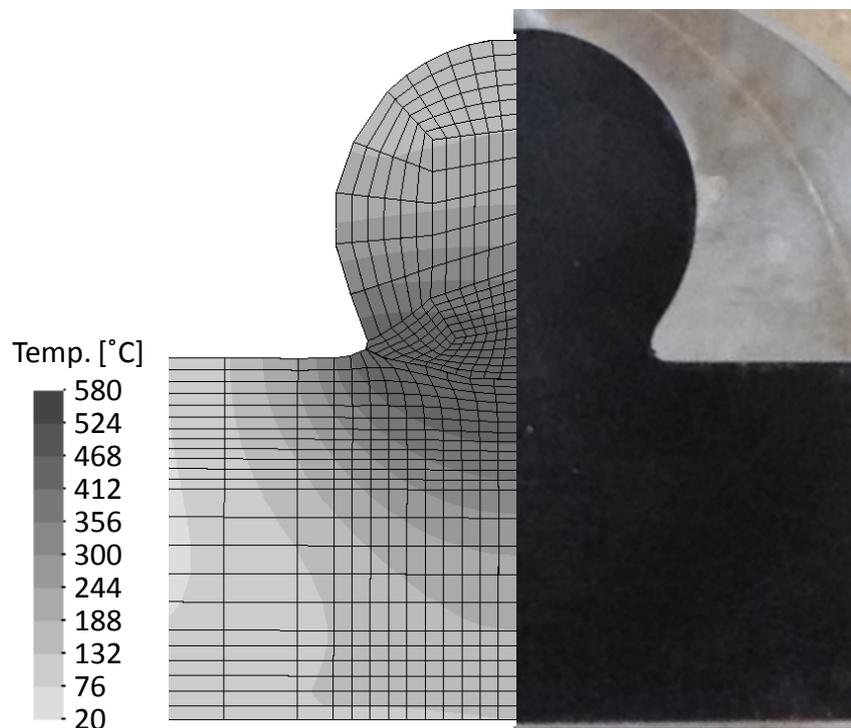


Fig. 9.15. Comparison of simulation and experiment for weld G (Fig. 9.5b and Fig. 9.9a) consisting of two aluminum rods welded at the low settings. The simulation shows the process peak temperature reached at the end of the welding time.

When it comes to cross-wire welding with high weld settings, the simulations cannot predict the actual heat generation and the final setdown. When examining the cross-sections corresponding to medium or high weld settings in Fig. 9.6, 9.7 and 9.9, it appears that material is squeezed out between the two wires because of the high degree of local softening and eventual melting. The local type of deformation is suited for a typical simulation by itself and thus not possible to simulate by the few elements available for the local details. Hence, the discussion given in relation to the initiation of flash formation for the stainless steel welding under low weld settings (see Fig. 9.13) also applies for high weld settings of all the materials. Under high weld settings the consequence is bigger than above and reasonable simulation is not possible. The elements encounter too heavy mesh distortion because they cannot gradually adapt to the geometrical changes, resulting in excess of heat generation and softening to a degree that the applied force cannot be withstand.

A numerical experiment is presented in the following to support the above discussion. It is emphasized that this is not prediction of the welding case, but only for discussion. The numerical experiment is based on cross-wire welding of stainless steel with a welding current of $12kA$ during $600ms$ weld time with an electrode force of $10kN$ (the case corresponds to the above weld C). The mechanical material properties of the stainless steel are limited to remain constant after $1170^{\circ}C$, such that the material

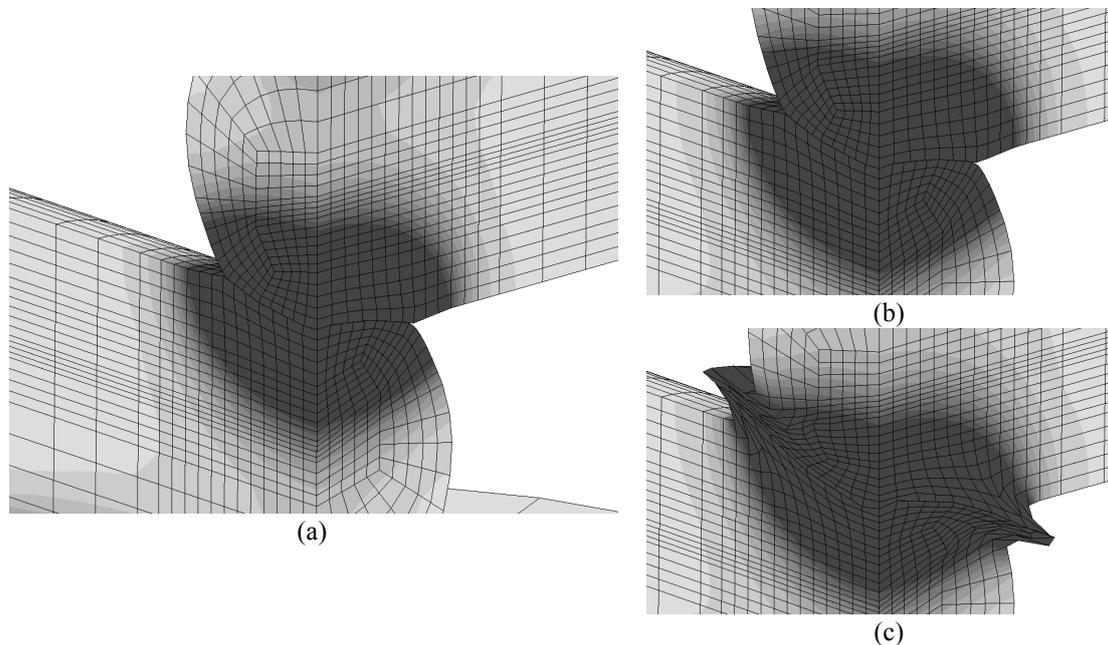


Fig. 9.16. Numerical experiment with material and weld settings corresponding to weld C (Fig. 9.5a and Fig. 9.6c) consisting of two stainless steel rods welded at high weld settings. The darkest color identifies melting. (a) Simulated $200ms$ of the weld time with artificially stiff material at high temperatures. (b) $25ms$ continuation from (a). (c) $25ms$ continuation from (a) with material properties reflecting the actual temperatures, i.e. with softer material.

has enough strength to withstand the electrode force without the need to deform heavily. After one third of the welding time ($200ms$), the geometry and developed temperature field looks like depicted in Fig. 9.16a. It is naturally unrealistic because the molten ball would collapse if not kept artificially stiff, but a gradual collapse would not be possible to simulate due to the above discussion. Further $25ms$ simulation leads to Fig. 9.16b. If instead, the realistic material properties are assumed to apply after Fig. 16a, a constant velocity of $50mm/s$ produce the geometry depicted in Fig. 9.16c after $25ms$.

By doing this, a geometry (Fig. 9.16c) similar to the real geometry after welding with high current settings is obtainable, although after reaching an artificially high temperature. This cannot be used for prediction, but it proves that the deformation pattern including the contact between the two objects is potentially simulated as long as the softened volume is discretized by a sufficient number of elements.

Fig. 9.17 provides additional views of the geometry obtained in Fig. 9.16c for showing the similarity to the real case.

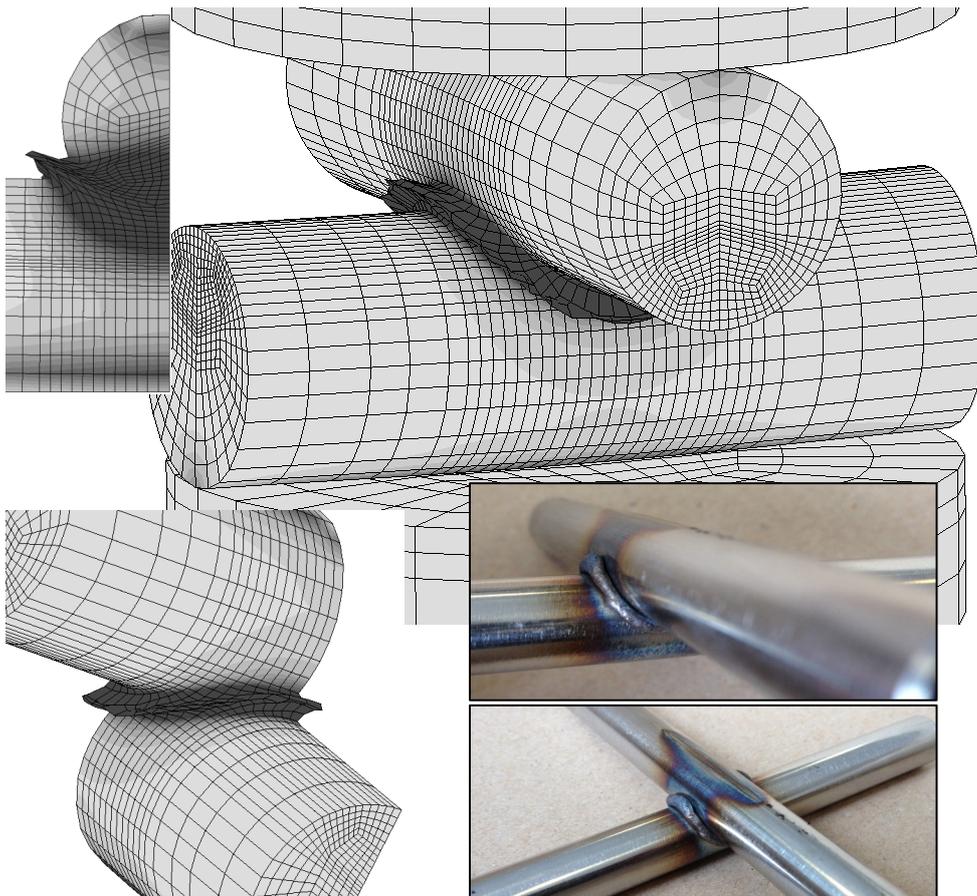


Fig. 9.17. Additional views of the numerical experiment corresponding to Fig. 9.16c. Two photos of the real weld are included.

9.2 Projection welding of square nut to sheet

Projection welding of a square nut to a sheet is considered as a second challenging projection welding case. Nuts, being annular, squared or hexagonal, are welded to sheets in automotive industry for further assembly of other parts. The projections on nuts are specifically made for the purpose of welding; as they would not naturally appear as the projection in the above presented cross-wire welding. The projections are either formed during forging of the nuts or by machining.

The challenges for the numerical modeling of a square nut to a sheet are due to the collapse of the square nut legs (projections) when they are sufficiently softened by the heating.

9.2.1 Experimental analysis

The experiments are performed by M10 square nuts and simplified models hereof welded to 1.5mm thick DC06 steel sheets of size $60 \times 60 \text{mm}^2$. The standard square nuts follow the shape specified by the standard *DIN 928 – Square nut weld specifications*. The standard specifies a recommended range of sheet thicknesses, 1.25mm – 5mm, that can be joined with the square nut. The chosen thickness, 1.5mm, is in the lower end of this range as it resembles sheet thicknesses commonly used in the automotive industry.

The shape of the nuts can vary significantly within the standards. Of particular relevance to projection welding is the shape of the legs, which can be either formed by forging (as the nuts utilized in this work) or they can result from machining, where the latter produces a circular edge on the inside due to turning. Furthermore, the initial contact area can vary significantly because it is not specified whether the tip of the leg should be flat or rounded. The forged ones are typically rounded.

The uncertainty in the geometry may be acceptable for projection welding because the area towards the bulk part of the nut may be of more importance than the exact tip geometry because the leg collapses in the early stages of welding and produces the necessary area towards the sheet. However, for a comparison with numerical simulations, the exact geometry of the legs counts because it influences the final shape of the leg that is compared with the simulation. Square nut models are therefore produced with well-defined shape of the legs as shown in Fig. 9.18. The shape of the legs follows the specifications set by *DIN 928*, although deviating from the standard nuts. The legs are produced with a flat contact face towards the sheet, such that the initial contact area is known, and the hole is produced without thread.

While the material of the standard square nut is only specified by its class (here class 8), the models of the square nut are made in known materials; namely the three test materials also used for the cross-wire welding and the contact experiments of Section 8.1. The test materials used for the nut are aluminum AA6060-T6, structural steel S235JR+AR and stainless steel AISI 316L.

The setup for projection welding of square nut to sheet is shown in Fig. 9.19. The same guiding system and the same electrodes as for the cross-wire welding are uti-

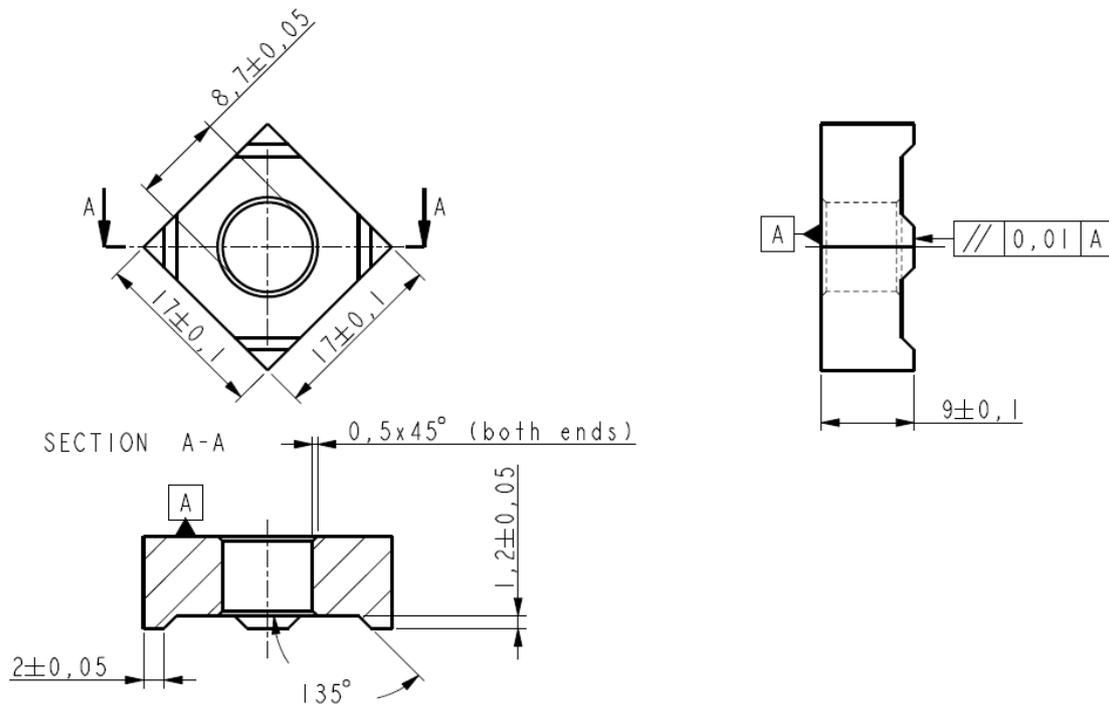


Fig. 9.18. Geometry of simplified square nut models.

lized. The bottom electrode has an additional feature in form of a centering pin placed in a hole drilled through the electrode center. The centering pin is shown in Fig. 9.19a together with its function, which is to center the square nut (exemplified by the standard nut in the figure). The nut is placed on the bottom electrode with its legs upwards. The sheet is placed in the guidance system as shown in Fig. 9.19b on top of the nut, such that it is resting on the legs of the square nut. Fig. 9.19c shows a photograph of the welding process.

The square nut welding was preliminary tested in both the Tecna (AC current and pneumatic force control) and the Expert (middle frequency current and hydraulic force control with disc springs). Welding of the square nuts to the sheet was possible in both systems. The Expert welding machine was chosen for the experiments for obtaining more stable welding conditions in terms of constant current and force for the comparison with numerical simulations. This choice was made without considering weld quality, which Tolf and Hedegård [11] reported to be enhanced on an AC-machine compared to a middle frequency machine.

An example of the measured process curves is given in Fig. 9.20, where Fig. 9.20a shows the measured current obtained with a machine setting of $11kA$ over $180ms$. The average measured current is $10.6kA$. The measured force is $6kN$ before the current is applied. When the current is applied, the force is increased due to thermal expansion for a short period of time until the softening results in a lowered force. The force stabilizes towards $5.4kN$ resulting in an average force of $5.6kN$ during the welding time.

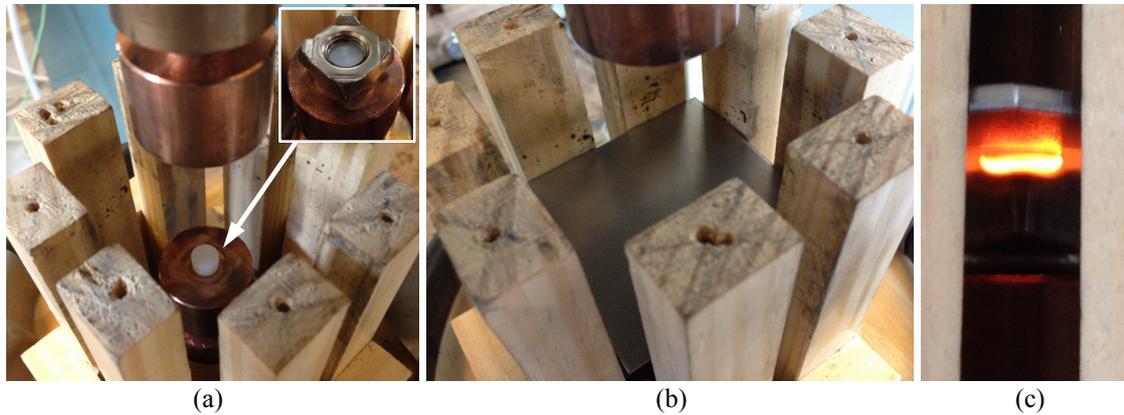


Fig. 9.19. Resistance projection welding of square-nut to sheet in the same guidance and overall setup as presented in Fig. 9.1 for the cross-wire welding. (a) View to the two flat C0-type electrodes, where the lower has a centering pin placed in a hole drilled through the electrode center. The subfigure in the upper right corner shows the positioning of the square nut on the lower electrode via this centering pin. (b) Positioning of sheet by centering and alignment in the guidance support. (c) Photograph of the welding process taken between two of the guides of the support system. Only one of the four legs of the nut is visible.

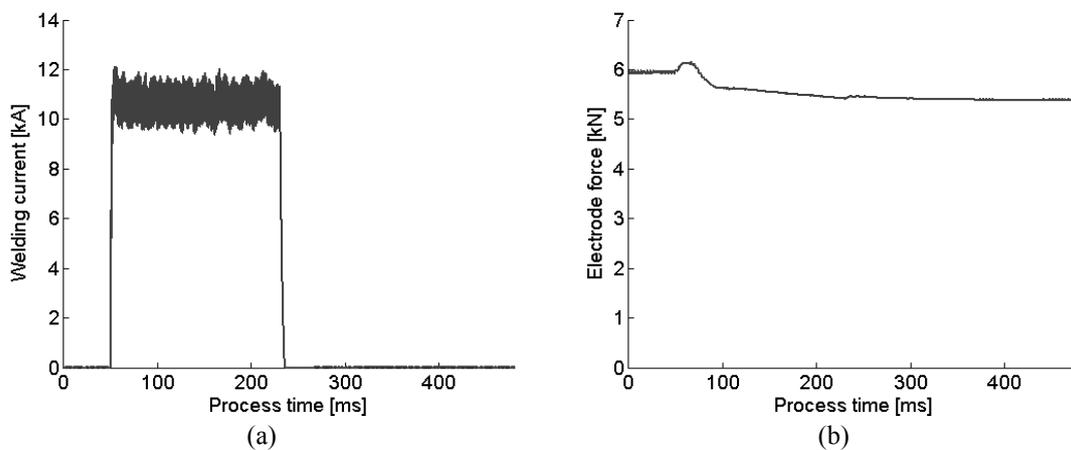


Fig. 9.20. Example of measured welding current (a) and force (b) as function of process time.

Resulting square-nut welds

The setdown for the evaluation of the square-nut welds is defined as follows with parameters given in Fig. 9.21;

$$S = \frac{A - B}{A} \cong \frac{A - (h - c - t)}{A} \quad (9.2)$$

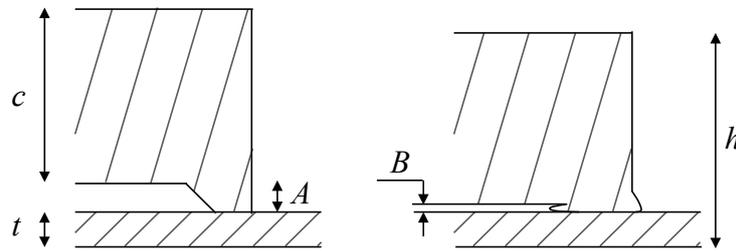


Fig. 9.21. Parameters A and B for definition of setdown and c , h and t for practical measurement of B .

The definition is a measure of the degree of collapse of the projection and it is practically obtained by the expression in the right side of (9.2). As a result of this, the collapse is not only measuring the collapse of the projection, but also the indentation into the sheet and the bulk deformation of the nut itself near the projection. Global bulk deformation of the nut is considered negligible.

Fig. 9.22 presents the obtained setdown for all the square-nut welds performed on the Expert welding machine. Fig. 9.22a contains all the steel nuts, including the standard nut and the AISI 316L and S235JR+AR nut models, and Fig. 9.22b-c contain the results obtained with aluminum AA6060-T6 nut models. All nuts fall along the same path on the curve despite the different materials and despite the difference in geometry of the projection tip between the standard nut and the nut models. This suggests that the projection height and the area of the projection towards the bulk of the nut are governing the setdown of steel nuts. They are the two factors kept constant among all nuts, being standard nuts or nut models. The overall path shows increasing setdown with increasing current due to the larger heat input and resulting softening of the material with the projection in particular.

The aluminum nuts are in another range of material properties than the steels, including the DC06 deep drawing steel sheet that they are welded to. As a result of these differences, it is generally difficult to weld aluminum to steel, and so is the case in the square nut welding. The relevance of having an aluminum nut may also be questioned. However, for comparison purposes with the developed simulation software, it is interesting to evaluate the reaction when dealing with such different materials. The current was initially varied for the evaluation of different levels of setdown (Fig. 9.22b), but even at 100% setdown, there was limited bonding between the aluminum nut and the steel sheet. Alternatively, the influence of welding time was varied (Fig. 9.22c), but still with poor bonding. In fact, all weld settings result in poor bonding between the aluminum nut and the steel sheet, and they could easily be separated by hand.

Two points on each curve in Fig. 9.22a are selected for analysis by their cross-sections and three points are selected among the welds including aluminum nuts in Fig. 9.22b-c. The cross-section is taken similar to the cross-section shown schematically in the lower left part of Fig. 9.18. Among the nuts in Fig. 9.22a, points A-B are selected for the standard nuts, points C-D are selected for the stainless steel AISI 316L nut models and points E-F are selected for the structural steel S235JR+AR. For the aluminum nut models in Fig. 9.22b-c, points G-I are selected for analysis.

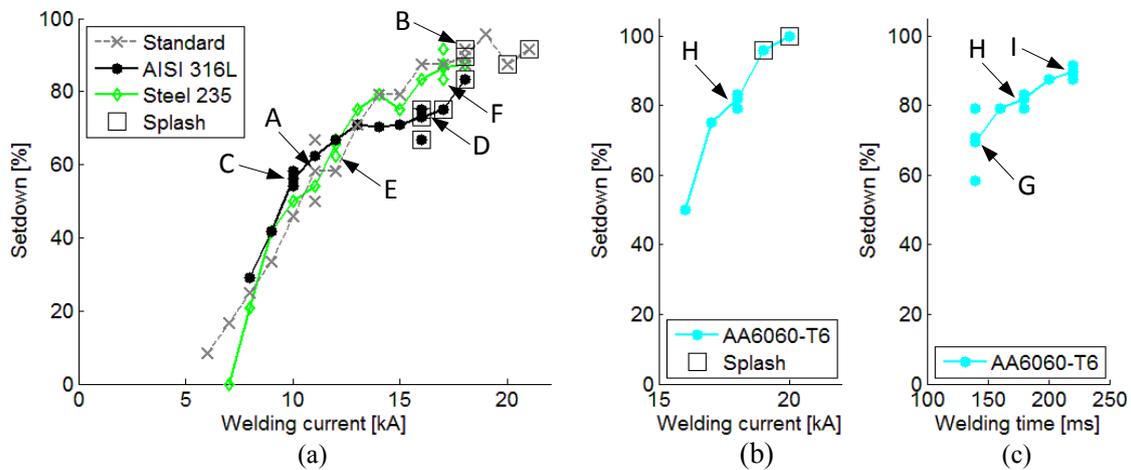


Fig. 9.22. Setdown for welds of square nut to sheet. (a) Setdown as function of current for standard nuts and stainless steel AISI 316L and structural steel S235JR+AR nut models welded to deep drawing steel DC06 sheet. (b-c) Setdown for welding of aluminum AA6060-T6 nut model to DC06 sheet (b) as function of current at constant weld time 180ms and (c) as function of weld time at constant current 18kA .

