

Finite Element Simulation of Heat Transfer

Jean-Michel Bergheau
Roland Fortunier

ISTE

 **WILEY**

This page intentionally left blank

Finite Element Simulation of Heat Transfer

This page intentionally left blank

Finite Element Simulation of Heat Transfer

Jean-Michel Bergheau
Roland Fortunier

ISTE

 **WILEY**

First published in France in 2004 by Hermes Science/Lavoisier entitled "Simulation numérique des transferts thermiques par éléments finis"

First published in Great Britain and the United States in 2008 by ISTE Ltd and John Wiley & Sons, Inc.

Translated from the French by Robert Meillier

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988, this publication may only be reproduced, stored or transmitted, in any form or by any means, with the prior permission in writing of the publishers, or in the case of reprographic reproduction in accordance with the terms and licenses issued by the CLA. Enquiries concerning reproduction outside these terms should be sent to the publishers at the undermentioned address:

ISTE Ltd
27-37 St George's Road
London SW19 4EU
UK

www.iste.co.uk

John Wiley & Sons, Inc.
111 River Street
Hoboken, NJ 07030
USA

www.wiley.com

© ISTE Ltd, 2008

© LAVOISIER, 2004

The rights of Jean-Michel Bergheau and Roland Fortunier to be identified as the authors of this work have been asserted by them in accordance with the Copyright, Designs and Patents Act 1988.

Library of Congress Cataloging-in-Publication Data

Bergheau, Jean-Michel.

[Simulation numérique des transferts thermiques par éléments finis. English]

Finite element simulation of heat transfer / Jean-Michel Bergheau, Roland Fortunier.

p. cm.

Includes bibliographical references and index.

ISBN 978-1-84821-053-0

1. Heat--Transmission--Mathematical models. 2. Finite element method. I. Fortunier, Roland. II. Title.

TJ260.B45413 2008

621.402'2015118--dc22

2008025105

British Library Cataloguing-in-Publication Data

A CIP record for this book is available from the British Library

ISBN: 978-1-84821-053-0

Printed and bound in Great Britain by CPI/Antony Rowe Ltd, Chippenham, Wiltshire.



Mixed Sources

Product group from well-managed
forests and other controlled sources

Cert no. SGS-COC-2953

www.fsc.org
© 1996 Forest Stewardship Council

Table of Contents

Introduction	11
PART 1. Steady State Conduction	17
Chapter 1. Problem Formulation	21
1.1. Physical modeling	21
1.1.1. Thermal equilibrium equation	21
1.1.2. Fourier law	22
1.1.3. Boundary conditions	23
1.2. Mathematical analysis	24
1.2.1. Weighted residual method	25
1.2.2. Weak integral formulation	27
1.3. Working example	30
1.3.1. Physical modeling	30
1.3.2. Direct methods	32
1.3.2.1. Analytical integration	32
1.3.2.2. The finite difference method	33
1.3.3. Collocation methods	34
1.3.3.1. Point collocation	35
1.3.3.2. Sub-domain collocation	36
1.3.4. Galerkin method	37
1.3.4.1. Polynomial functions	37
1.3.4.2. Piecewise linear functions	39
Chapter 2. The Finite Element Method	43
2.1. Finite element approximation	43
2.1.1. Mesh	43
2.1.2. Nodal approximation	46
2.2. Discrete problem formulation	48