Standard square nut

The selected nut welds including the standard square nut, A-B, are shown in Fig. 9.23 by their cross-sections. Both cases show a heat affected zone covering the projection and part of the bulk material of the nut near the projection, while little or no heat affected zone is observed in the sheet. This is a result of high heat conduction away from the sheet through the bottom electrode with large contact area towards the sheet and high heat conductivity compared to the steels as the electrode is a copper alloy.

The shape of the interface between the sheet and the nut shows a waved sheet surface resulting from the initial indentation of the rounded square nut leg into the sheet because of the small initial contact area. Later softening of the material has resulted in collapse of the projection and material flow out of the initial contact area. On the outside of the nut, the material flow has been free in terms of normal pressure towards the sheet and as result of that, there is no bonding between the nut material and the sheet in this zone as depicted by the magnifications in Fig. 9.23e-f.

The cross-sections do not reveal any melting during the welding process, but it is possible that the initial small contact area has resulted in local melting, which has been squeezed out together with following material flow. Comparing the flow towards the outside of the nut in Fig. 9.23c-d shows that in weld B with the higher current, the flow has been more fluid-like because the flow has resulted in a narrower shape despite the low load in this region. It is not clear if the material has been melted or if it has been close to melting but still in the mushy state. The latter can be the case, and in any case it seems that the weld, like in the case of cross-wire welding, are close to friction welding facilitated by the high normal pressure and elevated temperatures.

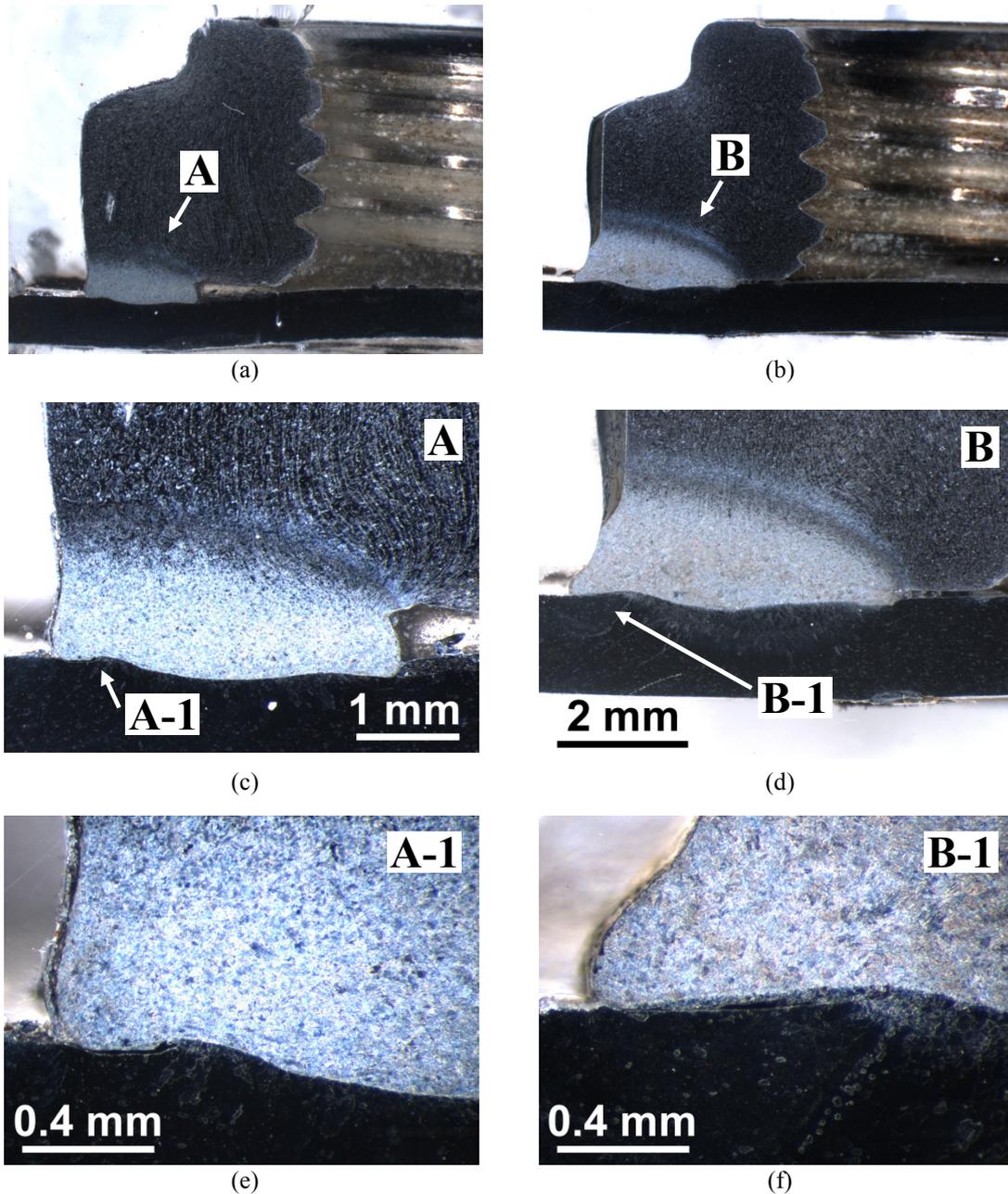


Fig. 9.23. Cross-sections of selected standard square nuts welded to a DC06 steel sheet. (a) Weld A with 58% setdown. (b) Weld B with 92% setdown. (c-d) Magnifications with positions marked in (a-b). (e-f) Further magnifications with positions marked in (c-d).

Tolf and Hedegård [11] conclude similar on the basis of their experiments with square nut welding by observing *that the joints were created more by forge welding than fusion welding.*

Stainless steel AISI 316L

Cross-sections including the selected stainless steel nut models, C-D, are shown in Fig. 9.24. As an immediate comparison between the standard nuts and the nut models, it is

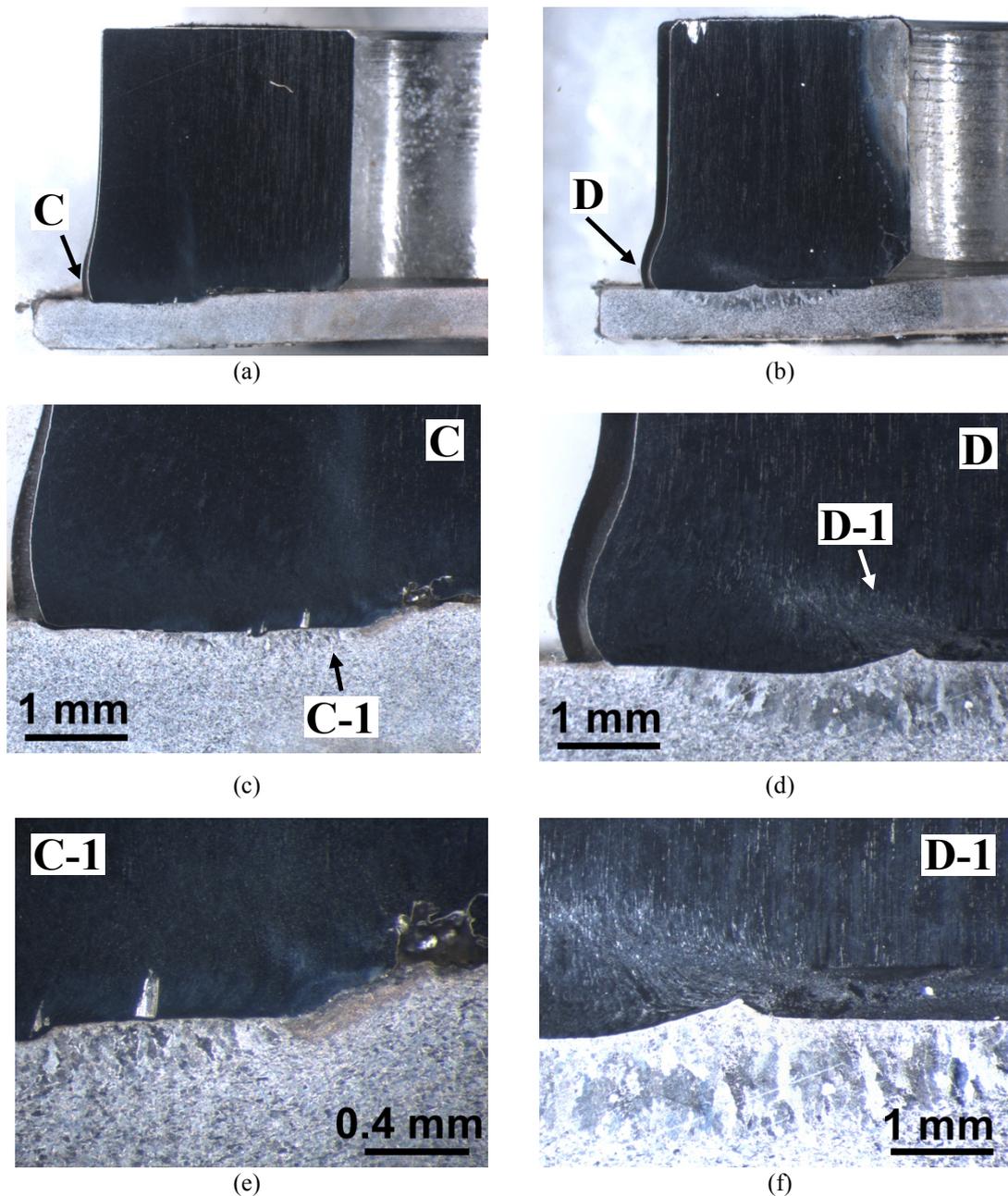


Fig. 9.24. Cross-sections of selected stainless steel AISI 316L square nut models welded to a DC06 steel sheet. (a) Weld C with 54% setdown. (b) Weld D with 75% setdown. (c-d) Magnifications with positions marked in (a-b). (e-f) Further magnifications with positions marked in (c-d).

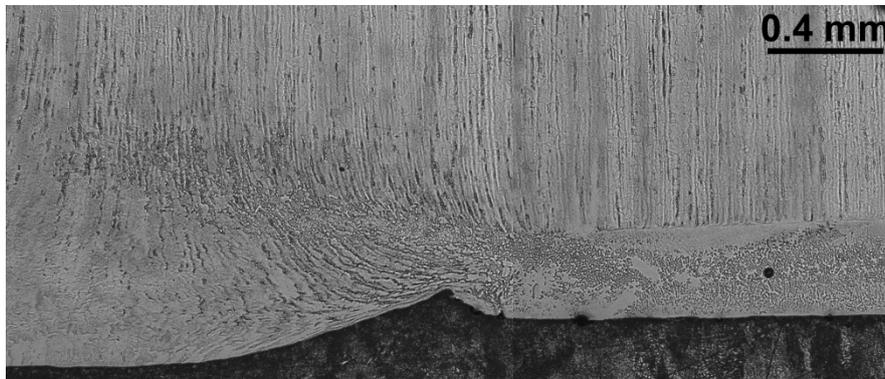


Fig. 9.25. Magnified detail of Fig. 9.24f showing material flow from the collapsed projection underneath the nut itself towards the nut center (towards right in the figure).

seen that the models are fabricated without thread and that the overall shape is more regular for simplifications. The major change is that the initial projection is flat in the contact interface with the sheet (as drawn in Fig. 9.18).

The relatively harder stainless steel compared to the deep drawing steel results in an indent in the sheet, although the initial contact area is larger than in the above case. It is visible in both cases in Fig. 9.24 and more distinct in weld D with higher heat input; see Fig. 9.24d. The shape of the projection shows restriction of outward flow due to friction, which results in a barreling effect of the resulting shape. This effect is visible in all the nut welds, but is more evident in case of the stainless steel. This may be a result of a larger normal pressure built up before the major material flow.

In Fig 9.24.d, it appears that there has been melting in the inside of the nut projection, which has caused material flow away from the projection towards the center of the nut. This flow, which is magnified in Fig. 9.24f, is between the bulk part of the nut and the sheet. An even closer magnification of this flow is provided in Fig. 9.25, where the microstructure illustrates the underlying material flow. The melting has taken place in the stainless steel, potentially a bit above the interface because of the cooling effect of the electrode, and it is squeezed out by the applied electrode force. The material flow passes the sheet material piled up due to prior indentation and continues below the nut towards right in the figure.

Structural steel S235JR+AR

Fig. 9.26 includes the cross-sections for the structural steel nut models, E-F. It appears that in this case there has been no melting and that the flow has been accommodated by softening of the material by temperatures resulting in austenitization of the entire projection and part of the bulk material of the nut. In the magnifications (especially Fig. 9.26c) pieces of the original projection material are visible between the bulk part of the nut and the sheet. It is likely that the initial sawing in the preparation of the cross-section has resulted in pieces breaking off from the projection. The austenitization and subsequent rapid cooling of the material through the sheet and the underlying

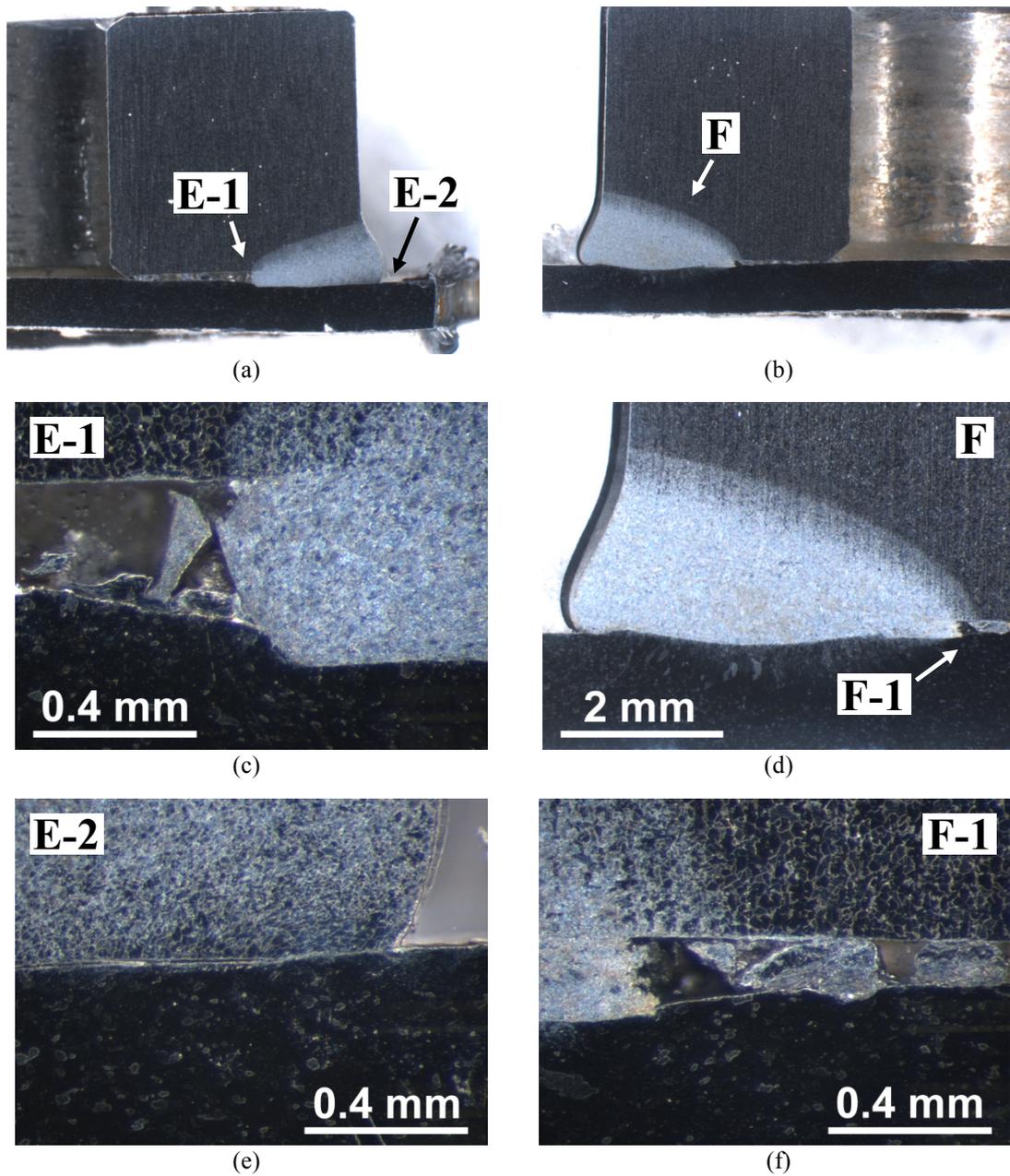


Fig. 9.26. Cross-sections of selected structural steel S235JR+AR square nut models welded to a DC06 steel sheet. (a) Weld E with 63% setdown. (b) Weld F with 83% setdown. (c-e) Magnifications with positions marked in (a-b). (f) Further magnification with position marked in (d).

electrode may have turned the material more brittle, such that those pieces were easily broken off.

Aluminum AA6060-T6

As mentioned earlier the examples of aluminum nut models welded to the steel sheet are less relevant for industrial applications, but they are relevant for comparison to the numerical simulations. Fig. 9.27 shows the resulting cross-sections for welds G-I.

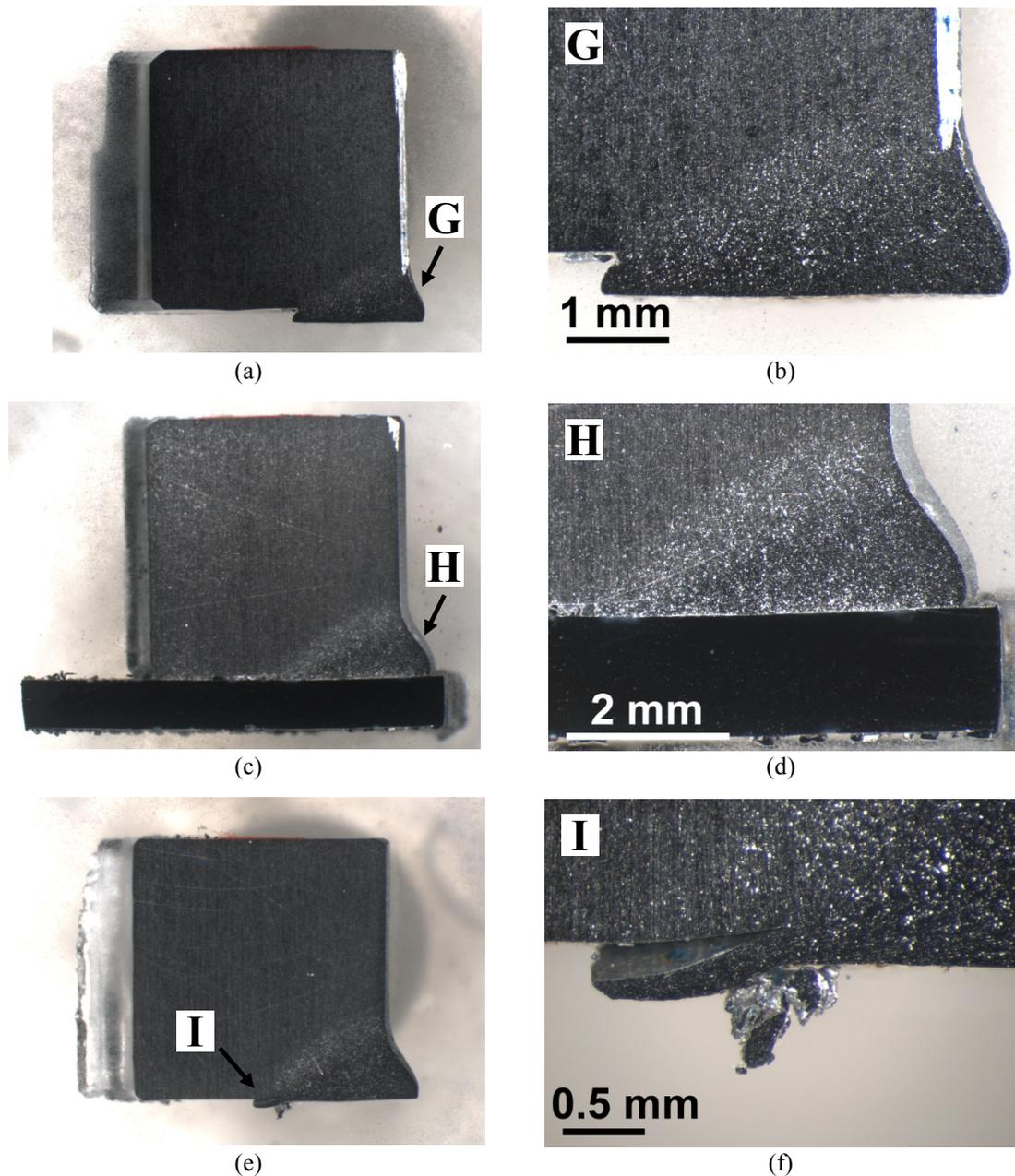


Fig. 9.27. Cross-sections of selected aluminum AA6060-T6 square nut models welded to a DC06 steel sheet. (a) Weld G with 71% setdown and (b) magnification hereof. (c) Weld H with 83% setdown and (d) magnification hereof. (e) Weld I with 90% setdown and (f) magnification hereof.

Compared to the above welds including steel nuts, an immediate difference is that none of the aluminum nut models were welded to the sheets with sufficient bonding. Out of the three presented cross-sections, only one of them got through the preparation of the cross-sections without falling apart from the sheet, and the repetitions made under similar weld settings reveal that the sheet and the nut model only remained together by very careful preparation. All weld settings do therefore result in no practical strength. Because of the difference in the materials, the aluminum is softened and deformed without remarkable deformation of the sheet surface.

9.2.2 Numerical modeling

The projection welding of the square nut models to the sheet is simulated by the finite element mesh shown in Fig. 9.28. Two vertical symmetry planes are identified such that one quarter of the full geometry is simulated. The amount of bulk material in the nut models compared to the projections is quite large, and for that reason it is justified that the electrode towards the nut will have only limited effect during the short process times. Therefore, the model only includes the electrode on the sheet side. This electrode has on the other hand a large effect on the process because of the large heat conduction. The model further includes thin layers of elements between this electrode and the sheet and between the sheet and the nut to simulate the interface properties.

Simulations and experiments are compared in Fig. 9.29 by their cross-sections de-

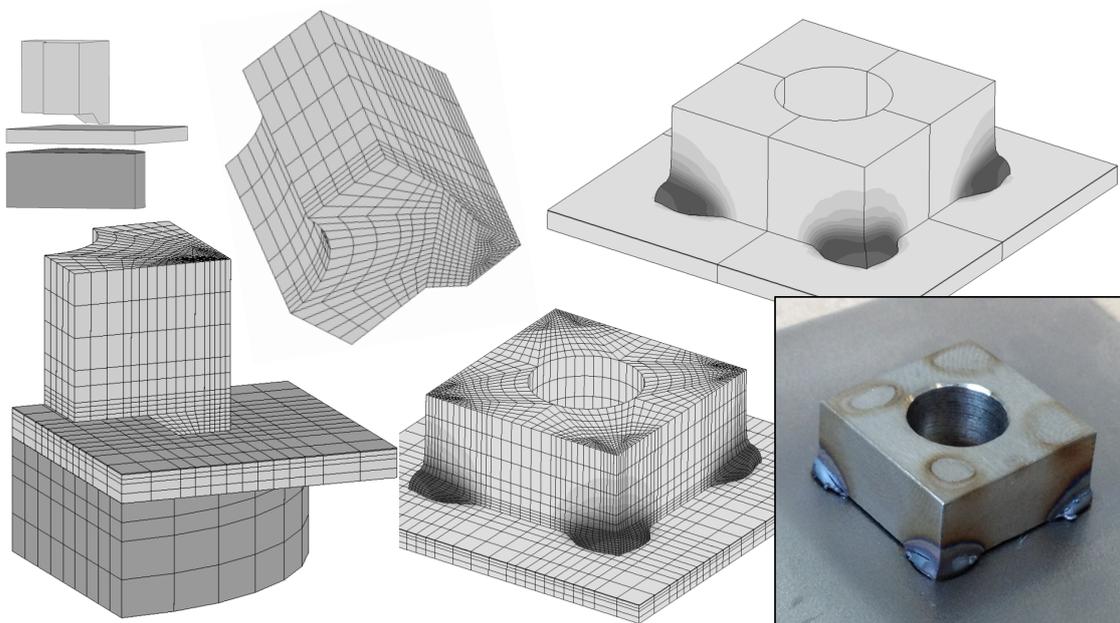


Fig. 9.28. Finite element mesh of square nut, sheet and lower electrode for simulation of projection welding. One quarter is simulated by utilization of symmetry planes. The two figures showing the full nut show examples of simulated geometry and temperature field. A picture of a real weld is also included.

defined as section A-A in Fig. 9.18. The comparisons are made for the low weld settings and include the square nut models represented by each of the materials. The simulations include both frictionless conditions and full sticking conditions for a comparison. The stainless steel (weld C in Fig. 9.22a and 9.24a) is shown in Fig. 9.29a, the structural steel (weld E in Fig. 9.22a and 9.26a) is shown in Fig. 9.29b and the aluminum (weld G in Fig. 9.22c and 9.27a) is shown in Fig. 9.29c.

Common for the stainless steel and the structural steel (Fig. 9.29a-b) is that simulation with full sticking appears to represent experiments better than simulation with frictionless conditions in contact interfaces. When simulating the contact between nut and sheet as frictionless, the nut leg is much easier collapsed because of the absent of

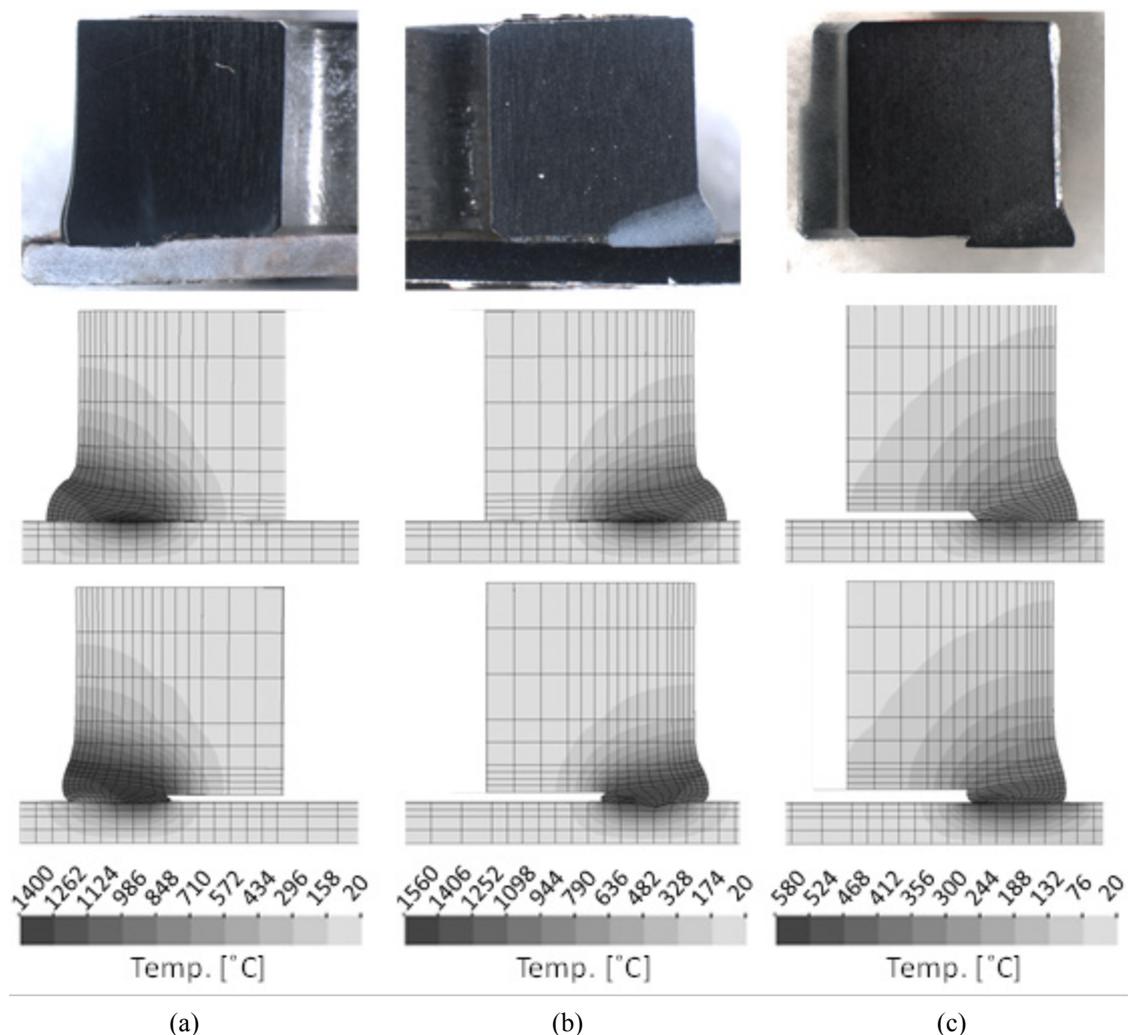


Fig. 9.29. Comparison of experimental (upper) and simulated cross-sections by assuming frictionless contact (middle) or full sticking (lower) for (a) AISI 316L stainless steel nut model, (b) S235JR+AR structural steel nut model and (c) AA6060-T6 aluminum nut model. All nut models are welded to DC06 steel sheets and the weld settings are the low settings of the experiments.

restriction to the outwards material flow. The simulated stainless steel nut welding (Fig. 9.29a) has almost complete collapse of the nut leg (99.6%), while the simulation with full sticking results in 80% setdown. When compared to the experiment with only 54% setdown, it is clear that the full sticking assumption resembles more accurately the experiment. A comparison of the final height between the experiment and the simulation with full sticking reveals a height difference of 0.31mm . When it comes to the structural steel nut weld (Fig. 9.29b), the simulated setdown under frictionless conditions is 91% and under full sticking conditions it is 71%. The assumption of full sticking is again closer to the experiment, which shows a setdown of 63%. The difference in the final height between the experiment and the simulation including full sticking is 0.096mm .

Comparison of the overall shape supports the above conclusions. The simulations with assumption of frictionless contact show unrestricted material flow away from the center, whereas the experiments include barreling of the nut legs because of restricted material sliding along the sheet because of high friction at elevated temperatures and the initial indent in the sheet that provides restriction to material flow. The indents are not as large as the photos may reveal at first because the sheets are bended during preparation of the cross-sections. The simulations with assumption of full sticking reproduce the barreling effect and some material flow towards the center of the nut.

A natural issue to address is that the two extremes in terms of frictionless and full sticking assumptions are not resulting in simulated setdowns being larger and smaller than the experiment, respectively, but both result in too large simulated setdown. In other words, it appears natural that the simulations with frictionless contact result in too large setdown, but it may not appear natural that the simulations with full sticking also result in too large setdown. An explanation from a process point of view is that the degree of softening of the material has been larger in the simulation with full sticking because of the resulting smaller contact area, and during the process simulation, this has resulted in larger setdown. On top of that are more trivial explanations as the modeling of the material properties, especially at high temperatures.

To conclude the discussion of the two steel nut model welds, the simulated temperatures are evaluated. The simulated peak temperature in the stainless steel nut is 1574°C with frictionless treatment and 1692°C with full sticking, which is 174°C or 292°C above the melting temperature of the stainless steel. This, however, is only in a few nodes in the interface between the nut and the sheet. Hence, the overall amount of simulated heat is expected to be in correspondence with the experiments, where no melting was observed. The same conclusion holds for the simulated structural steel nut model, where the peak temperature reached in a few nodes is above the melting temperature. The simulated process peak temperature is 1716°C with frictionless treatment and 1731°C with full sticking corresponding to 156°C or 171°C above the melting temperature.

The aluminum square nut model welded to the sheet shows a different behavior than the above steels. The barreling effect on the outside of the nut leg is less than observed for the two steels, while a larger amount of material flow is observed towards the cen-

ter below the bulk part of the nut. In terms of setdown, the simulation with frictionless treatment of contact is the most accurate prediction with 72% setdown, which is to be compared to 71% setdown in the experiment. The difference in the final height is only 0.012mm . The simulated setdown with full sticking in the contact interfaces deviates more, but is still relatively close taking into account the small nut legs. The simulated setdown is in this case 61% with a final height difference of -0.12mm . Note the negative sign indicating less simulated than experimental setdown as opposed to all other simulations showing too large setdown.

The compared setdowns indicate better simulation by frictionless conditions. This can be related to the absence of any significant indentation in the sheet caused by the aluminum and also because the friction between the aluminum and the steel sheet is less than between the above hot steel to steel contact. Complete satisfaction by simulation with frictionless contact is not the case as seen when comparing the resulting geometry. Due to friction in the real case, part of the material flow of the deforming nut leg is towards the center below the nut. This is not simulated with frictionless treatment of the contact.

Simulation with full sticking shows the effect of material flow towards the center, but only due to barreling and not due to frictional sliding, and hence the effect is much less than in the experiment. The overall material flow is therefore overly constrained and the resulting setdown is simulated too small under full sticking.

The simulated peak temperature with frictionless and full sticking treatment are 514°C and 583°C , respectively, which is 66°C below and 3°C above the melting point of the aluminum. As above, the peaks are local temperatures in the interface. The simulation shows that the nut material do not melt, which is in agreement with the experiment where the projection keeps strength enough for supporting the applied electrode force.

References

- [1] Scotchmer N (2007) The other resistance process: Cross wire welding. *Welding Journal* 86(12):36-39.
- [2] Goodman IS (1950) Variables in cross-wire welding of dissimilar metals. *Welding Journal* 29(10):863-875.
- [3] Knowlson PM (1967) Welding in electronics. *British Welding Journal* 14(7):398-400.
- [4] Jordan RH (1964) Productivity in multiple cross-wire welding. *Welding and Metal Fabrication* 32(1):19-26.
- [5] Resistance Welding Manufacturers' Association (2003) Resistance welding manual. RWMA, Philadelphia.
- [6] Batten GH (1958) Some aspects of cross wire welding of aluminium alloys. *British Welding Journal* 5(9):417-420.
- [7] Thornton PH, Krause AR and Davies RG (1996) Contact resistances in spot welding. *Welding Journal* 75(12):402s-412s.

- [8] Pan R, Watt DF (1995) Simulating microstructure development in high-carbon steel cross-wire welding. *Welding Journal* 74(12):385s-395s.
- [9] Fukumoto S, Tsubakino H, Zhou Y (2006) Heat input and deformation control in resistance microwelding of fine nickel wires. *Welding International* 20(9):692-697.
- [10] Khan MI, Kim JM, Kuntz ML, Zhou Y (2009) Bonding mechanisms in resistance microwelding of 316 low-carbon vacuum melted stainless steel wires. *Metallurgical and Materials Transactions A* 40A:910-919.
- [11] Tolf E, Hedegård J (2007) Resistance nut welding: Improving the weldability and joint properties of ultra high strength steels. *Welding in the World* 51(3/4):28-36.

10. Applications

Following the previous chapter with discussions related to local deformation, load bearing volume with high degree of softening and frictional aspects of sliding surfaces, this chapter presents applications with greater degree of success. The presented examples include spot welding, projection welding and a micro joining process. The examples range from pure numerical studies to industrial cases with experimental verification.

10.1 Resistance spot welding

Resistance spot welding is a key technology in automotive assembly production, and it is by number the most used welding process. According to Zhu et al. [1], more than 200 sheet metal parts are spot welded together resulting in 4000-7000 spot welds of two and three sheet combinations in each car.

The development of new materials (such as e.g. advanced high strength steels (AHSS)) presents challenges to the resistance spot welding process when combined with other materials. These new steel types are often used in supporting parts of the car and in safety parts that are designed to absorb the impact of a crash. The parts are typically joined to considerably thinner and softer low-carbon sheet materials that act as the outer panels of the car. There is therefore an increasing trend of assembling three sheets by spot welding, which typically involves two thicker, high strength steels and one, thin mild steel as one of the outer sheets. This combination has attracted a lot of attention because of the difficulties in attaining a weld nugget at both interfaces as illustrated by Nielsen et al. [2].

The following three subsections deal with different challenges in resistance spot welding. An example consisting of three sheets as described above is dealt with in the first subsection by comparison of simulation and experiment. The second subsection elaborates on the same example by showing the effect of electrode misalignment, which is an important issue in production where the flexibility of the welding gun arms can result in a slight rotation of the electrodes. Another complication in industrial spot welding is the shunt effect between two consecutive spots. This is illustrated in the third subsection by spot welding a two sheet assembly.

10.1.1 Three sheet spot welding

As already outlined above, spot welding of three sheets is the main challenge in automotive spot welding. Two thicker, high strength steels and a thin, low carbon steel is the typical combination. The specific combination chosen in the present example consists of a 1.5mm DP600 dual phase steel (advanced high strength steel) as the bottom sheet, a 0.8mm HSLA340 (high strength low alloy) steel in the middle and a 0.6mm DC06 (low carbon deep drawing steel) as the top sheet. This combination is welded between two type B conical electrodes with tip diameter $\varnothing 8\text{mm}$ towards the DP600 and tip diameter $\varnothing 6\text{mm}$ towards the DC06 as illustrated in Fig 10.1a with its finite element discretization and numerically predicted weld nugget in Fig. 10.1b. This example is reproduced from Nielsen et al. [2] with weld settings as follows. The weld force is constant 3.5kN and the weld current is applied during 180ms at 7.2kA RMS through an AC welding machine with estimated conduction angle of 75% .

As the DC06 sheet is considerably thinner than the DP600 sheet, the interface between the DC06 and the HSLA340 is located closer to the neighboring electrode than the interface between the DP600 and the HSLA340 is to its corresponding neighboring electrode. This results in larger heat conduction to the upper electrode and thus an asymmetric heat distribution. In the particular case (Fig. 10.1b), the heat input was too small to create a nugget that develops into the thinner sheet. On the other hand, if the heat input was too large, splash would be likely to occur between the two thicker sheets, leading to uncontrollable material removal, loss of strength, and excessive electrode wear. Compared to welding of two sheets, these restrictions result in a rather narrow window of applicable weld settings.

The simulated temperature distribution and weld nugget are compared to the corre-

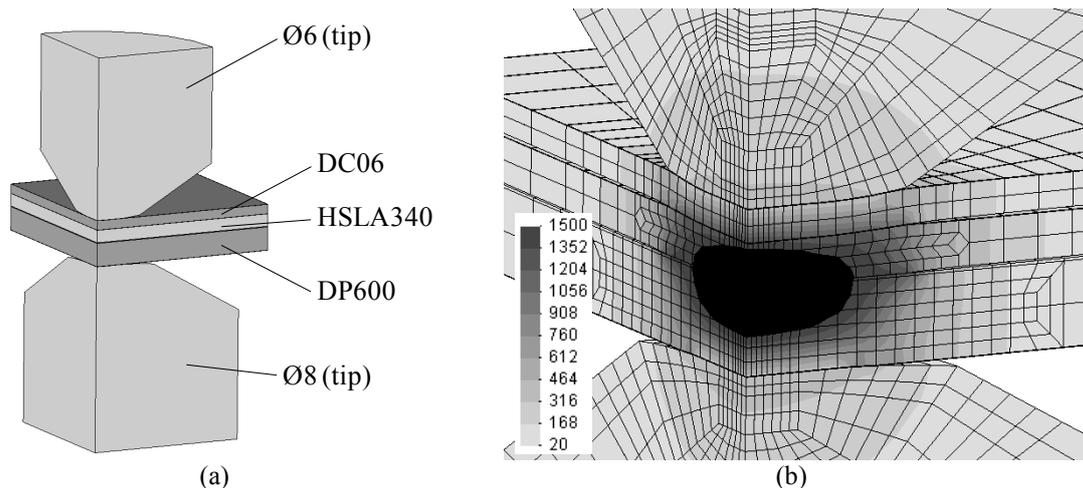


Fig. 10.1. Example of three sheet spot weld consisting of a thin 0.6mm DC06 steel sheet, a 0.8mm HSLA340 steel sheet and a 1.5mm DP600 steel sheet welded between two type B electrodes with tip diameters $\varnothing 6\text{mm}$ and $\varnothing 8\text{mm}$. (a) Quarter of the geometry showing material combination. (b) Detail showing the finite element predicted temperature field with scale bar in degrees Celsius.

sponding experiment in Fig. 10.2a (cf. [2]). The overall weld nugget size is matching between the experiment and the simulation, and of specific interest in this case is that the finite element simulation reproduces the fact that the nugget does not develop into the thin sheet. This is in many cases a reason to reject the weld settings in order to achieve a weld nugget that covers both interfaces.

Due to the narrow window of appropriate weld settings (if any), innovative solutions have been developed to initiate the weld nugget in the interface towards the thin sheet as noted by Nielsen et al. [2], who at the same time proved that plug failure mode (the

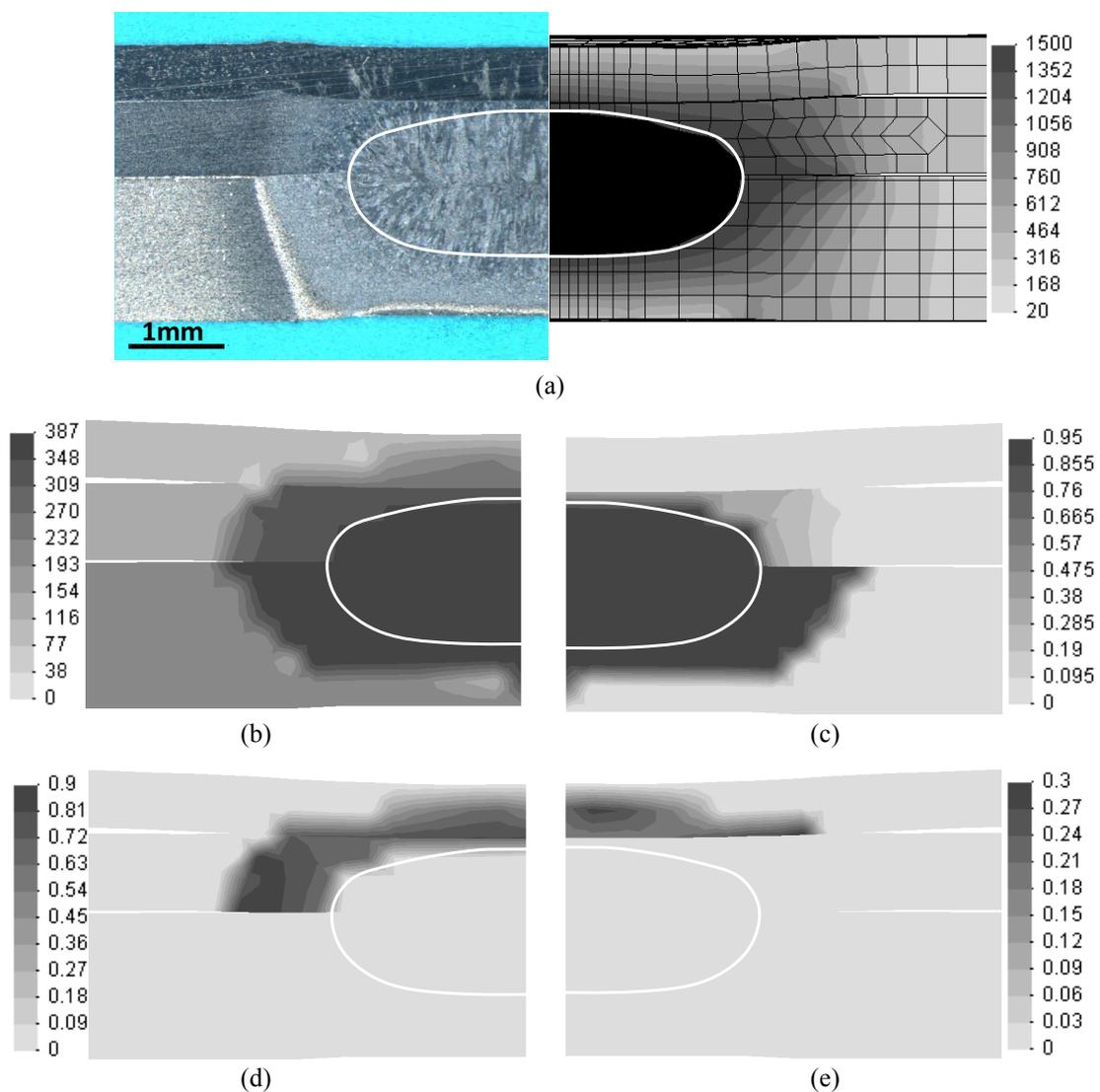


Fig. 10.2. Three sheet spot welding industrial test case. (a) Comparison between experimental and simulated maximum temperature distribution (degrees Celsius) with indication of simulated weld nugget mirrored onto the experiment. (b) Simulated hardness distribution in Vickers. (c) Simulated martensite distribution. (d) Simulated bainite distribution. (e) Simulated pearlite distribution.

desired failure type in tensile-shear testing) can be achieved without melting into the thin sheet. The strength of the weld interface towards the thin sheet is in those cases achieved by solid state bonding facilitated by heat and plastic deformation.

Having the temperature history simulated including maximum temperatures and cooling rates as well as knowing the compositions of the base materials [2], it is possible to calculate the resulting hardness distribution (Fig. 10.2b) and microstructure distributions. The individual fractions of selected phases of the microstructure are shown in terms of martensite (Fig. 10.2c), bainite (Fig. 10.2d) and pearlite (Fig. 10.2e) and demonstrate the potential of predicting the metallurgical behavior of materials. The simulations of hardness and microstructure are performed by the existing 2D version of SORPAS.

Following time-temperature-transformation (TTT) diagrams for the specific steels, the fractions of the different phases are found by comparison with critical cooling rates. Typically, and also in Fig. 10.2c, the center of the nugget consists mainly (here 95%) of martensite due to the prior full transformation into austenite followed by rapid cooling. Outside the nugget, the material may form bainite and pearlite depending on the initial composition and the actual cooling rates (Fig. 10.2d-e) or more martensite as in the DP600 steel.

The estimation of the quantities inside the nugget is complicated by the presence of more than one material. The contribution of each material to the combined microstructure and hardness distribution is evaluated by volume weighting assuming that the material inside the nugget is fully mixed in its molten stage.

The hardness is evaluated by the model by Blondeau et al. [3] based on the actual cooling rate and the carbon equivalent. More details of the evaluation of microstructure and hardness distributions can be found in the work by Pedersen et al. [4], who also compare experimental and simulated results.

10.1.2 Electrode misalignment

The above industrial case with three sheets will now be analyzed under the assumption of electrode misalignment, which is relevant to assembling in a production line. A potential source of electrode misalignment is the flexibility of the welding machine arms for positioning the electrodes. These arms are necessary in order to reach the locations of the spots on larger panels. Rotation of the electrodes can occur as illustrated in Fig. 10.3 when applying the electrode force through these arms.

A situation like the one illustrated in Fig. 10.3 is simulated with the same sheet setup and weld settings as in the above case. The electrode rotation is assumed to be 2.5° for each of the electrodes as illustrated in Fig. 10.4a showing the detail of Fig. 10.3. The finite element model including the electrode misalignment consists of 5542 elements and is shown in Fig. 10.4b with simulated process peak temperature and indication of the weld nugget. The weld nugget can be compared to Fig. 10.2a because all other parameters than the electrode misalignment are identical.

The resulting weld nugget is clearly asymmetric as a result of the angled electrodes, which are initially only touching the sheets on the outer edge until a certain indentation

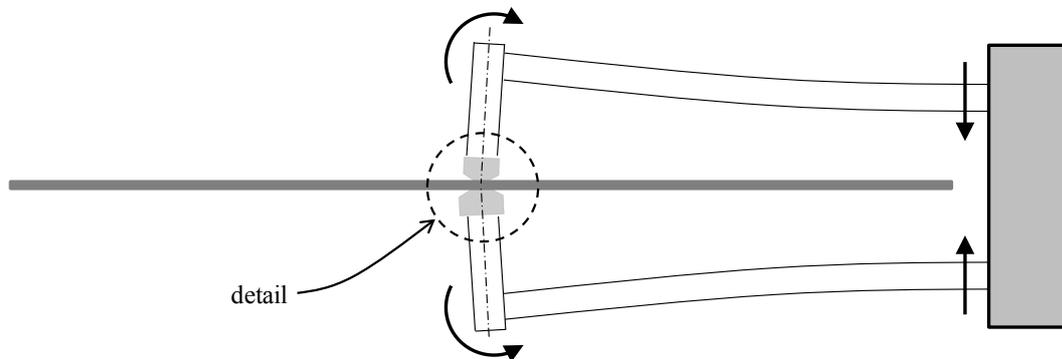


Fig. 10.3. Electrode misalignment due to rotation caused by flexibility of the arms of a welding gun typically applied in production in order to reach the location of the spots. The detail enclosing the electrodes is enlarged in Fig. 10.4a.

has developed. Fig. 10.4c shows a close up of the weld nugget as well as a clear angled indentation of the upper electrode into the thin low carbon steel sheet. By comparison to the symmetric weld in Fig. 10.2a, the indentation is more severe in case of electrode misalignment due to the small initial contact area.