2.2.1. Element quantities	49
2.2.2. Assembly	51
2.3. Solution	53
2.3.1. Application of temperature boundary conditions	53
2.3.2. Linear system solution	56
2.3.2.1. Direct methods	58
2.3.2.2. Iterative methods	60
2.3.3. Storing the linear system matrix	62
2.3.4. Analysis of results	63
2.3.4.1. Smoothing the heat flux density	64
2.3.4.2. Result accuracy	66
2.4. Working example	68
2.4.1. Finite element approximation	70
2.4.1.1. Mesh	70
2.4.1.2. Nodal approximation	71
2.4.2. Discrete problem formulation	72
2.4.2.1. Element quantities	72
2.4.2.2. Assembly	74
2.4.3. Solution	75
2.4.3.1. Application of boundary conditions	75
2.4.3.2. Solution	77
Chapter 3. Isoparametric Finite Elements	79
3.1. Definitions	79
3.1.1. Reference element	79
3.1.1.1. Triangular element with linear transformation functions . . .	81
3.1.1.2. Quadrangle element with linear transformation functions . .	82
3.1.1.3. Quadrangle element with quadratic transformation functions	84
3.1.2. Isoparametric elements	85
3.1.3. Interpolation function properties	89
3.2. Calculation of element quantities	90
3.2.1. Expression in the reference frame	91
3.2.2. Gaussian quadrature	93
3.2.2.1. 1D numerical integration	94
3.2.2.2. 2D and 3D numerical integration	97
3.3. Some finite elements	99
PART 2. Transient State, Non-linearities, Transport Phenomena	101
Chapter 4. Transient Heat Conduction	105
4.1. Problem formulation	105
4.1.1. The continuous problem	105
4.1.2. Finite element approximation	107
4.1.3. Linear case	109

4.2. Time integration	111
4.2.1. Modal method	111
4.2.1.1. Determining the modal basis	112
4.2.1.2. Projection on the modal basis	114
4.2.2. Direct time integration	115
4.2.3. Accuracy and stability of a direct integration algorithm	119
4.2.3.1. Accuracy	120
4.2.3.2. Stability	121
4.2.3.3. Simplified analysis of the stability condition	122
4.2.4. Practical complementary rules	124
4.2.4.1. Space oscillations during thermal shock simulation	124
4.2.4.2. Discrete maximum principle	128
4.2.4.3. Initial temperatures during thermal contact simulation	130
4.3. Working example	135
4.3.1. Physical modeling and approximation	135
4.3.2. Numerical applications	139
Chapter 5. Non-linearities	143
5.1. Formulation and solution techniques	143
5.1.1. Formulation	143
5.1.2. Non-linear equation system solution methods	144
5.1.2.1. Newton-Raphson method	147
5.1.2.2. Substitution method	149
5.1.2.3. Quasi-Newton methods	150
5.1.3. Line search method	152
5.2. Traditional non-linearities	153
5.2.1. Physical properties	153
5.2.2. Flux or volumetric heat source boundary conditions	155
5.2.3. Modeling state changes	157
5.2.3.1. Equivalent specific heat method	158
5.2.3.2. Enthalpy solution method	160
5.3. A temperature-enthalpy formulation	162
5.3.1. Mathematical formulation	163
5.3.2. Example	166
Chapter 6. Transport Phenomena	169
6.1. Highlighting instabilities	169
6.1.1. Thermal balance	169
6.1.2. Treating a simple case	171
6.2. Resolution techniques	174
6.2.1. Upwind technique	175
6.2.2. SUPG method	177
6.2.3. 2D and 3D Petrov-Galerkin formulation	180

PART 3. Coupled Phenomena	183
Chapter 7. Radiation Exchanges in a Chamber	189
7.1. Modeling radiative heat exchanges in a cavity	189
7.1.1. Posing the problem	190
7.1.2. Calculation of view factors	194
7.1.3. Diffusion-radiation coupling	197
7.1.3.1. Tangent matrix	198
7.1.3.2. Substitution matrix	199
7.2. Examples	200
7.2.1. Radiation between two walls	200
7.2.2. Cylinder quenching	203
Chapter 8. Fluid-Structure Coupling in a Pipe	207
8.1. Modeling the fluid	207
8.1.1. Physical model and mathematical formulation	207
8.1.2. Modeling the coupling	210
8.2. Example	212
8.2.1. Physical and geometric modeling	212
8.2.2. Results	213
Chapter 9. Thermometallurgical Coupling	215
9.1. Modeling phase changes	215
9.1.1. Rate of phase changes	215
9.1.1.1. Avrami kinetics	215
9.1.1.2. Martensitic kinetics	217
9.1.2. Numerical integration	217
9.1.3. The case of several phase changes	220
9.1.4. Modeling the coupling	221
9.2. Examples	222
9.2.1. Phase transformation diagrams	223
9.2.2. Steel quenching	227
Chapter 10. Thermochemical Coupling	231
10.1. Finite element simulation of simultaneous diffusion and precipitation	231
10.1.1. Governing equations	232
10.1.2. Finite element formulation	234
10.2. Calculation of precipitation	236
10.2.1. Mathematical formulation	236
10.2.2. Numerical scheme	238
10.3. Examples	239
10.3.1. Calculation of a phase diagram	239
10.3.2. Carbon diffusion in a titanium steel	240