The larger and localized indentation causes the thin sheet to lift more (right side in Fig. 10.4c) and complicates the overall assembly process because distortion can create relative movement of the sheets to a degree that makes the sheets off position at the location of following spots. This is already an issue under ideal conditions that need to be taken care of in the planning of the sequence of the welds. The procedure is further complicated by the additional distortion due to eventually misaligned electrodes.

The simulation also shows that the gap between the two high strength steels is increased by the introduction of angled electrodes, while at the same time, the nugget forms towards the gap opening. This increases risk of splash significantly, which would lead to uncontrolled joining conditions.

From the discussion in Section 10.1.1, it is clear that the chosen weld settings are too low to form a weld into the thin upper sheet, and that the weld settings should be increased (i.e. increased current/weld time or lowered electrode force). However, due to the lowered splash limit by the electrode misalignment, it might not be possible to increase the weld settings in this case. Formation of a weld nugget into the thin sheet may therefore be impossible in case of electrode misalignment (leaving out of account innovative solutions to initiate the weld nugget at the interface towards the thin sheet).

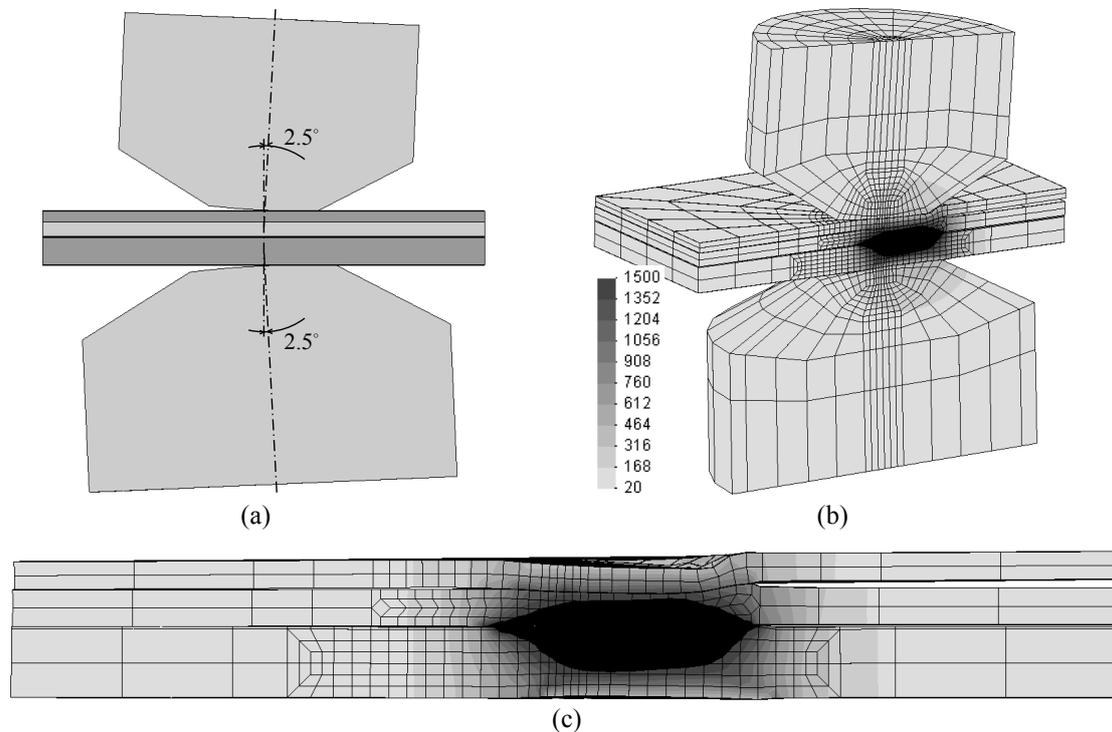


Fig. 10.4. Electrode misalignment in the three sheet spot welding case of Section 10.1.1. (a) Electrode misalignment by rotation of each electrode by 2.5° . (b-c) Simulated peak temperature distribution (degrees Celsius) with indication of asymmetric weld nugget and visible excessive electrode indentation of the upper electrode into the upper thin sheet (c).

10.1.3 Shunt effect

Shunt effect is taken as another complication occurring in industrial joining with multiple spot welds. The effect is considered in a case with two sheets welded between two type B electrodes (cone shaped as in Fig. 10.1a) with tip diameter $\varnothing 6\text{mm}$. The two sheets are chosen to be different steels with different thicknesses. The bottom sheet is a 1.2mm DP600 steel and the upper sheet is a 0.7mm DC06 steel.

The squeeze time is simulated as 40ms to reach the constant welding force 2.5kN . The AC welding current is kept constant at 8kA RMS for 160ms , such that the welding current of the first spot is ending at time 200ms (temperature field shown in Fig. 10.5). The electrode force is kept during a hold time of 80ms finishing the first weld at time 280ms . Hereafter follows 3s where the electrodes are repositioned to the location of the second spot (temperature fields at selected instants of time, 370ms , 1310ms , 2230ms , and 3190ms , during the repositioning are shown in Fig. 10.5).

The location of the center of the second spot is 12mm away from the center of the first spot. This distance corresponds to two electrode tip diameters, which is closer than the recommended minimum distance between spots [5]. This is chosen in order to

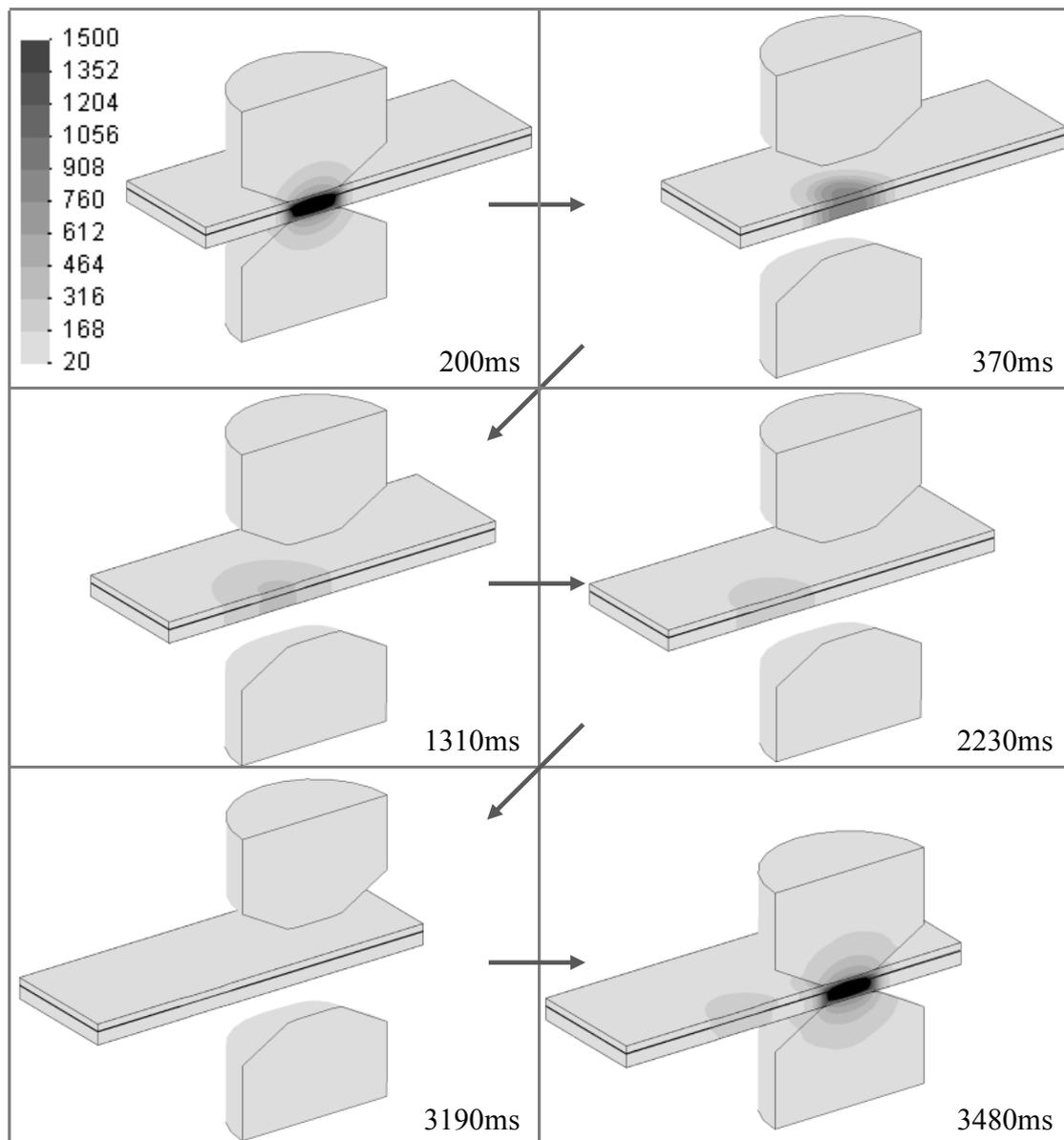


Fig. 10.5. Shunt effect between two consecutive spot welds illustrated by the temperature field (degrees Celsius) at different instants of time.

magnify the shunt effect to support the presentation. After moving the electrode, the squeeze time is again initiated and the same weld schedule is applied as for the first weld, implying that also the weld current is kept at the same level. The shunt effect is therefore not compensated by an increased current and a comparison of the weld nuggets will show the effect of shunting. The weld current of the second weld ends at time 3480ms (temperature field shown in Fig. 10.5).

The above referred temperature fields and instants of time are collected in Fig. 10.5 to give an overall representation of the shunt effect. The upper left temperature field

shows the ending of the first weld current. Following the arrows, the following four temperature fields illustrate the temperature evolution during the movement of the electrodes to the location of the second spot. The first weld cools while the surrounding sheet material is moderately heated due to heat conduction. The last temperature field shown in the figure corresponds to the ending of the second weld current. At this stage the second weld nugget has formed, but also the temperature in the first spot has risen as seen by a comparison between the two last instants of time. This is due to electrical heating caused by the shunting current flowing through the first spot while welding the second spot.

The shunting current is shown in Fig. 10.6a by the current density at the peak current of the third half cycle. While the majority of the current flows through the sheet interface at the location of the second spot, it is seen that a considerable amount of current flows through the first spot because of the absence of an interface after welding. At the location of the first weld, the current density is seen to be higher where the sheets start to separate towards the second spot due to the singularity. The amount of shunting current varies during the welding time of the second spot. The contact resistance between the sheets is larger at low temperatures indicating a larger shunting current in the beginning, but on the contrary the bulk resistivity increases with temperature, which indicates a larger shunting at the later stages because the material between the spots as well as the first spot remain at moderate temperature.

The peak temperature distribution achieved during the second weld is shown in Fig. 10.6b, where it is shown that the temperature in the first spot raises to 330°C during the second spot welding. As a result of the shunting current and the temperature increase in the first spot, less heat is dissipated in the second weld compared to the first one. This is also directly readable from the resulting weld nuggets shown in Fig. 10.6c, where the overall peak temperature distribution of the entire welding process is shown. The nugget sizes measured at the interface between the sheets are 4.84mm in the first spot and 4.18mm in the second spot, which is a decrease of 14%.

In a production line, where the shunt effect will play a role due to the necessity of having spot welds located close to each other, the current can be increased to compensate for the heat dissipated in the neighboring spot(s).

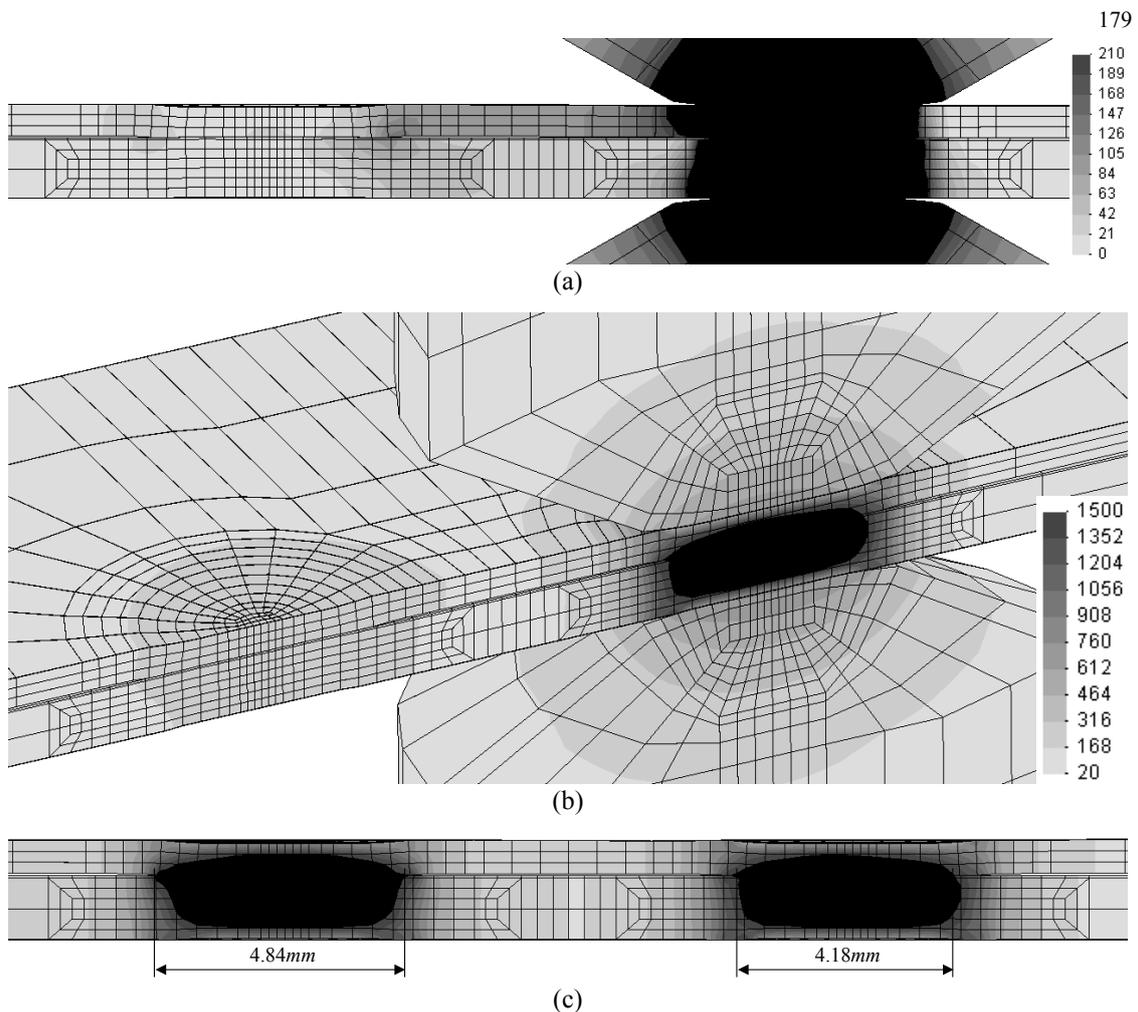


Fig. 10.6. Detail of the finite element predicted simulation of the shunt effect between two consecutive spot welds of two sheets. (a) Current density after $5/4$ cycles of the second weld current on a $0-210 A/mm^2$ scale out of maximum $1683 A/mm^2$. (b) Peak temperatures reached during the second weld showing the weld nugget achieved in the second weld while reaching $330^\circ C$ in the first spot. (c) Comparison of the two weld nuggets by overall peak temperatures (first spot to the left and second spot to the right).

10.2 Single-sided spot welding of sheet to tube

This section deals with another form of spot welding; namely single-sided spot welding, which is here applied to the joining of a sheet to a tube as presented by Nielsen et al. [6].

With intensive focus on weight and cost reductions in automobile manufacturing, it is becoming common to utilize hydroformed tubes as structural parts of the body-in-white as presented by Shah and Bruggemann [7]. The closed tubular shape provides

sufficient stiffness with relative low weight compared to conventional stamped parts and is therefore attractive in vehicle design. The utilization of closed tubular parts, however, presents new challenges to the assembling procedures as e.g. discussed by Poss and Lendway IV [8] and Cho et al. [9]. The resistance spot welding process is the preferred joining technique in automotive assembly lines due to production cost and efficiency, versatility and robustness, but the typical application of an electrode from each side of the weld is not feasible when welding sheet materials to closed tubular components. Single-sided spot welding is therefore utilized in order to facilitate welding of weld flanges or panels to the tubular structures.

In single-sided spot welding, a primary electrode is applied from the sheet side to impose the weld force and supply the weld current while a secondary electrode is introduced at another available location of the tube structure to supply electrical connection. This implies current flowing from the weld region to the secondary electrode with moderate heat generation in the intermediate material as a result. Different locations may be utilized for the secondary electrode. In some cases it may be convenient to place it from the opposing side of the tube, while in other cases it may be more convenient to place it on the same side of the tube as the primary electrode. An example of an experimental weld setup for the single-sided sheet-to-tube welding is provided by the laboratory of Dortmund FuE-A Pressfügen/Kleben, ThyssenKrupp, Germany (Fig. 10.7). The tube is resting in a V-support while the sheet is positioned and welded from the top by applied force and current through the primary electrode. A flexible secondary electrode consisting of copper blades can be positioned arbitrarily on the tube with good electrical connection ensured by the mechanical flexibility.

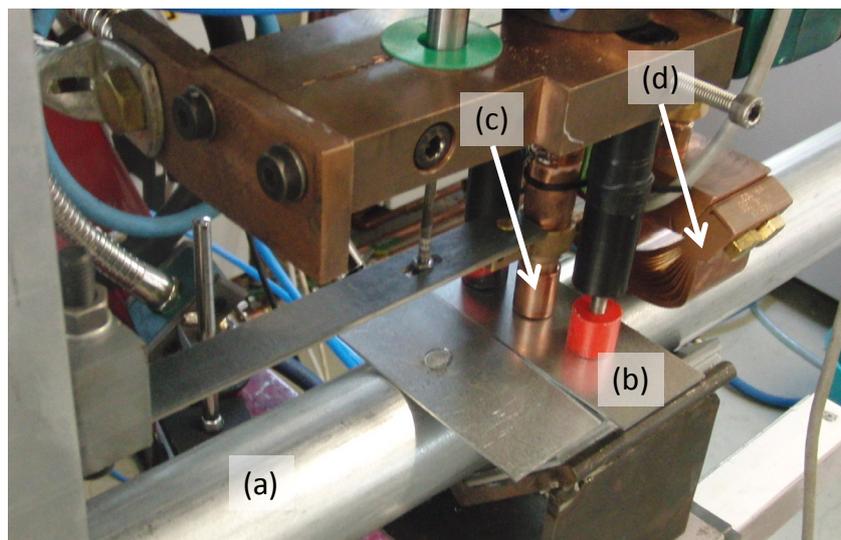


Fig. 10.7. Experimental setup in laboratory of Dortmund FuE-A Pressfügen/Kleben, ThyssenKrupp, Germany for single-sided spot welding of sheet to tube consisting of (a) tube, (b) sheet, (c) primary electrode and (d) secondary electrode.

The primary electrode is controlled by a servo gun as suggested by Sun and Wang [10]. It can supply a varying electrode force during the welding schedule. It is thereby possible to apply a larger electrode force in the early stage to ensure mature contact conditions between the sheet and the tube and to apply a lower electrode force in the later stage to avoid severe indentation due to softening of the sheet and tube without support on the inside.

Fig. 10.8 shows an example of a spot weld obtained in the experimental setup. The cross-section in Fig. 10.8a shows the indentation of the primary electrode into the sheet and corresponding local deflection of the tube. Fig. 10.8b shows plug failure when tearing the sheet apart from the tube, which indicates a sound developed weld nugget.

The single-sided spot welding process is preferred prior to alternative welding techniques as presented by Cho et al. [9]. Arc welding is inducing larger thermal distortion and presenting quality control issues, and laser welding is expensive. Both alternative welding procedures may face problems in case of a gap between the sheet and tube because there is no applied force to bring the parts in proper contact.

A number of contributions [8-9,11-15] present experimental analysis of single-sided spot welding of sheet to tube combinations with squared or cylindrical cross-sections of the tube. The work by Rudolf [11] is accentuated as a thorough investigation of different weld parameters, welding positions, sheet and tube thickness ratios, and electrode shapes for squared and cylindrical cross-sectioned tubes with analysis of heat generation, contact development and deformation. Common for the experimental investigations by different authors is the narrow weldability lobes reported. The electrode force is critical in obtaining a proper weld. If the force is too low, splash is likely to occur between the sheet and the electrode, while if it is too large, the indentation is easily too big because of the absence of support on the inside of the tube. The local and global deformation is therefore largely depending on the stiffness of the tube as well as

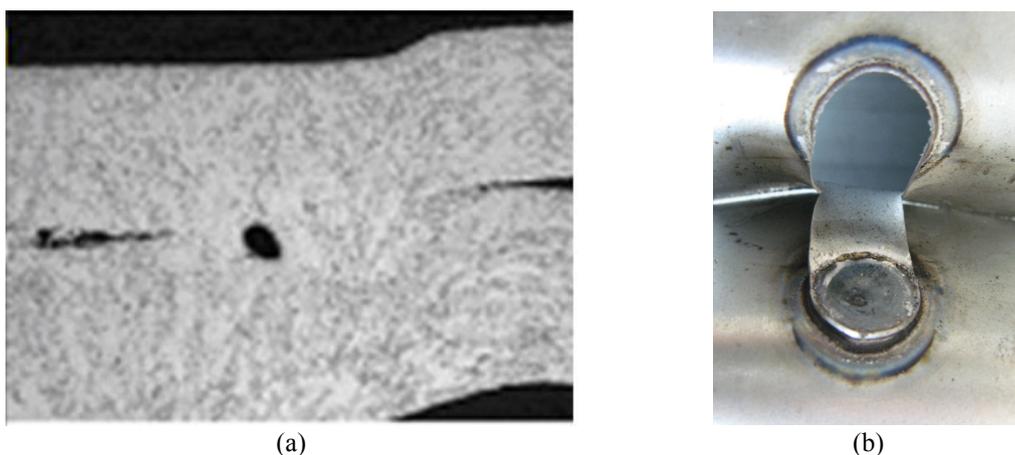


Fig. 10.8. Example of obtained weld by (a) cross-section of sheet (upper) and tube (lower) and (b) resulting plug failure upon separation. The photos are provided by Dortmund FuE-A Pressfügen/Kleben, ThyssenKrupp, Germany.

the degree of softening due to elevated temperatures. Large indentation can lead to torus shaped weld nuggets and cracks on the inside of the tube or in the sheet.

10.2.1 Numerical simulations of sheet-to-tube welding

Numerical simulations based on finite element modeling are suited for assisting experimental investigations for further understanding and improvement of the process. Rudolf [11] presented numerical simulations performed in SORPAS, which at that time was only available in 2D in terms of axisymmetric or block modeling. The work by Rudolf [11] concluded that the block model was not applicable for this geometry, while the axisymmetric model was able to provide useful information about the process. This is despite the fact that a sphere is actually simulated when modeling a cylindrical tube axisymmetric. Liang et al. [13,16] based their numerical analysis on ANSYS, but also with simplification to axisymmetric modeling.

With increasing insight into the process, the details that are sought for further understanding can only be simulated by 3D models that include the real geometry, facilitating the simulation of contact development, current density and heat development in three dimensions.

An example of single-sided sheet-to-tube spot welding is analyzed in the following by 3D numerical simulations. The example consists of a DP600 steel tube of inner diameter $\varnothing 52\text{mm}$ and outer diameter $\varnothing 55\text{mm}$, such that the wall thickness is 1.5mm . The sheet is 1mm thick and made of DX54 steel. The primary electrode is modeled as $\varnothing 16\text{mm}$ F1-type with tip diameter $\varnothing 5.5\text{mm}$ and the secondary electrode is modeled as being placed 40mm apart from the primary electrode center axis by a boundary condition defined ahead in terms of a tool.

A section of the tube is cut out for analysis as shown in Fig. 10.9a, where potential symmetry planes are also identified. The zx -plane is a pure symmetry plane, which is only violated if imperfections or misalignment is present. The yz -plane is more questionable because there is only a secondary electrode on one side (cf. Fig. 10.7). If the asymmetries presented by this are of interest, the yz -plane cannot be utilized as a symmetry plane. However, in the present analysis, this effect is ignored and the focus is put on the contact development and heat development arising from the original line contact between the sheet and the tube. Hence, the yz -plane is also utilized as a symmetry plane, and a finite element mesh as shown in Fig. 10.9b is established when utilizing both symmetry planes. A third plane could be utilized as a symmetry plane; namely the plane parallel to the sheet cutting the tube into half. Experience from the simulations showed that the deformation field violated this symmetry condition due to local deformation propagating through the tube circumference.