Chapter 11. Electrothermal Coupling	243
11.1. Electrokinetic modeling	243
11.1.1. Weak formulation	243
11.1.2. Modeling the coupling	244
11.1.3. Solving the coupled problem	246
11.2. Resistance welding	248
11.2.1. Implementing the model	249
11.2.2. Results	251
Chapter 12. Magnetothermal Coupling	253
12.1. Introduction	253
12.2. Magnetic vector potential formulation for magnetodynamics	254
12.3. Coupled finite element-boundary element method	257
12.3.1. Finite element formulation	259
12.3.2. Boundary element formulation	260
12.3.3. FEM-BEM coupling	261
12.4. A harmonic balance method for the magnetodynamic problem	261
12.5. Coupling magnetodynamics with heat transfer	263
12.5.1. Iterative coupling	263
12.5.2. A direct method for magnetothermal coupling	265
12.6. Application: induction hardening of a steel cylinder	266
Bibliography	269
Index	277

This page intentionally left blank

Introduction

In their daily practice, professional engineers are often confronted with problems involving complex physical phenomena. Whether they have to analyze the behavior of a product or optimize an implementation process, modeling these phenomena often makes it possible to solve these problems:

- by gaining access to the physical magnitudes characterizing the product and the consequences of their modification,
- by highlighting the main parameters of the process and thus its optimal conditions of use.

Physical modeling seems to be the key point when an engineer is thinking of how to solve a scientific problem. Generally speaking, modeling follows these three steps:

- 1) identifying the essential physical phenomena related to the behavior of a product or process within a given context,
- 2) defining the mathematical equations (domain of study, partial differential equation(s), boundary conditions and stresses, initial conditions, etc.) governing these phenomena,
- 3) validating the equations with respect to the necessary data which must be available or accessible by means of measurements as well as the results provided, which must be reliable and reproducible.

The first two points need commenting upon. Firstly, physical modeling has to meet a set objective. It is then obvious that several modeling methods are possible to meet a similar objective. The most relevant model is the simplest one making it possible to meet the set objective. Modeling is first and foremost making (simplifying) assumptions. Some are obvious. For instance, it is unnecessary to take into account heat radiation phenomena in low temperature applications; a mean

exchange coefficient, independent of the temperature, will be sufficient. In other cases, it is far more difficult. For example, how can we define the analyzed field and include the behavior of the parts not addressed by means of carefully selected boundary conditions? This is exactly where the professional engineer's contribution comes in. We will come back to this later on.

The last point of this definition is about the model input data and results. The selection of input data (availability, validity, etc.) often guides the users in the choice of their modeling methods. It is unnecessary, for instance, to use very sophisticated metallurgical transformation models if sufficiently accurate data on the processed material are not available. Result reliability and reproducibility guide the users in the choice and use of the model solution method. The aim of this method (numerical or analytical) is to determine a solution of the problem with minimum approximation. It is obvious that an accurate analytical solution will always be given priority; however, the growing complexity of the model often leads today's users towards approximate numerical solutions. Moreover, commercial numerical simulation software often includes the functions and models required for industrial applications, thus making it possible to take full advantage of the progress of information technology (rapid calculation, memory resources, graphic visual aids, etc.). Consider for instance the boundary value problem related to steady state heat exchanges in a solid occupying a limited field (equation [1.8]), which reveals:

- a volumetric heat source Q most often resulting from other physical phenomena such as the Joule effect in conduction heating or induction heating applications. Is it necessary to model these phenomena by adding the corresponding partial differential equations to the heat equation? This is not certain but then we require simple analytical models to evaluate Q ;

- a surface density of the heat flux q applied to a portion of the solid boundary, reflecting the heat exchange between this solid and the outside medium. This can be for instance heat radiation towards an infinite medium or a liquid flow into a quenching bath. Here again, is it necessary to include those phenomena in the modeling process with the addition of complementary equations? It depends on the application considered and the feasibility of such an approach. However, if, whatever the reason, fine modeling of these phenomena was not envisaged, other approaches would then be necessary to determine the boundary conditions. These could be, for instance, simplified analytical models;

- a temperature prescribed on the other portion of the solid boundary. This condition is a limit case of the previous condition. As a matter of fact, either the solid is steeped at this spot in a fluid with a very high exchange coefficient so as to prescribe the fluid temperature on the corresponding boundary, or the user has a measurement.