Fig. 10.9b shows three tools supplying boundary conditions besides the two aforementioned symmetry planes. The top and bottom tools act as the connection to the machine from the primary electrode and the support on the underside of the tube. The third tool is specially developed for simulation of the secondary tool in an efficient and simple way. It is specified as an electrical connection without mechanical influence to the simulation. This will simulate the good contact conditions resulting from the sec-

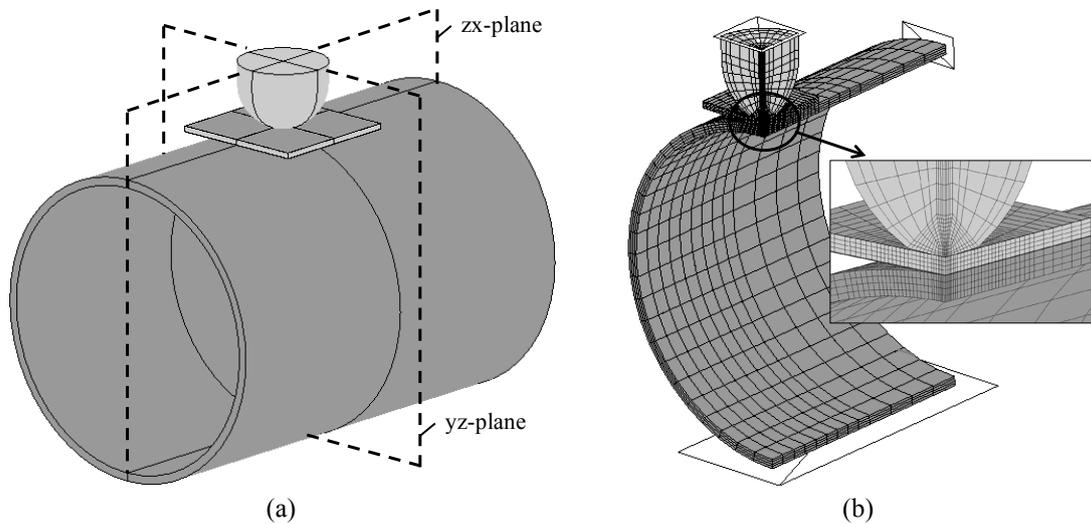


Fig. 10.9. Numerical model in terms of (a) identification of potential symmetry planes and (b) finite element mesh.

ondary electrode consisting of thin copper blades (see Fig. 10.7). The deflection of the tube that can be caused by the secondary electrode is minimal and ignored by this tool. The heat absorption by the secondary tool is also ignored. The magnification in Fig. 10.9b shows thin layers of elements on each side of the sheet to simulate the interface conditions between the objects.

The welding process is analyzed for ten different weld settings, spanned by two force levels and five current levels while the process time is kept constant. The force and current profiles are illustrated in Fig. 10.10. The force is raised to $1.5kN$ or $1.8kN$ before the current is applied. The current is applied with an up-slope of $60ms$ to a level of 4, 5, 6, 7 or $8kA$ and then kept constant for additionally $140ms$. The force is constant during the up-slope current and then decreased during $30ms$ to either $1.2kN$ or $1.5kN$. The force is hereafter constant until the end of the hold time.

The electrode force of the primary electrode is assumed to be controlled by a servo gun such that the two levels can be used during one weld as suggested by Sun and Wang [10], and a DC current profile is assumed with the possibility of using an up-slope. The higher level of the force in the early stage of welding is, together with the up-slope of the current, prescribed in order to ensure a mature contact area before reaching the full current level. The applied force is then lowered in the later stage of welding in order to decrease the indentation and local deformation due to softening and lack of support on the inside of the tube.

Simulations based on the ten different weld settings are presented in Fig. 10.11 by the process peak temperature in the end of the weld time, i.e. at process time $260ms$ in Fig. 10.10, where the current is switched off. Each subfigure in Fig. 10.11 shows the resulting peak temperature field in a view similar to that utilized in Fig. 10.9b, and they are organized such that the left column contains the welds performed at the low force level and the right column contains the welds performed at the high force level.

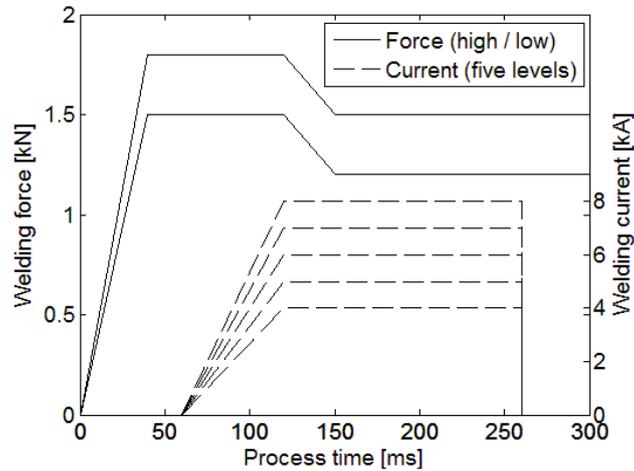


Fig. 10.10. Two selected electrode force levels and five selected current levels. All ten combinations are simulated.

Each row corresponds to one level of the current as specified in the left side of the figure.

At both force levels it is clear that the weld nugget increases in size with increasing current and eventually penetrates through the tube thickness. At the low force level, this happens already between 6 kA and 7 kA , while at increased force level it is postponed to between 7 kA and 8 kA . As it will be discussed ahead, it is also necessary to examine the subsequent cooling time for the evaluation of melting through to the inside of the tube because lack of an electrode on the inside of the tube results in poor cooling on the inner tube surface.

Fig. 10.11 shows that the weld nuggets resulting from a higher force level are smaller than the similar welds resulting from lower force level, which is to be expected due to larger and better contact obtained at the higher electrode force. Larger indentation and local deformation can be expected as a consequence of the high force, but this is partly compensated by less softening due to smaller heat generation. Increased indentation and local deformation is therefore not considered to be a problem when going from the low to the high force level in Fig. 10.10. However, this conclusion is closely related to a moderate increase in the force (20% in the early stage and 25% in the later stage).

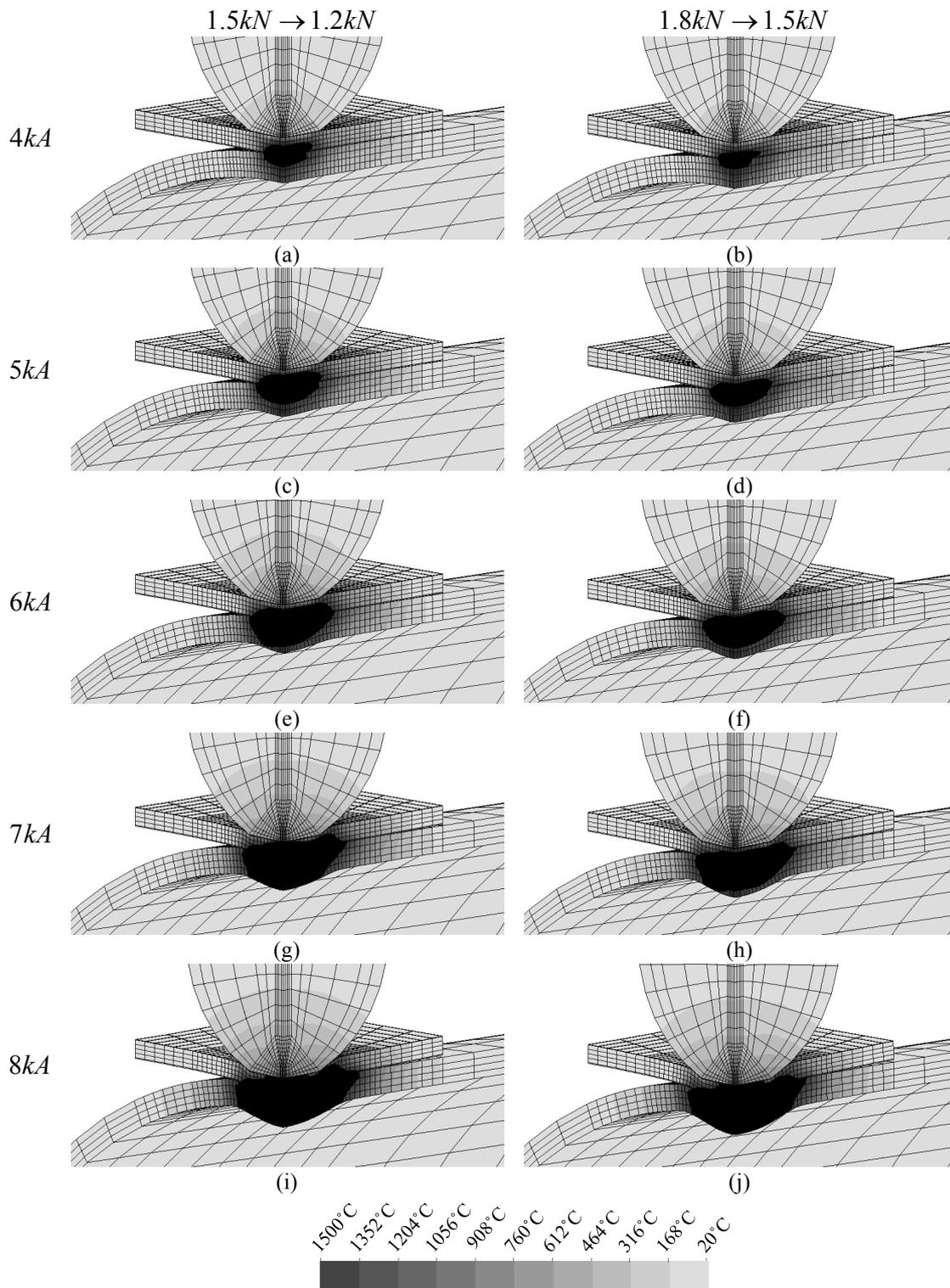


Fig. 10.11. Process peak temperatures in the end of the weld time for different weld settings. The force and current profiles follow Fig. 10.10. Force levels are indicated above the columns and current levels are indicated in the left side.

From Fig. 10.11 it appears that the higher force level in the present study results in better weld quality than the lower force level because of the increased contact area. Fig. 10.12 exemplifies this for the cases with $6kA$ weld current, which seems to present mature weld nuggets. Because the contact area in the early stages is smaller with the low electrode force (Fig. 10.12a), the heat generation is more intense and the nugget height and risk of penetration through the tube are larger than with the high electrode force level (Fig. 10.12b). The smaller contact area in the later stages combined with the larger weld nugget result in increased risk of splash between the sheet and the tube near the gap resulting from the curvature of the tube. The solid load bearing area that encloses the liquid in order to avoid splash is marked in both cases in Fig. 10.12. The comparison reveals that at low force (Fig. 10.12a), the nugget is larger while the contact area is smaller, and therefore this weld setting has higher risk of splash.

Spot welding the sheet to the tube under the high electrode force setting with a current level of $6kA$ seems from Fig. 10.11f and Fig. 10.12b to result in a well-shaped nugget of proper size while sufficient contact area is formed to avoid splash. It is nevertheless necessary to consider the cooling of the spot weld as unveiled above. The hold time ensures that a spot weld cools and solidifies to get enough strength before

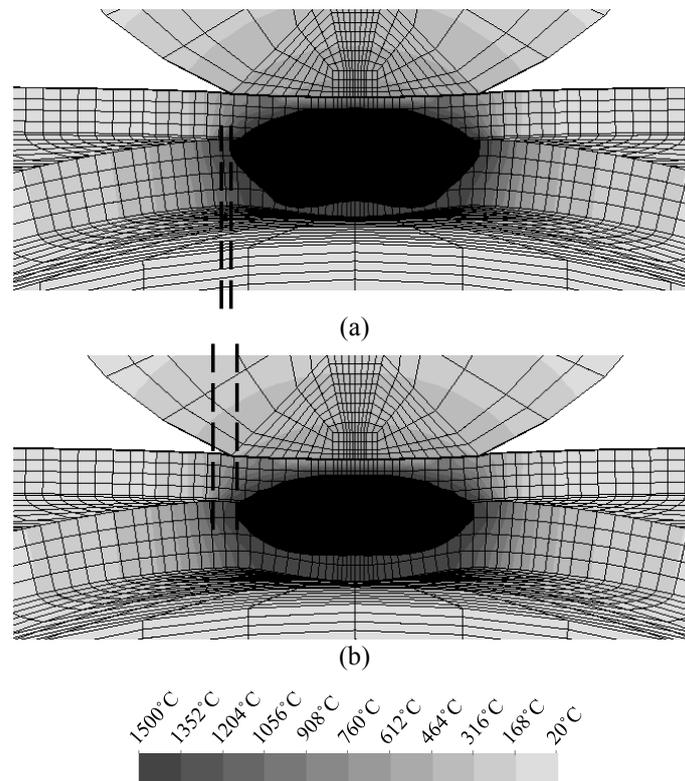


Fig. 10.12. Comparison of cross-sections (corresponding to the yz -plane in Fig. 10.9a) with process peak temperatures shown for the cases with $6kA$ current level at (a) low force level and at (b) high force level.

the electrode force is released and the electrode moved. Electrodes absorb the majority of the heat and lead it away due to high thermal conductivity. However, in the single-sided spot welding there is no electrode on the inside of the tube resulting in poor cooling on this side. The resulting cooling process during the hold time is depicted in Fig. 10.13 for the case with $6kA$ weld current and high electrode force. The sequence of depicted temperature fields starts from the end of the weld time and goes into the hold time in steps of $30ms$.

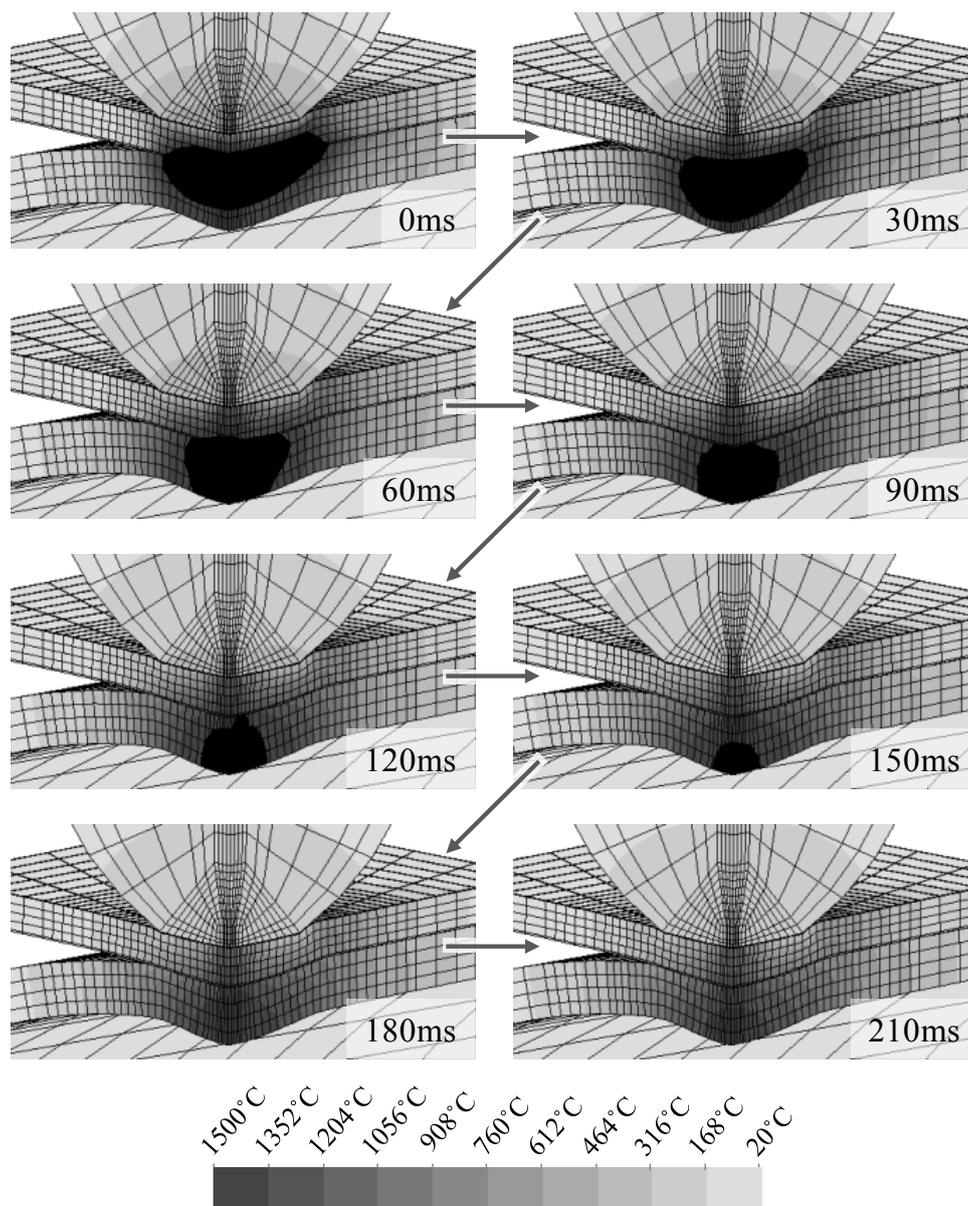


Fig. 10.13. Temperature during the hold time for the case with $6kA$ and high force. The hold time is indicated in each subfigure.

The sequence of temperature fields shows that the overall temperature naturally decreases, but it also shows that high temperature initially develops through the tube thickness as a result of heat conduction in the tube and poor cooling on the inner surface. The temperature field corresponding to 60ms into the hold time shows the penetration through the tube thickness to the inner surface. The two last temperature fields, corresponding to 180ms and 210ms into the hold time, are after complete solidification of the weld nugget. The solidification at the original interface position of the two materials takes place in-between the depicted time instants corresponding to 120ms and 150ms into the hold time.

The numerical study has been used to show the effect of selecting a proper force level. A sufficient force is required for establishing an initial contact area large enough to stand the applied welding current, but on the other hand a too large force would result in excessive indentation and local deformation in the tube. A certain amount of indentation is though needed because this is allowing the sheet and tube in the later welding stage to develop contact over an area large enough to contain the weld nugget and avoid splash. The numerical study was also used to highlight the importance of the cooling during the hold time. Since there is no electrode on the inside of the tube, the weld nugget has potential of penetrating through the tube thickness to the inside surface as illustrated in Fig. 10.13.

Besides future comparisons to corresponding experiments with specific welding parameters, it is of interest to simulate the effect of the position of the secondary electrode. The secondary electrode presents asymmetry which was neglected in the present study, but a future study could involve this effect.

10.3 Projection welding with unequal projection heights

Projection welding is an alternative to spot welding when joining sheets. The current is concentrated to the weld region by projections instead of being concentrated through the electrodes as in the above spot welding examples. An example of a type of application is shown in Fig. 10.14a, where projection welding is applied for joining the two flanges from each of the two parts. Such parts can be joined for creating a closed sectioned structure to enhance overall stiffness. In some cases, this is preferred prior to hydroformed parts as dealt with in Section 10.2. The reason for applying projection welding instead of spot welding in Fig. 10.14a can be the width of the flanges, because projection welding can be applied to narrower flanges than spot welding. Another type of application is illustrated by Fig. 10.14b, where two parallel sheet ends are to be joined. If one side needs to remain flat without electrode indents, projection welding can be employed as illustrated in the figure, where one sheet is preformed to fit on top of the other sheet.

Three projections on a row are considered in the following. Whether they are the only three or they are three out of a longer row is not considered important. The main interest is to simulate the effect of one projection being smaller than the others due to

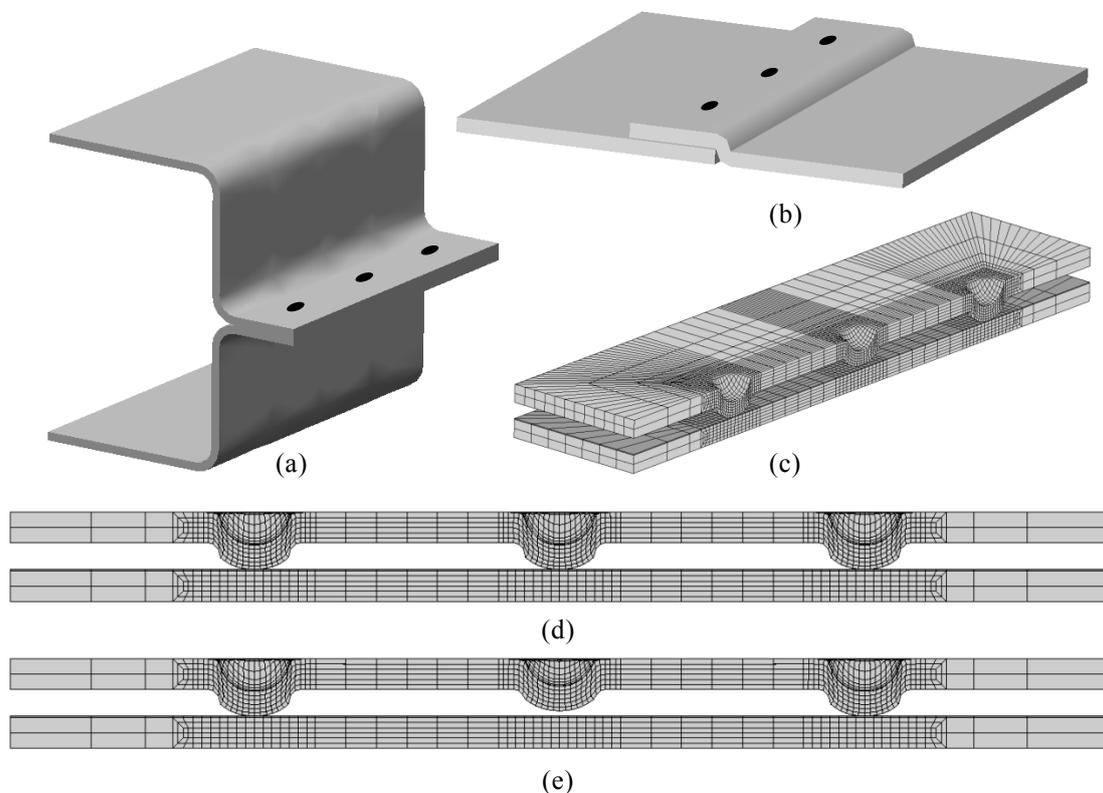


Fig. 10.14. Joining of sheets by projection welding. (a-b) Examples of application for joining two sheets by projection welding. (c) Finite element model of two sheets joined by three projections in a row, which ideally have equal heights (d), but may have unequal heights (e).

prior inaccurate embossing. For this purpose three projections are sufficient. The overall shape and setup is shown by the finite element mesh in Fig. 10.14c. This figure also shows that the following analysis is made without taking into account any specific configuration (e.g. Fig. 10.14a-b), or in other words that stiffness and material of the remaining structure are neglected. The ideal case with equal projection heights as shown in Fig. 10.14d and the case with unequal heights as shown in Fig. 10.14e are both simulated for comparison.