Experimental measurements will frequently be used to determine the missing input data, but these will be obtained by alternative routes as the data required for modeling are often inaccessible to direct measurements. The methods called inverse methods have developed significantly over the last few years. They are able to couple physical modeling with accessible magnitude experimental measurements by adjusting iteratively an input data until the whole set of calculated results are as close as possible to the measurements.

It is tempting to say after these comments that the easiest model to elaborate is one in which all influential physical phenomena are finely modeled. This is partly true; however:

- Are we really in a position to define the mathematical equations governing these phenomena with a sufficient degree of accuracy to make it worthwhile?
- Are we certain that we are able to have the data required to feed these models all the more so that these models will frequently require unusual data on a different scale, which is thus hardly accessible? Using inverse methods may solve this type of difficulty.
- Will it not be often more difficult to interpret the results than with a simpler model given the amount of information to process? Indeed, will it not be necessary to conduct *a posteriori* the analysis which will not have been carried out *a priori* to eliminate insignificant results?

In fact, everything depends on the problem to solve. Moreover at which point of the domain of study considered should we stop? First, we will take into consideration all the symmetries presented by the problem for which the corresponding boundary conditions are written clearly from a mathematical point of view:

- revolution symmetry: the most productive one as it allows us to carry out 3D analysis on a plane model representing a meridian section of the studied structure,
- symmetry to a plane: very commonly used,
- anti-symmetry to a plane: less natural and often forgotten,
- periodicity conditions on a repetitive structure.

The relevance of a 3D model will then be questioned. If it is true that calculation software and computers make it easy to carry out this type of analysis, a 2D (and even 1D) model is often sufficient (in the case of a repetitive structure in one direction for example) and therefore preferable as the analysis and interpretation of results will always be easier with a 2D than with a 3D model. In all other cases, restricting the domain of study will lead to the definition of appropriate boundary conditions.

It is therefore in the interest of the engineers in charge of physical modeling to think in detail about the relevance of their choices and assumptions. It is the price to pay to take maximum advantage of their models. In any case, before reaching the intensive exploitation stage, it is imperative to ensure the model prediction quality by comparing it with the results of one experiment (at least!).

Most often, analytical methods cannot be applied to the solution of the mathematical equations governing a set of physical phenomena, except if major assumptions reducing the modeling validity are made. The analog method takes advantage of the fact that conduction heat exchanges and electrical conduction phenomena are governed by the same equations. It is thus possible to study conduction heat exchanges by means of more easily accessible measurements carried out on a similar electrical device. However, the field of application of this method is far too restrictive. Therefore, numerical techniques whose use is made possible by the performance of today's computers are used to determine an approximate solution of the set of mathematical equations governing the problem.

The finite difference method which replaces partial derivatives with finite differences at different points of a grid is highly regarded by mechanical engineers. The discrete equation system obtained can also be interpreted as resulting from a complex electrical diagram combining resistances and capacities. This technique is hardly possible with complex geometries.

The finite volume method is also highly regarded by mechanical and thermal engineers. It is based upon a previous division of the geometric domain of study into element volumes. On each element volume, the thermal balance equations are solved. This method is particularly efficient in the case of structured geometries. It is widely used in thermal science and fluid mechanics, but more rarely in other disciplines.

Among the various numerical techniques available today, the finite element method [TOU 81, ZIE 91] is the most widespread owing to:

- its general fields of application (thermal, electromagnetic sciences, solid mechanics, fluid mechanics, etc.),
- its capacity to treat problems with complex geometries,
- its easy implementation.