The simulated sheets are 1mm thick stainless steel AISI 304 sheets of dimensions $16 \times 36\text{mm}^2$. The projections are formed with 10mm between their centers. This is slightly closer than the recommended 12.7mm [5], but for evaluation of the effects, the distance is chosen short. The recommended shape of projections [5] is given in terms of the punch shape, the die shape and the overall dimensions of the projection. Hence, the embossing of the projections is simulated for prediction of the actual shape as depicted in Fig. 10.15a. After simulation of the embossing, the deformed mesh is saved as an object mesh that is used for further simulation of the projection welding. This is a generic procedure that can be utilized whenever it is relevant to simulate multiple process steps. The embossing in this example is simulated with rigid tools, thereby benefiting of the strong mechanical forming options available due to the kinship with I-

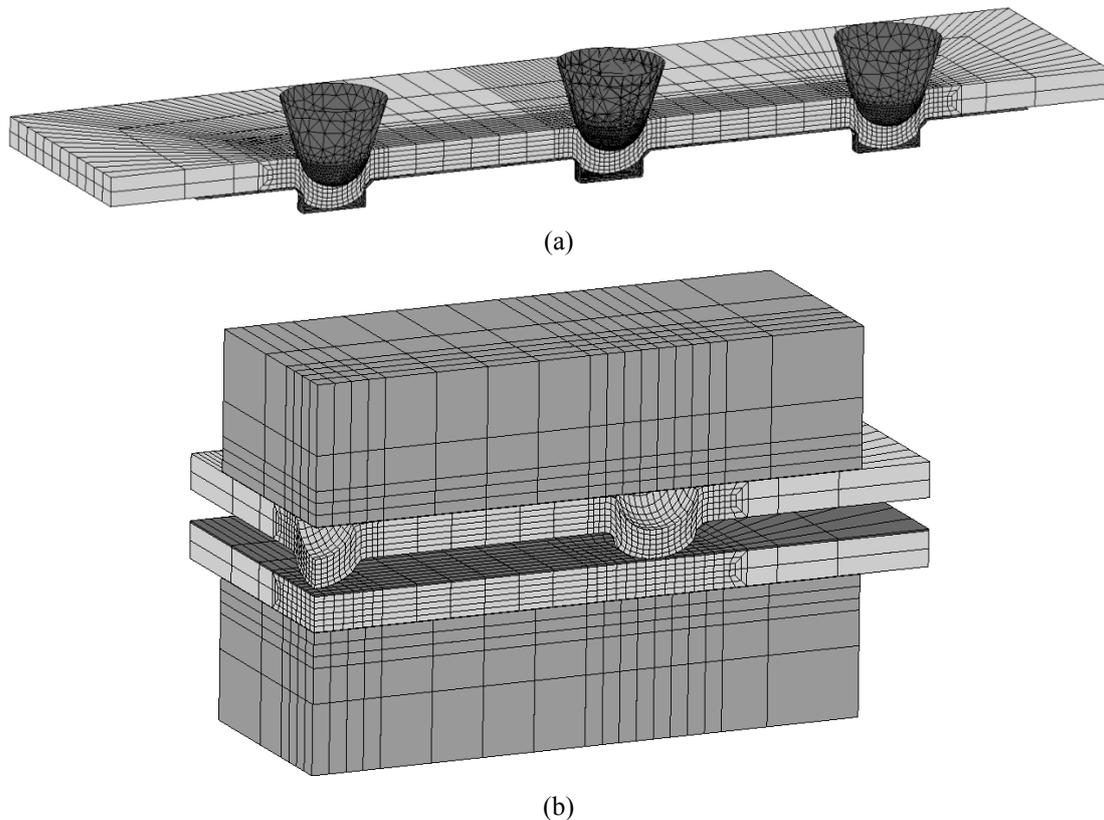


Fig. 10.15. Finite element mesh for simulation. (a) Simulated embossing of the three projections for prediction of the actual shape. (b) Finite element assembly for simulation of projection welding.

Form. Another benefit is the easy mesh generation, because an initial flat sheet is meshed (the same as used for the lower sheet in the welding simulation), and the simulated embossing provides the mesh of the actual projections.

After embossing, the projection diameter on the other side of the punch is 2.8mm and the projection height is 0.884mm , except for the one that is simulated smaller. The smaller projection has obtained a height of 0.708mm , corresponding to 80% of the intended height. The height of the projections is controlled by the closed bottoms of the discretized die in Fig. 10.15a. When the projections of full height reach the closed bottom, the embossing is stopped and the mesh is taken from the time step just before the projection tips experience compressive stresses. The closed bottoms of the die are only for this purpose. The height of the smaller projection is controlled by the relative position of the relevant punch.

The finite element model of 7749 hexahedral elements in Fig. 10.15b is used for simulation of the projection welding process. It consists of the two sheets, two flat copper alloy electrodes and interface layers in-between all objects. Two symmetry planes are utilized as evident from the finite element mesh resembling one quarter of the total geometry. The above embossing is also performed by simulating one quarter, although shown with all three projections visible for illustration.

The weld settings are based on the recommendations given in [5] for low-carbon steel as a starting point and then adjusted to match the simulated stainless steel. The welding time, $160ms$, is chosen within the recommended interval, the welding force, $2.4kN$, is adjusted towards the higher end of the range, and the current is selected considerably lower. The current is applied as a $50Hz$ AC current of $8kA$ RMS with an assumed conduction angle of 80% . After ended welding time, the electrode force is kept for additional $80ms$.

The resulting simulated projection welds are compared in Fig. 10.16 by the final shape and the process peak temperature with indication of weld nuggets. The simulation of equal projection heights (Fig. 10.16a) reveals similar weld nuggets in all three projections as expected, whereas the simulation of unequal heights (Fig. 10.16b) shows different nugget sizes. The nugget associated with the smaller projection is smaller than the others, but develops reasonable into the lower sheet, at least locally. The two neighboring nuggets are larger, but critical differences are found when compared to the ideal case; namely less development into the lower sheet across the interface and more

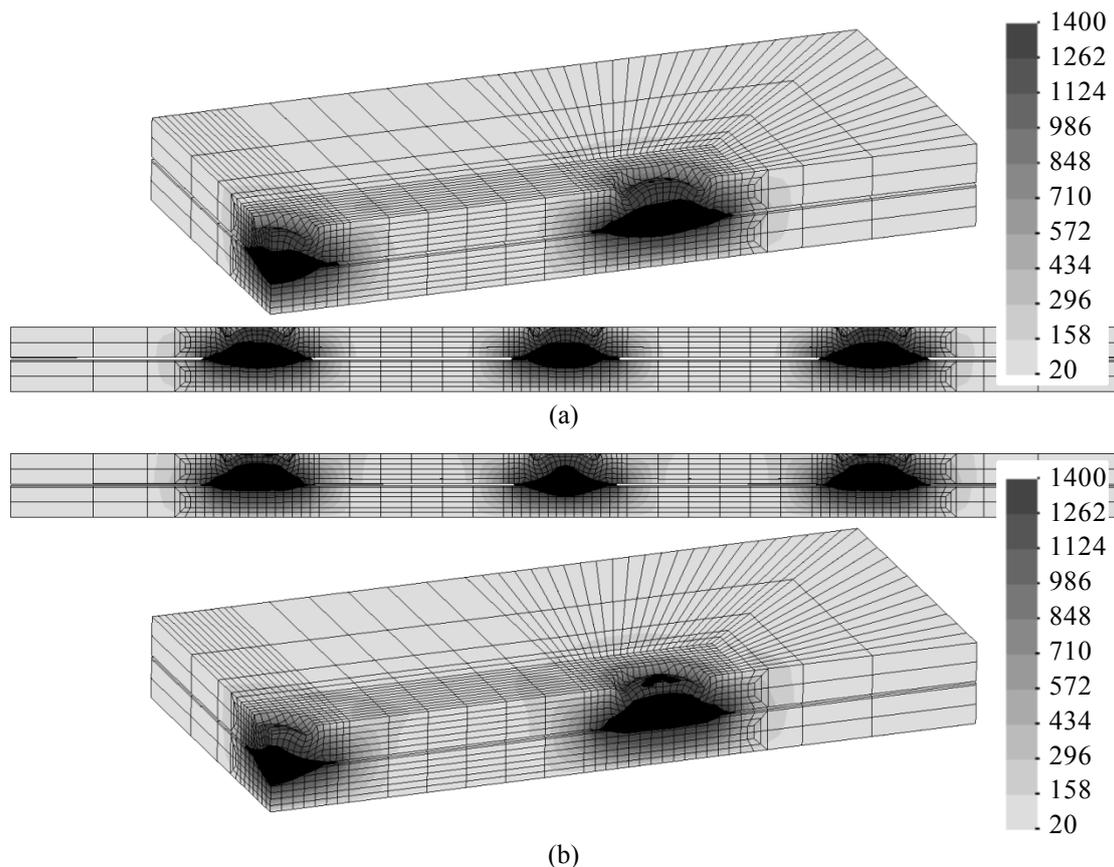


Fig. 10.16. Comparison of resulting projection welds by final geometry and process peak temperature with scale bars in degrees Celsius. The final welds are comparing welding with (a) initially equal projection heights and (b) initially unequal projection heights.

development into the material of the projection. The strength of the combined three projections with unequal heights are therefore expected much lower than when welded under ideal conditions because the center weld is smaller and the two neighboring welds develop less across the interface. For practical purposes, the current could be chosen at higher level to accomplish full collapse of the projections and larger nugget development.

Evaluation by Fig. 10.17 of the heat development during the first cycle of the applied current shows the causes for the different weld nugget developments. The initial contact between the projections and the sheet is shown just before the onset of welding current by the upper figures in Fig. 10.17a for equal projection heights and in Fig. 10.17b for unequal projection heights. The smaller projection in Fig. 10.17b is just touching the lower sheet with resulting small contact area. As a consequence, the neighboring projections are deformed more to carry the applied load by the electrode force, and hence the contact area is larger than in the ideal case with equal projections. The following subfigures show the heat development in intervals of $5ms$, ending at the end of the first cycle at $20ms$ (equal to one period of $50Hz$ AC current).

After $5ms$, severe melting of the small projection is observed, while the neighboring projections sustain the electrode force by the larger contact area. The neighboring projections do not experience melting at this stage, whereas melting is initiated in all contact interfaces of the projections in the ideal case. The severe melting of the smaller projection may cause expulsion and further uncontrolled conditions already at this stage.

Further melting of the small projection happens before the neighboring projections experience melting. The situation $15ms$ into the welding time shows that melting of the neighboring projections initiate on the outer ring of contact because of higher current density than in the central region, where good contact conditions are developed. The final situation showed in the figure is in the end of the first cycle. The melting of the neighboring projections has developed over the entire area, but only limited across the interface. The projections with equal heights, on the other side, develop regularly from the centers by nugget growths that also cover the interfaces to a larger degree. Further process simulation results in the final configurations shown in Fig. 10.16.

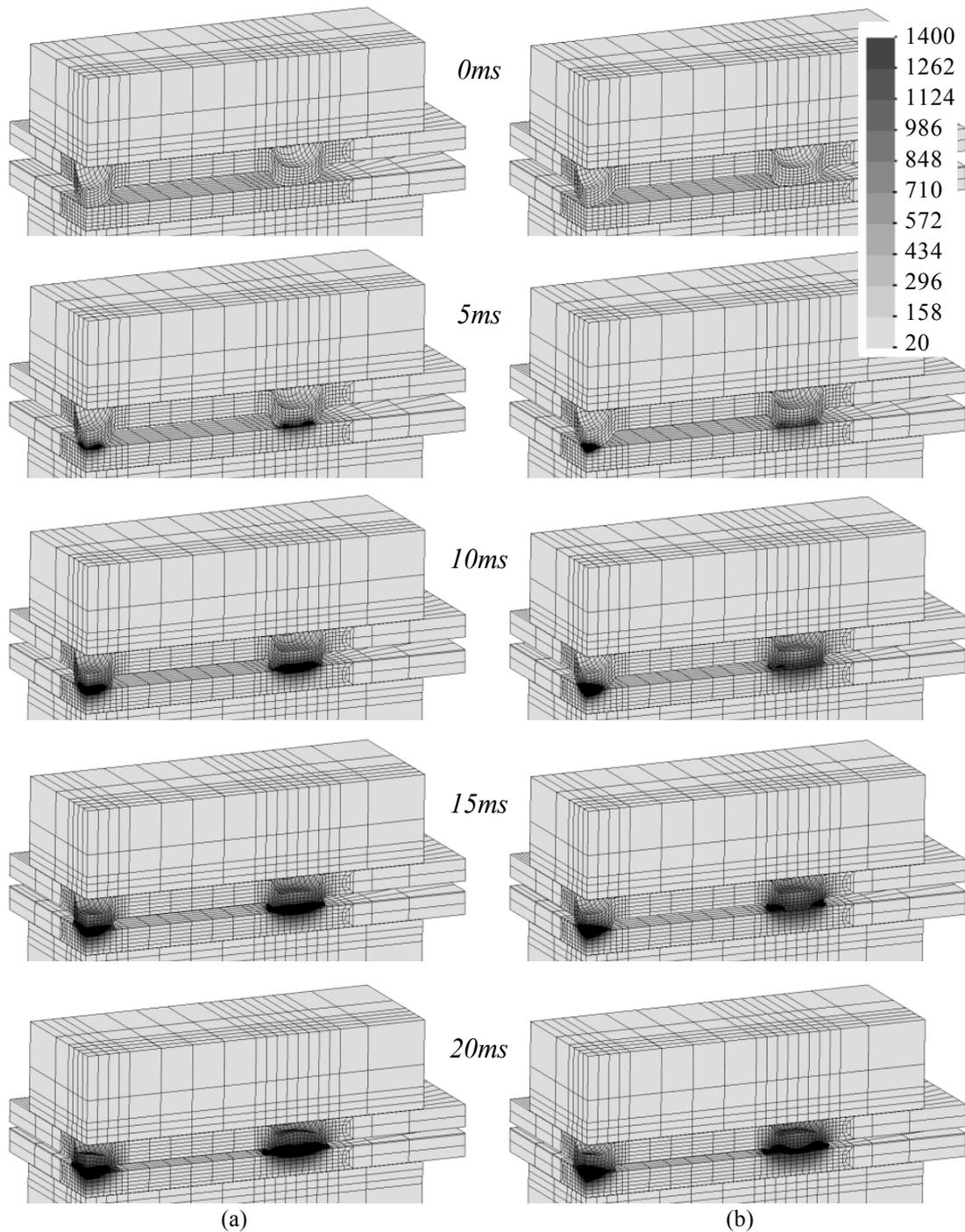


Fig. 10.17. Comparison of the initial welding stage with (a) initially equal projection heights and (b) initially unequal projection heights. The indicated process time is relative to the welding time, such that 0ms corresponds to the onset of the welding current after the squeeze time, and 20ms corresponds to completion of the first period of the 50Hz AC current.

10.4 Projection welding by longitudinal embossment

Another application of projection welding is presented in Fig. 10.18, where two sheets are joined perpendicular to each other. This is relevant to e.g. fabrication of housings and containers that are not required to be water or air tight and to the addition of perpendicular stiffeners to sheet panels.

The presented example is an industrial case provided by a Japanese company. In order to facilitate joining of two sheets perpendicular to each other by projection welding, one of the sheets is embossed as shown in Fig. 10.18a. When the other sheet is positioned as shown in Fig. 10.18b, the longitudinal embossments ensure local contacts between the two sheets. Resistance projection welding is carried out under constant weld force and DC current resulting in the joint shown in Fig. 10.18c. A close up of one of the projection welds is shown in Fig. 10.18d and a cross-section is shown in Fig. 10.18e.

The two sheets are 0.8mm thick high strength low alloy steel sheets (grade similar to HSLA340). The welding parameters are as follows: 700N weld force, 3.5kA DC weld current and 30ms weld time.

Fig. 10.19a shows the finite element discretization of one of the projection welds by 5630 hexahedral elements. The simulation utilizes a natural symmetry plane along the longitudinal projection (that is the cutting plane utilized to show the cross-section in Fig. 10.18e). An additional symmetry plane is assumed in the simulation to reduce the model size. It is introduced in the center of the vertical sheet in Fig. 10.18e, such that the final model utilizing both symmetry planes is as shown in Fig. 10.19a. The round

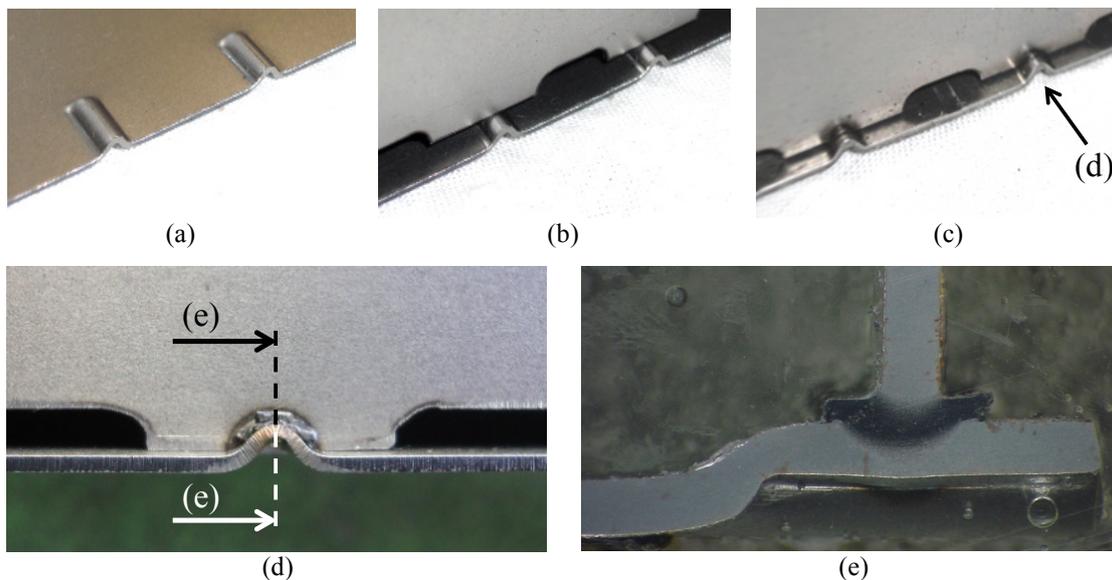


Fig. 10.18. Industrial example of projection welding of two sheets perpendicular to each other. (a) Sheet with embossed longitudinal projections. (b) Positioned perpendicular sheets before welding and (c) after welding. (d) Side view after welding (view indicated in (c)). (e) Cross-section as defined in (d).

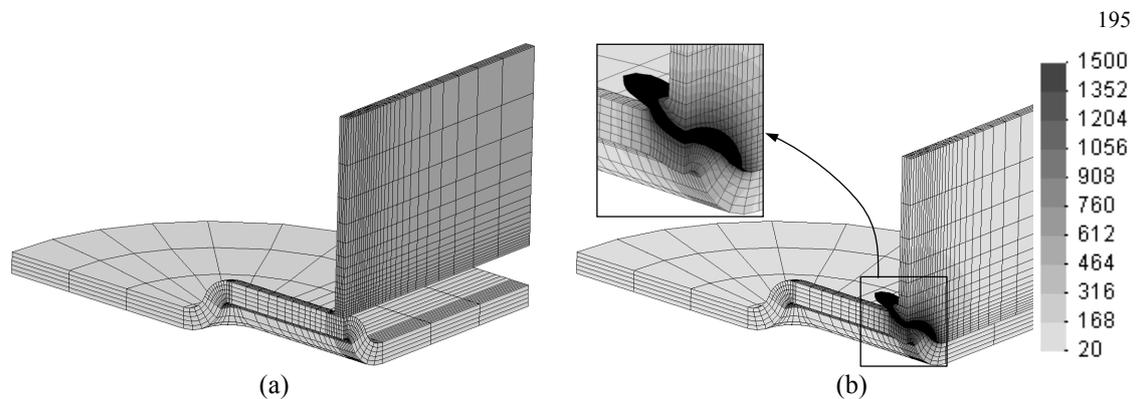


Fig. 10.19. Projection welding of two perpendicular sheets. (a) Initial finite element mesh and (b) predicted peak temperature field (degrees Celsius) at the end of welding showing molten volume squeezed out between the two sheets.

periphery of the lower sheet is only to make structured meshing of the round end of the projection easier. It does not influence the simulation due to the distance from the weld.

The second symmetry plane is justified as follows with reference to Fig. 10.18e. The differences on each side of the vertical sheet in terms of the electrical and thermal fields are considered negligible. This is in light of the short process time (weld time is $30ms$) and considering the distance to the end of the embossment on one side and to the end of the bottom sheet on the other side. As regards the mechanical aspects of the assumed symmetry plane, the geometry after welding (Fig. 10.18e) is symmetric around the vertical sheet. The longitudinal embossment does not bend towards the free end, and the free end can therefore be omitted from the simulation. The side including the rounded end of the embossment is included to prevent the embossment from flattening, and the mirroring of that does not affect the overall deformation.

The simulated weld is shown in Fig. 10.19b with the peak temperature distribution shown. In the interface of the two sheets, the material melts and squeezes out as in the real case (compare detail in Fig. 10.19b to the cross-section in Fig. 10.18e) while the upper sheet closes towards the bottom sheet (compare Fig. 10.19b to the side-view in Fig. 10.18d).

A detailed comparison of the real example (Fig. 10.18) and the simulated projection weld (Fig. 10.19) is presented in Fig. 10.20 in the cross-section similar to Fig. 10.18e. The comparison covers the final geometry as well as the peak temperature field. As regards the geometry, the main difference is the shape of the metal that is squeezed out between the two sheets in a molten or mushy state. The exact shape might be of less importance compared to the volume squeezed out and the formed contact area during welding as it relates to the heat development. In the specific example, more elements would be required in the volume that is squeezed out if the details of the squeeze out are of importance.

The heat development and the heat balance were of more importance when doing the presented simulation in collaboration with the company. The simulated process peak temperature field and the resulting microstructure of the real case are compared.

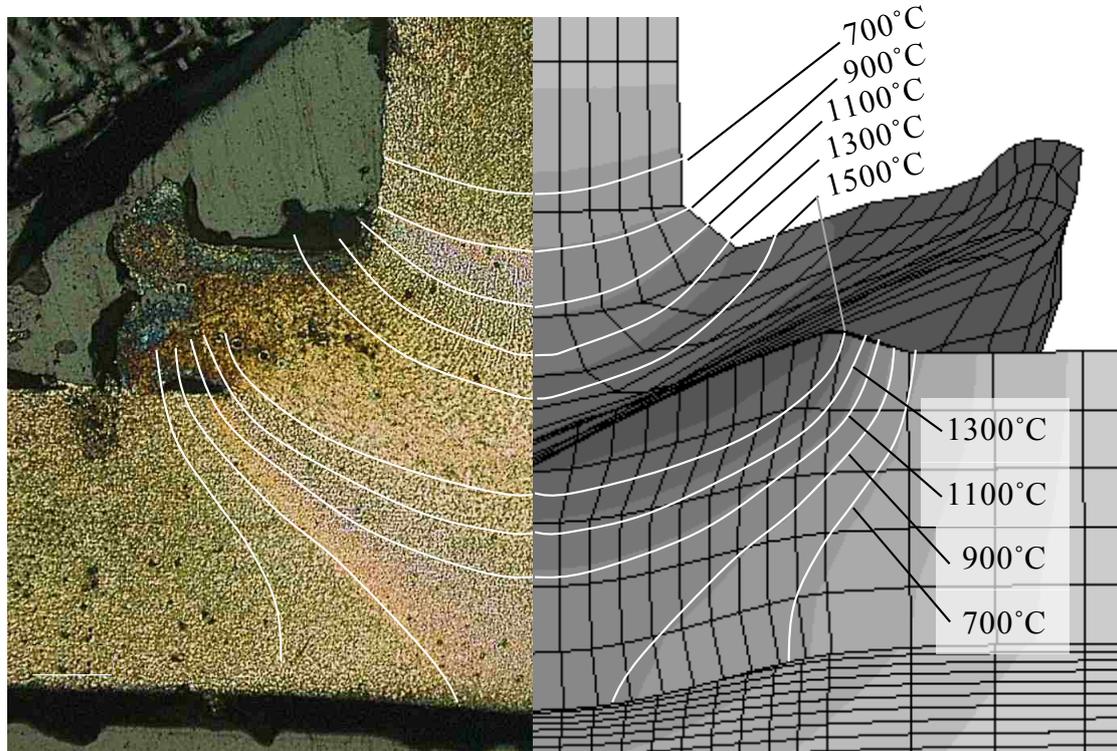


Fig. 10.20. Comparison of cross-section of the real component and simulated peak temperature distribution in the cross-section view similar to Fig. 10.18e. The simulated peak temperature field is shown on a 20-2000°C scale with selected isothermal lines. These lines are mirrored onto the actual cross-section.

The selected isothermal lines in the simulated temperature field are mirrored onto the cross-section of the real case, revealing that the temperature gradients are simulated correctly as the isothermal lines of the simulated temperature field match the shape of the border lines between the different microstructures.

10.5 Welding of bellow to disc by natural projection

This section presents an industrial resistance welding case from a Danish company. The specific welding case is part of the production of thermostat valves for radiators. Inside the thermostat valve is a bellow that expands or contracts due to temperature changes and thereby opens or closes the valve controlling the heating of the radiator. A few steps of this production are illustrated in Fig. 10.21.

A tin-bronze bellow tube with a conical collar (2) is resistance welded to a steel ring (3) between electrodes (1) and (4) as schematically shown in Fig. 10.21a by its setup. The result of this welding process is the joined bellow tube and steel ring shown in Fig. 10.21b (upside down compared to Fig. 10.21a). The bellow is hereafter formed as shown in Fig. 10.21c before it is mounted in a container as depicted in Fig. 10.21d (turned back to the same orientation as Fig. 10.21a). The joint between the steel ring

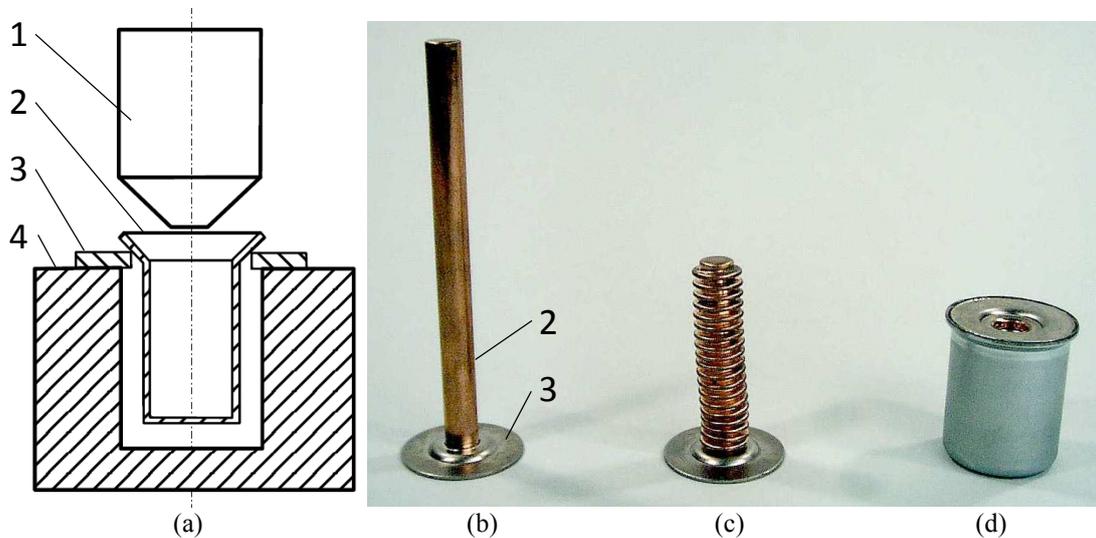


Fig. 10.21. Selected process steps utilized in the production of thermostat valves. (a) Resistance projection welding of bellow tube to steel ring. (b) Joined bellow tube and steel ring. (c) Formed bellow. (d) Mounting in container by resistance projection welding of steel ring to container.

and the container is also accomplished by resistance projection welding. However, this section focuses on the resistance projection welding of the bellow tube to the steel ring (Fig. 10.21a-b).

The $\varnothing 8\text{mm}$ bellow tube is produced in tin-bronze CuSn6 (W.Nr. 2.1020) with a wall thickness of 0.14mm . A 90° conical collar is formed prior to welding such that the contact to the steel ring forms a natural projection. The 1mm thick mild steel (W.Nr. 1.0338) ring has outer diameter $\varnothing 29\text{mm}$ and hole diameter $\varnothing 8.3\text{mm}$ and it is coated with a $2 - 6\mu\text{m}$ thick layer of electroless deposited Ni-P alloy (8-12% P) to facilitate welding. The upper 90° conical electrode is a standard copper alloy for resistance welding, CuCr1Zr, A2/2 after ISO 5182:1991.

This welding case was analyzed by Rasmussen [17] and Bay et al. [18] with focus on electrode wear and the influences on the weld quality. The joint is tested for leakage in the production by an applied pressure. Very few (of the order of per thousand) defects are observed when welding up to 40,000 pieces, but the defect rate increases with electrode wear. The influence of electrode wear is therefore analyzed and presented in the following by new finite element simulations. By assuming axisymmetry, these simulations are performed in 2D.

The electrode geometry changes significantly due to electrode wear as illustrated by a new and a worn electrode in Fig. 10.22a-b. A cross-section of a worn electrode after 580,000 welds is shown in Fig. 10.22c showing severe change in electrode geometry from the original conical shape. This number of welds is well beyond the normal tool life and is made for the analysis such that clear effects are noticed.

The differences in the resulting welds are analyzed by a combined metallographic study of selected cross-sections and a numerical study, which is based on the axisymmetric finite element models shown in Fig. 10.23. The setup including a new electrode

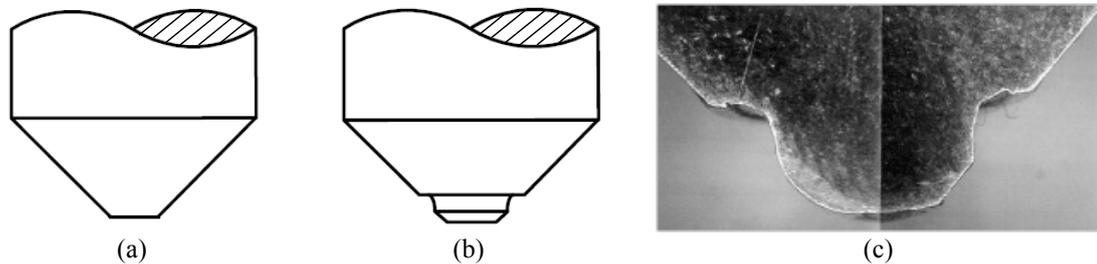


Fig. 10.22. Illustrations of electrode wear by (a) new electrode and (b) model of a worn electrode. (c) Cross-section of a real worn electrode after 580,000 welds, which is well beyond the normal tool life.

is shown in Fig. 10.23a with a close-up of the simulated deformation and temperature in the end of the weld time in Fig. 10.23b. The model including a worn electrode is shown in Fig. 10.23c with the shape of the electrode equal to the worn electrode shown by its cross-section in Fig. 10.22c. Fig. 10.23d shows the simulated deformation and final temperature as a result of welding with the worn electrode.

The simulated temperature fields of Fig. 10.23b and Fig. 10.23d are compared with the cross-sections of the real welds in Fig. 10.24 contributing to the overall analysis. Deformation and microstructure show clear differences between the welds stemming from a new electrode and the worn electrode.

After welding with a new electrode, an investigation of the microstructure (Fig. 10.24a) shows that the steel adjacent to the weld interface has a very coarse, ferritic grain structure, which is observed as a thin, bright zone. Apparently, the steel in this zone has been heated close to, but not above, 900°C leading to grain growth in the fer-

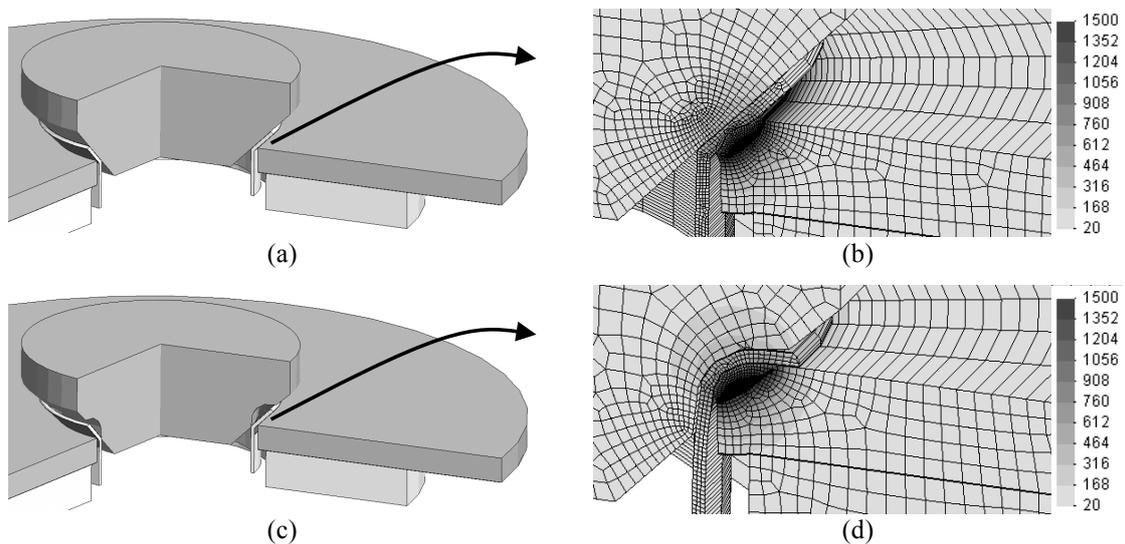


Fig. 10.23. Simulated projection welding of bellow tube to steel ring with new and worn electrode. (a) Model including new electrode with detail shown in (b) including the simulated deformation and final temperature field. (c) Model including worn electrode with detail shown in (d) including simulated deformation and final temperature field.

rite. This is confirmed by the corresponding numerical simulation (Fig. 10.24b), where the white isothermal line corresponds to 900°C process peak temperature. This isothermal line is seen not to cross the interface confirming the above hypothesis.

Below this zone, the microstructure appears to be dark indicating that the peak temperature in this part of the steel has been raised to above 900°C causing a phase trans-

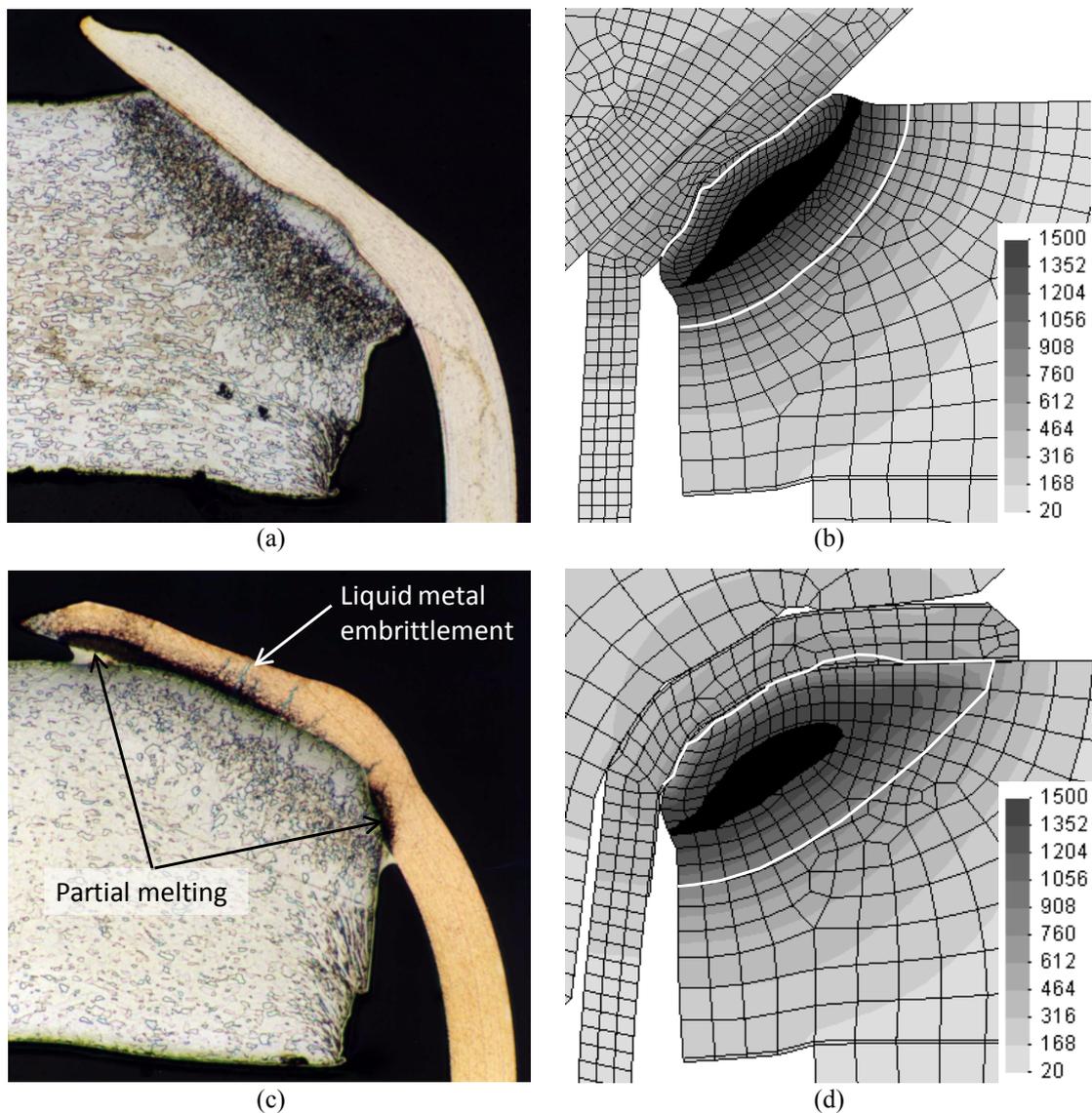


Fig. 10.24. Cross-sections of welded bellow tube to steel ring. (a) Weld performed with a new electrode. (b) Simulation with new electrode. (c) Weld performed by worn electrode and indication of partial melting and liquid metal embrittlement in the bellow tube. (d) Simulation with worn electrode. Figures (b) and (d) show the final temperature field together with contour lines corresponding to the 900°C isothermal line of the process peak temperature field.

formation to fine-grained austenite. During the subsequent rapid cooling this austenite transformed to very fine-grained ferrite, which appears dark on the micrograph. The microstructure of Fig. 10.24a indicates that the highest temperature during the welding process was reached inside the steel at a certain distance from the weld interface, and not at the interface itself. This is due to the large difference in electrical resistivity of the tin-bronze bellow tube and the steel ring and the cooling through the bellow tube and electrode with high thermal conductivity. This effect is also seen in the simulation (Fig. 10.24b). No phase transformations are observed in the tin-bronze bellow tube, which is thus kept in a good condition.

Welding with the heavily worn electrode results in a microstructure (Fig. 10.24c), which as regards the steel contains the same microstructural elements as those seen in Fig. 10.24a. Due to the poor contact between the upper electrode and the tin-bronze tube in the first phase of welding, the tube experiences higher temperatures than with a new electrode. The dark areas in the tin-bronze represent areas of partial melting and hot cracking. This occurs for the bellow tube material when the temperature is above approximately 900°C , cf. the Cu-Sn phase diagram provided in Fig. 10.25 [19], where the actual tin-bronze alloy (CuSn6) is marked by the dashed line.

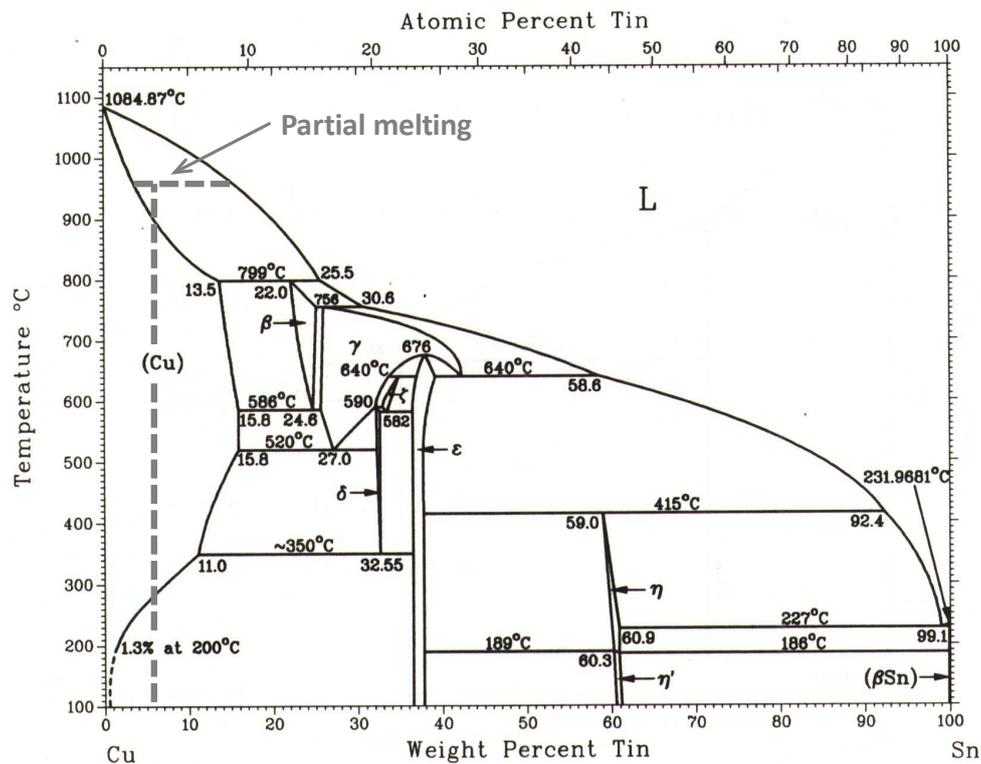


Fig. 10.25. Phase diagram of copper (Cu) and tin (Sn) with the actual tin-bronze alloy marked by the dashed line at 6 weight-% tin. The specific alloy experiences partial melting in the region above approximately 900°C .

The simulated weld by the worn electrode (Fig. 10.24d) also reveals peak temperatures above 900°C in the tin-bronze by the white isothermal line. The real weld is seen to have experienced heavier partial melting than the simulation shows. This can stem from asymmetric wear of the electrode, which will result in further localization of the heat along the circumference, whereas the axisymmetric simulation distributes the heat evenly along the circumference.

Besides the partial melting, liquid metal embrittlement is noticed in Fig. 10.24c. This is caused by penetration of melted Ni-P coating into the grain boundaries of the tin-bronze. These phase transformations are explained by the elevated temperatures reached when welding with a worn electrode.

10.6 Micro joining of fork and wire

An industrial case from the electronics industry is provided by means of a collaborative work with a German company. The application is micro joining of a fork to a wire as shown in Fig. 10.26 in its configuration before joining. The wire (1) is pure copper of diameter $\varnothing 0.73\text{mm}$ coated by a $5\mu\text{m}$ thick polyimide plastic (2). It is joined to an alloyed copper fork (3) between two tungsten electrodes (4 and 5). The tungsten electrodes close the fork legs around the wire by an applied force to form the joint. A current is simultaneously applied for two reasons. The resistance heating caused by the current facilitates the closing of the fork around the wire due to softening of the material. At the same time, the induced temperature melts the polymer coating locally on the wire to create electrical connection between the wire and the fork, which is required for the use of the component while the polymer keeps the remaining wire isolated.

Although no weld is created (there is no melting except for the coating), the joining is facilitated by the principles of resistance welding, and the multi-object numerical

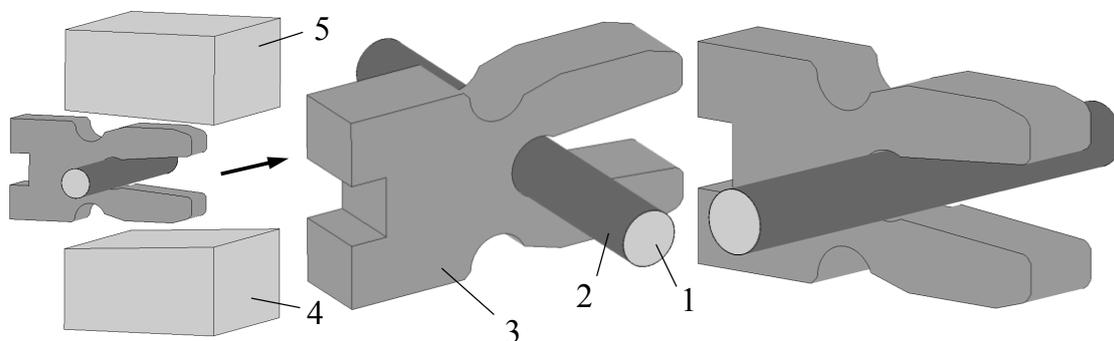


Fig. 10.26. Initial configuration of micro joining process of copper wire (1) coated by polyimide plastic (2) to a copper alloy fork (3) between two tungsten electrodes (4 and 5).

simulation based on the electro-thermo-mechanical coupled finite element flow formulation appears effective for performing the analysis of the process. Two natural symmetry planes are utilized to reduce the finite element model, such that the finite element mesh in Fig. 10.27a consisting of 4856 elements represents the overall geometry. The process conditions shown in Fig. 10.27b are applied in the simulation.

The applied force is built up to a level of $120N$ and kept constant until and during the first current pulse. The first current pulse has an up-slope time of $80ms$ reaching $0.75kA$ DC, which is kept constant for additionally $80ms$. The force is raised to $150N$ before the second current pulse, which is applied as a constant current of $1.2kA$ DC during $50ms$.

The effect of the two applied pulses is shown in Fig. 10.28 by the simulated process. Fig. 10.28a shows the moment where the tungsten electrodes just touch the legs of the fork. This corresponds to time $0ms$ in Fig. 10.27b where the force is applied. After $80ms$, the applied force has been kept constant for $30ms$ and the first current pulse is ready to be applied. At this stage (Fig. 10.28b) the deformation of the fork is enough to close the initial gap towards the wire such that a sound contact is setup before applying the current.

At the end of the first current pulse, the tips of the fork legs are closed (Fig. 10.28c). This deformation happens under the same applied force due to softening of the material. The temperature field after the first current pulse is shown in Fig. 10.28c with a maximum reached temperature $334^{\circ}C$. In order to perform the final closing of the fork, the second pulse is applied while at the same time increasing the applied load. This results in the final geometry shown in Fig. 10.28d. The figure also shows the temperature field with a maximum reached temperature $507^{\circ}C$, which is sufficient to

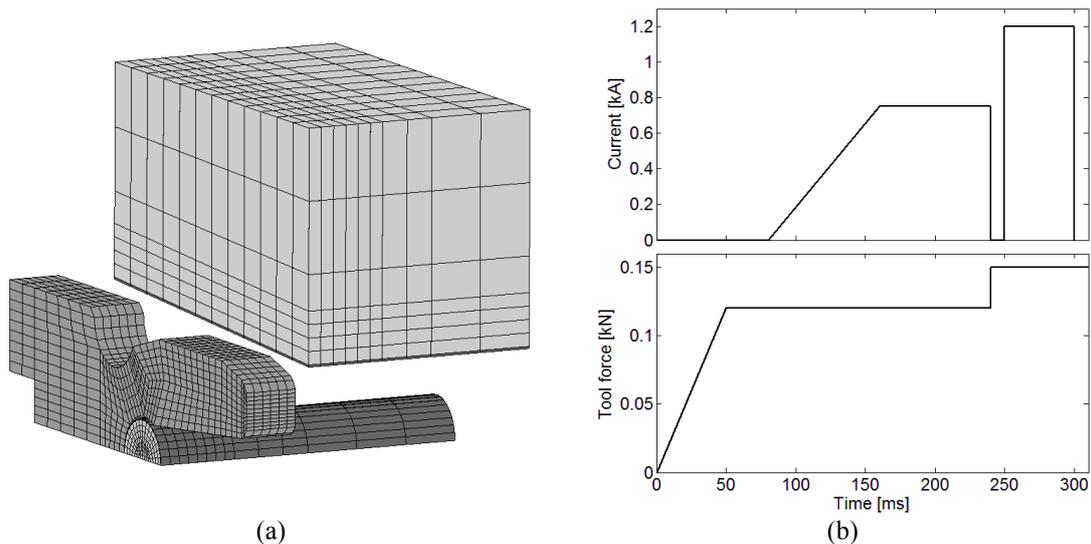


Fig. 10.27. Simulation of micro joining of fork to wire. (a) Initial mesh by utilization of two natural symmetry planes. (b) Applied current and force as function of process time.

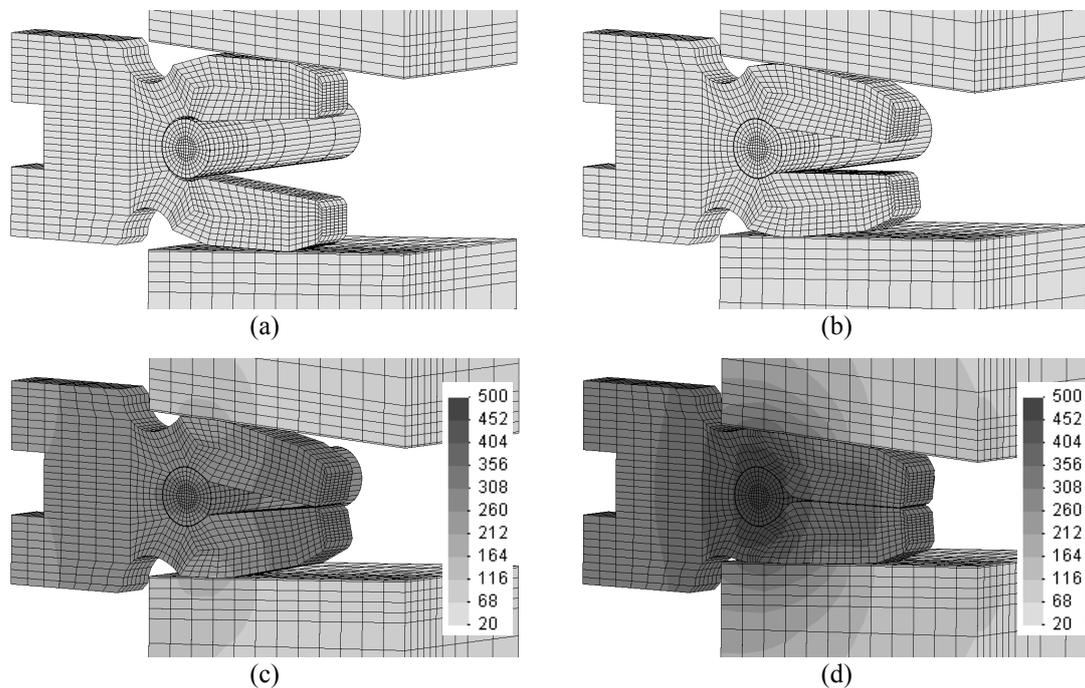


Fig. 10.28. Finite element predicted temperature in micro joining of fork to wire. (a) Start of the joining process at time $0ms$, where the electrodes just touch the fork. (b) Joining process after $80ms$ corresponding to the onset of the first current pulse. (c) End of first current pulse at time $250ms$. Maximum temperature reached at this stage is $334^{\circ}C$. (d) Completion of the joining process including the two current pulses at time $300ms$. Maximum temperature reached is $507^{\circ}C$.

melt the polymer coating to create electrical contact between the fork and the wire of importance to the final component.

The final geometry is compared to the real component in Fig. 10.29. The left figures compare the overall deformation showing that both the simulation and the real joint result in closing of the fork to a degree where the fork legs touch each other along the majority of their length. A detailed view of the region near the wire is shown in the right figures, where it is seen that the fork is closed around the wire with almost no deformation of the wire, which has part of its stiffness from the wire outside the contact area to the fork. The right figures also show that the amount of closing of the fork is simulated correctly near the wire.

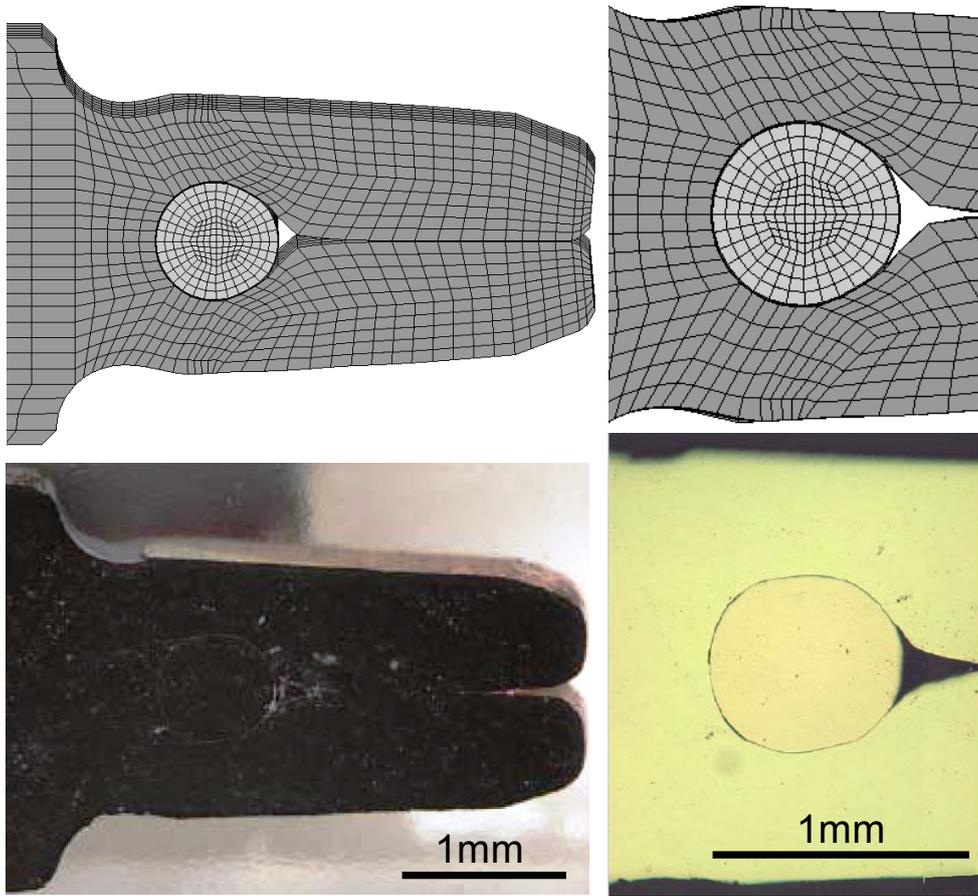


Fig. 10.29. Comparison of simulation (upper) and the real component (lower) in terms of the final geometry of the joined wire and fork.

References

- [1] Zhu W-F, Lin Z, Lai X-M, Luo A-H (2006) Numerical analysis of projection welding on auto-body sheet metal using a coupled finite element method. *International Journal of Advanced Manufacturing Technology* 28(1-2):45-52.
- [2] Nielsen CV, Friis KS, Zhang W, Bay N (2011) Three-sheet spot welding of advanced high-strength steels. *Welding Journal* 90(2S):32s-40s.
- [3] Blondeau R, Maynier P, Dollet J (1973) Prediction of the hardness and strength of plain and low-alloy steels from their structure and composition. *Memoires Scientifiques de la Revue de Metallurgie* 70(12):883-892.
- [4] Pedersen KR, Harthøj A, Friis KL, Bay N, Somers MAJ, Zhang W (2008) Microstructure and hardness distribution of resistance welded advanced high strength steels. *Proceedings of the 5th International Seminar on Advances in Resistance Welding*, Toronto, Canada, pp. 134-146.
- [5] Resistance Welding Manufacturers' Association (2003) *Resistance welding manual*. RWMA, Philadelphia.

- [6] Nielsen CV, Chergui A, Zhang W (2012) Single-sided sheet-to-tube spot welding investigated by 3D numerical simulations. Proceedings of the 7th International Seminar on Advances in Resistance Welding, Busan, Korea, pp. 147-158.
- [7] Shah S, Bruggemann C (1994) Tube hydroforming – process capability and production applications, Proceedings of International Body Engineering Conference IBEC '94, Detroit, Michigan, pp. 26-30.
- [8] Poss MG, Lendway IV JM (1997) Closed tubular automobile parts demand innovative welding methods, *Welding Journal* 76(10):55-58.
- [9] Cho Y, Chang I, Lee H (1997) Single-sided resistance spot welding for auto body assembly. *Welding Journal* 85(8):26-29.
- [10] Sun PC and Wang PC (2006) Sheet-to-tube resistance spot welding using servo gun. US patent no. 7,060,929 B2.
- [11] Rudolf H (2007) Widerstandspunktschweissen von Blech-Hohlprofil-Verbindungen an verzinkten Stahlblechen (in German), Dr.-Ing. Dissertation, TUDpress, Dresden.
- [12] Yang HG (2010) Investigations on sheet to cylindrical tube single-sided spot welding using servo gun. 2010 International Conference on Computer, Mechatronics, Control and Electronic Engineering (CMCE), pp. 256-258.
- [13] Liang C, Sun F, Liu X (2011) Effect of weld parameters on weld quality in sheet to tube joining with simulation and experimental research. 2011 International Conference on Consumer Electronics, Communications and Networks (CECNet), pp. 302-305.
- [14] Yang HG, Hu SJ, Zhang YS, Li YB, Lai XM (2007) Experimental study of single sided sheet to tube resistance spot welding. *Science and Technology of Welding and Joining* 12(6):530-535.
- [15] Li YB, Liang CP, Zhang YS, Lin ZQ (2008) Application of electrode force change in single sided resistance spot welding using servo gun. *Science and Technology of Welding and Joining* 13(7):671-678.
- [16] Liang C-P, Lin Z-Q, Chen G-L, Li Y-B (2006) Numerical analysis of single sided spot welding process used in sheet to tube joining. *Science and Technology of Welding and Joining* 11(5):609-617.
- [17] Rasmussen MH (2000) Kvalitetssikring af pressvejseprocessen (in Danish). PhD Thesis, Technical University of Denmark.
- [18] Bay N, Zhang W, Rasmussen MH, Thorsen KA (2003) Resistance welding: Numerical modelling of thermomechanical and metallurgical conditions. DMS Winter Annual Meeting, Danish Metallurgical Society, pp. 87-100.
- [19] ASM International (1992) Alloy phase diagrams. ASM Handbook 3:2167-2182.

11. Conclusions and Future Work

The presented work covers complete computer simulation of resistance welding. The process involves mechanical, thermal and electrical aspects of physics together with dynamically changing contact between multiple objects with changing material properties. The simulations are facilitated by three-dimensional electro-thermo-mechanical coupled finite element modeling. The core of the finite element program is the flow formulation for handling the mechanics. Resistance welding and particularly projection welding involve large plastic deformations, and the elastic effects can hence be neglected in the majority of cases. The flow formulation, which is based on rigid-plastic/viscoplastic constitutive equations, is therefore suited for modeling the process. At the same time, the flow formulation offers computational advantages compared to traditional solid formulations based on elasto-plastic constitutive equations.

Resistance welding involves interaction between multiple objects in contact. The dynamically changing contacts, both in terms of areas and conditions, are important for the process due to the current concentration and heat generation associated with contact interfaces. Finite element contact modeling and modeling of physical contact properties are therefore a necessity for simulation of the process. Contact between deformable objects, or self-contact of an object, is handled by an implementation based on the penalty method that suppresses penetration by penalization in the equation system. Two algorithms are implemented for the identification of contact pairs. One algorithm divides quadrilateral element faces by one of two diagonals for setting up the surface normal, while the other algorithm divides the quadrilaterals by a temporary center node. The former is more direct and expands the system of equations (the skyline profile) the least, whereas the latter is preserving symmetry better. The division by a diagonal has been found sufficient in the majority of cases, but the division by the temporary center node has been used by default because of the higher accuracy and the relatively small amount of extra CPU time after all.

The implemented contact algorithms are verified against experiments. The most direct verification is made by mechanical tests at room temperature. Later presented resistance welding cases verify the implementation indirectly because the interactions between the models play a large role. Mechanical contact experiments at room temperature have been performed with geometries of a hemisphere pressed into a cylinder end face and also a similarity test of a cross-cylinder compression. Both sets of geometries were tested with similar and dissimilar materials designed for dynamic contact development with increasing compressive force. Cross-sections of the contact experi-

ments and the simulations show overall good agreement when compared, and force-displacement curves further support the verification of the contact implementation.

Further verification of the contact models have been shown by simulation of real metal forming applications involving mechanical contact. One example is forming of reservoirs with an internal, recyclable, deformable mandrel providing support on the inside and another example is forming of asymmetric compression beads that can be utilized in mechanical joining of tubes. In the later stage of the compression beading, the tube material folds to get in self-contact and is hereafter further formed to finalize the compression bead. These two cases were not part of the present project, but they serve as verification of the presented contact algorithms because they were simulated in I-Form after implementation of the contact algorithms developed in the present project.

Besides the implementation in I-Form, the contact algorithms are also implemented in SORPAS 3D, which is developed in the present project to a complete, standalone software package dedicated to simulation of resistance welding processes. A wide range of simulations are presented to show the capabilities and current limitations of the new computer program. The simulated welding applications include spot welding and projection welding, and additionally, a micro joining application is included. As regards spot welding, a three-sheet combination of two thicker, high strength steel sheets and a thin, low-carbon steel sheet is first analyzed because it resembles current interest and challenges in automotive assembly. A comparison to a corresponding experiment shows good agreement under carefully aligned conditions. By simulation, it is further investigated how the nugget formation, indentation and sheet separation develop in case of electrode misalignment, which can be caused by flexibility of the welding machine. Another simulated effect relevant to real production is shunting between consecutive spots. To conclude spot welding, a numerical analysis is presented for single-sided spot welding of a sheet to a tube, showing that the electrode force is critical and that the cooling during the hold time is relevant for the analysis of the temperature development. The single-sided spot welding case was analyzed based on interest from a German steel producer.

Simulated projection welding includes joining of sheets by three projections, where the middle projection is smaller than the two projections on the outside due to uneven punch movement when creating the projections. Another example utilized longitudinal embossments to facilitate joining of two perpendicular sheets. This example is an industrial case from Japan, which includes comparison between the simulated temperature field and the microstructural changes in the real case with good agreement observed. Another industrial case is joining of a bellows tube to a disc in the production of thermostat valves by a Danish company. This case was simulated by the existing two-dimensional version of SORPAS due to axisymmetry.

Two projection welding applications in form of cross-wire welding and welding of square nut to sheet are analyzed by experiments and used for exploring the limitations of the developed software. As regards cross-wire welding, good agreement between simulations and experiments is generally observed at low weld settings, that is for a combination of high electrode force and low weld current, where the material defor-

mation in terms of squeeze out between the two wires is moderate or absent. At high weld settings, where material is squeezed out, the presented simulations fail to model cross-wire welding because of lack of elements and mesh density control in remeshing to model the local details necessary for the squeeze out. The square nut to sheet projection welds present sliding surfaces, where the actual frictional behavior is of importance, while it can be ignored in many other cases due to symmetries. When simulated with frictionless contact, the simulations show too large collapse of the nut legs because of the easier outward deformation of the legs. When simulated with full sticking, the simulations capture the actual setdown better and include the barreling effect as also seen in the experiment. The actual contact behavior will be frictional and lie somewhere in-between the two extreme simulated cases, and based on the comparisons, it is expected to lie closer to the full sticking.

Finally, a last application is an industrial micro joining case by a German company. A fork and a wire are mechanically joined while softening the material by resistance heating, which at the same time is used to melt a polymer coating on the wire to establish local contact between the fork and the wire of importance to the actual end-product. This micro-joining case was successfully simulated with good comparison to the final geometry provided by the company.

As mentioned above, the resistance welding cases are simulated by the new computer program SORPAS 3D. This program is developed as a complete, standalone software that, besides the above electro-thermo-mechanical finite element core and contact modeling, comprises its own graphical interface, meshing and remeshing facilities and a parallelized skyline solver allowing parallel computing to speed up the calculations. The parallel computing is performed while keeping the solution of the main equation system direct, such that high accuracy is preserved. Speed-up and solution time has been evaluated by a benchmark test case and a welding case, both showing the obtained benefit utilizing parallel computing.

Meshing is accommodated by structured meshing or unstructured meshing for the initial geometry, and remeshing of selected objects is performed by unstructured meshing. The structured meshing is based on isoparametric meshing that was developed before, and outside, the present project. It has been adopted and built into the graphical interface that is enhanced to transfer element divisions and mesh density control automatically through neighboring superelements. Two neighboring superelements can be connected in 144 different ways (taking into account six faces and four different rotations), so it is important that this is taken care of automatically for easy setup of structured meshes. The unstructured meshing procedures are based on an existing all-hexahedral meshing algorithm that has been enhanced by adaptive selection of bounding boxes and the possibility of handling multiple objects in contact. The meshing procedures are improved in terms of versatility and robustness by replacing distance based criteria by mesh topology based criteria.

A graphical interface has been developed to operate the above mentioned facilities including setup of weld schedules and opening and showing of results in a mode allowing rotation, zoom and movement of the finite element model. It completes the finite element program as a standalone package, and it has given the possibility of pre-

senting the program to academia and industry as a useful tool for analyzing resistance welding processes. It was officially announced at the 7th International Seminar on Advances in Resistance Welding held in Busan, Korea in September 2012, where it was met with positive interest by the audience. A specific comment by a Canadian company with contacts in the automotive industry in Northern America highlights the immediate possibilities that SORPAS 3D offers; namely enhanced understanding and visualization of the process for the benefit of welding engineers. The comment was related to the example of spot welding simulated under conditions of electrode misalignment caused, e.g., by flexibility of the welding machine arms holding the electrodes. The interest by this example was drawn because customers of the Canadian company experience the problems related to electrode misalignment frequently, and it has been difficult to explain the causes. By the presented simulations, it was stated in the highlighted comment that it is now possible to show and explain the effects to the production companies experiencing these problems.

It is a wish by the presented project that many such problems can be further understood and improved by simulations of the process by SORPAS 3D. The long-term vision is that the utilization of SORPAS 3D will become a tool that can assist welding engineers in weld planning and that it can assist in solving complex problems where solutions have been difficult to explore by pure experimentation.

The first steps in this direction are taken by the fact that test users are already using the program and giving feed-back to guide further developments. The test users are among the contacts to SWANTEC Software and Engineering ApS. By the end of September, test users in Poland and Japan are using SORPAS 3D, and it is expected that a first license in Japan can start from the beginning of 2013.

These possibilities are given by the presented work, and they can be further extended in future work. The immediate improvements are seen on the surface of the program and include easier setup of the finite element meshes and welding parameters, and another immediate improvement is in terms of computation time as this will always remain an issue when dealing with three-dimensional problems. The parallelization of the skyline solver has improved the speed of solution significantly, such that simulated results of a complete welding case (with typical number of elements as presented throughout this thesis) can be available *tomorrow* rather than *next week*. However, further improvements will allow even further increase of solution speed and also the possibility of handling larger finite element models, which is necessary for analyzing more complex cases or further details of specific interest. Improvements of the solution time can be obtained by adopting a compressed row storage format for diminishing even more zeros in the system matrix than already achieved with the skyline format. Another possibility, which is more relevant for the larger models, is further development of the equation solver for parallelization on local memory processors, which will open for the possibility of running larger, complex simulations on clusters of computers.

From a process modeling view it will be natural to concentrate further on contact interfaces, both from a finite element point of view and from a physical modeling point of view. Among many published models for the electrical contact resistance, none are

able to model all experiments, and hence tuning is always necessary to smaller or larger degree. Similarly, better understanding of the thermal contact properties can enhance the overall modeling, though being of less importance than the electrical contact resistance.

Additional possibilities of implementation include prediction of microstructural changes, hardness prediction and strength prediction. Prediction of microstructure and hardness are potentially available by the already simulated temperature development if the initial microstructures of the simulated materials are known together with time-temperature-transformation (TTT) diagrams. Prediction of hardness may be related to microstructures and the chemical composition of the material as e.g. shown by the presented results simulated for spot welding by the two-dimensional version of SORPAS. The prediction of strength is more difficult and can take two different approaches. The first approach is intended to deal with simple geometries, e.g. spot welds, by empirical or semi-analytical models. The second approach is finite element modeling of the damage and fracture mechanisms involved, and if successful, this can handle complex geometries and testing under various conditions. This, however, is a major topic in itself and is considered out of scope of the near future, but is mentioned here for the (very) long-term possibilities of enhancing the program.

Another aspect of modeling is to enhance the versatility even more by being able to simulate local squeeze out of material as e.g. associated with the high weld settings in cross-wire welding. Besides the ability of handling larger models, it will also involve further development of remeshing procedures to follow the large deformation with hexahedral elements. With starting point in the presented all-hexahedral meshing techniques, the procedures may be enhanced by introducing octree based core meshes as an alternative to the implemented grid based core meshes. This will involve a systematic review of the meshing procedures to handle the different element sizes. It will allow mesh density control during remeshing (and in the initial meshing of objects as well), such that details can be modeled adequately while keeping the overall number of elements minimal.

To conclude the words on future work, it is inevitable to remind that complex finite element modeling goes hand in hand with continuous improvement, correction and further development when applied to more and more applications, which, each of them may present new challenges.

Appendix A

The FORTRAN source code including OpenMP instructions for the parallel skyline solver is listed as follows:

```

subroutine skyline_gauss_omp (skmatx, fmatx, &
                             maxa, nthreads, ntotv)

use omp_lib

implicit none

! -----
! This subroutine solves a regular system:
!   skmatx*x=fmatx
! where
! skmatx is the skyline vector of the system
!       matrix,
! fmatx  is the right hand side vector (in)
!       and later the vector of unknowns (out),
! x      is the vector of unknowns outputted
!       through fmatx
! The index vector pointing to the diagonal
! positions in skmatx is maxa.
! The number of degrees of freedom is ntotv.
! Number of threads to use during solving is
! nthreads.
! Method: Gaussian elimination with column
!       reduction.
!
! This skyline solver was originally provided
! in sequential form by
!   J.E.Akin, Finite Elements for Analysis
!   and Design, Academic Press, London, 1993.
! It is parallelized in the present work by
!   C.V.Nielsen and P.A.F.Martins

```

```

! -----

integer i, id, ie, ie0, ie0old, ih1, ih2, ihesitate, &
        iloop, iquit, ir, is, ithread, iwait, j, jd, &
        jh, jmax, jr, k, k0, k00, kmax, nthreads, ntotv
integer maxa(*)
double precision d, fmatx(*), skmatx(*)

! Set number of threads
call omp_set_num_threads (nthreads)

! Get (actual) number of threads
nthreads=omp_get_max_threads()

! Initializations
jmax=1
kmax=1

!$OMP parallel default (none) &
!$OMP private (d, i, id, ie, ie0, ie0old, ih1, ih2, ihesitate, &
!$OMP         iloop, iquit, ir, is, ithread, iwait, j, jd, jh, &
!$OMP         jr, k, k0, k00) &
!$OMP shared (fmatx, jmax, kmax, maxa, nthreads, ntotv, &
!$OMP         skmatx)

! Factorize skmatx and reduce fmatx
ithread=omp_get_thread_num()
iloop=0
iquit=0
do while (iquit.eq.0)
    iloop=iloop+1
    j=(iloop-1)*nthreads+ithread+2
    if (j.gt.ntotv) then
        iquit=1
        exit
    endif
    ! Characteristic positions in skyline
    jr=maxa(j-1)
    jd=maxa(j)
    jh=jd-jr
    is=j-jh+2
    ! Start of core code
    ie0=0
!$OMP flush (kmax)

```

```

do while (kmax.lt.jd)
  ! Initializations
  ihesitate=0
  ie0old=ie0
!$OMP flush (jmax,kmax)
  ! Judge if hesitation is necessary
  if (kmax.lt.jr) then
    ihesitate=1
    ie0=jmax
  endif
  if (jh.eq.2) then
    ! Reduce diagonal term
    iwait=1
    do while (iwait.eq.1)
!$OMP flush (kmax)
      if (kmax.ge.jr-1) then
        d=skmatx(jr+1)
        skmatx(jr+1)=d/skmatx(jr)
        skmatx(jd)=skmatx(jd)-d*skmatx(jr+1)
        iwait=0
      endif
    enddo
    ! Reduce right hand side (fmatx)
    fmatx(j)=fmatx(j)-skmatx(jr+1)*fmatx(j-1)
    ihesitate=0
  elseif (jh.gt.2) then
    ! Reduce all equations except diagonal
    ie=jd-1+(ie0-j+1)*ihesitate
    k00=jh-j-1+ie0old
    if (k00.lt.0) k00=0
    k0=0
    do k=max0(jr+2,jd-j+ie0old+1),ie
      ir=maxa(is+k0+k00-1)
      id=maxa(is+k0+k00)
      ih1=min0(id-ir-1,1+k0+k00)
      if (ih1.gt.0) then
        ih2=min0(id-ir-j+(j-1-k0-k00) &
          *ihesitate,&
          2-j+k0+k00+(j-1-k0-k00) &
          *ihesitate)
        if (ih2.lt.1) ih2=1
        skmatx(k)=skmatx(k) &
          -dot_product(&
            skmatx(k-ih1:k-ih2), &

```

```

                                skmatx(id-ih1:id-ih2)
                                endif
                                k0=k0+1
                                enddo
                                if (ihesitate.eq.0) then
                                ! Reduce diagonal term
                                ir=jr+1
                                ie=jd-1
                                k=j-jd
                                do i=ir,ie
                                id=maxa(k+i)
                                d=skmatx(i)
                                skmatx(i)=d/skmatx(id)
                                skmatx(jd)=skmatx(jd)-d*skmatx(i)
                                enddo
                                ! Reduce right hand side (fmatx)
                                fmatx(j)=fmatx(j) &
                                -dot_product(skmatx(jr+1:jr+jh-1), &
                                fmatx(is-1:is+jh-3))
                                endif
                                endif
                                if (ihesitate.eq.0) then
!$OMP critical
                                if (j.gt.jmax) then
                                jmax=j
                                kmax=jd
                                endif
!$OMP end critical
                                endif
!$OMP flush (jmax,kmax)
                                enddo
                                enddo

!$OMP end parallel

!      Divide by diagonal pivots
                                do i=1,ntotv
                                id=maxa(i)
                                fmatx(i)=fmatx(i)/skmatx(id)
                                enddo

!      Back substitution
                                j=ntotv
                                jd=maxa(j)

```

```
100      d=fmatx(j)
        j=j-1
        if (j.le.0) return
        jr=maxa(j)
        if (jd-jr.gt.1) then
            is=j-jd+jr+2
            k=jr-is+1
            fmatx(is:j)=fmatx(is:j)-skmatx(is+k:j+k)*d
        endif
        jd=jr
        goto 100

        return

    endsubroutine skyline_gauss_omp
```