This can be done as follows:

1) mesh-geometry: the geometric domain to be analyzed, most often resulting from CAD geometric modeling, is divided into a set of element sub-domains (finite elements) interconnected by nodes;

2) solution: the continuous functions sought (the temperature in a heat conduction problem) are replaced with a set of values estimated at the mesh nodes. This approximation, applied to an integral formulation of the problem, leads to a system of equations (linear or non-linear) whose number is equal to the number of values to be estimated;

3) analysis-interpretation of results: at this point, other results can be calculated (for instance, the heat flux density) from those obtained by the direct solution of the equation system. They are then analyzed and interpreted by the user, by means of very efficient graphic visualization resources.

The success of the finite element method is largely due to the significant progress of information technologies, both from a numerical point of view (rapid calculation, memory size available) and from a graphic point of view (3D visual resources). Today a large amount of software makes use of this method. They offer a growing number of functions and are increasingly user-friendly. It should be noted that calculation software displays clear and user-friendly interfaces nowadays and can be used by non-specialist engineers.

However, this apparent facility should not conceal the fact that, whatever the numerical method adopted, the discretization phase impairs the properties of the initial continuous model. Some phenomena present in the continuous model could be erased by the numerical model if we do not pay attention. The meshing phase in the finite element method is therefore very important and it is the user's know-how that will produce a quality approximate solution.

Therefore, the aim of this book is to present the basis and application of the finite element method to the solution of industrial thermal problems. It consists of three parts which the reader may possibly complete by reading a number of books related to this field [COM 94, RED 94, LEW 96, MIN 06].

Part 1, dedicated to the solution of steady state heat conduction problems, introduces the finite element method. Starting with the partial derivative problem and the related boundary conditions, various formulations upon which various discretization methods are based, are presented.

Part 2 extends the field of application of the method to transient state conduction problems, the most common non-linearities and transport phenomena (diffusion convection problems).

The last part, Part 3, deals with coupled problems:

- coupled by boundary conditions: radiation problems, fluid and structure coupling in a piping system,

- including additional state variables: thermometallurgical coupling,
- coupled by partial differential equations: electrothermal coupling, magnetothermal coupling and thermochemical coupling.

This book is a survey of the various thermal problems which professional engineers may have to simulate. The methods presented will allow readers to use in the best way possible a calculation software and design new calculation modules so as to complete their work.

PART 1

Steady State Conduction

This page intentionally left blank

Introduction

In this part we introduce the finite element method in the simplest field of thermal science: steady state conduction. We therefore consider a solid Ω undergoing thermal loads and try to determine the temperature field $T(\vec{x})$ in this solid, when it is in equilibrium with the external environment. In this part, it is assumed that no quantity involved depends on time. Thermal loads are therefore constant and the solid is motionless. The equilibrium state sought corresponds to a steady state. Besides, we will restrict our study to the linear case in which the mathematical field is clearly identified. The finite element method is introduced according to the following three steps.

Firstly, the physical problem to solve is analyzed from a mathematical point of view. This leads to a series of three problems (equations [1.8], [1.10] and [1.13]). Problem [1.8] is a partial differential equation resulting directly from physical modeling with its boundary conditions. Problem [1.10] turns this partial differential equation into a variational equation by means of the weighted residual method. Finally, problem [1.13], often called weak form, is the basis of the finite element method.

It is possible to find an approximate solution to each of the three problems [1.8], [1.10] and [1.13]. For this purpose, several methods are used. Each of them will be illustrated by a work example: an induction-heated plate. We will use a physical model to describe this example, then seek approximate solutions (the temperature field) of problems [1.8], [1.10] and [1.13] related to this physical model.

Chapter 2 deals with the finite element method. This method is based upon the weak form described in Chapter 1. The approximation used is called finite element approximation. It consists of discretizing the geometry (the mesh) and approximating the temperatures sought (nodal approximation). This step makes it possible to write

the problem to solve in a discretized form, which is very appropriate for a computer numerical solution. In this chapter, the various steps of the finite element method will be illustrated by a new work example: thermal conduction in a plate with holes.

In Chapter 3, we introduce isoparametric finite elements starting from the notion of the reference element. This makes it possible to develop simple and rapid methods for the calculation of the element quantities involved in the discrete problem formulation. The major types of isoparametric elements are also described in detail in this chapter.

Chapter 1

Problem Formulation

1.1. Physical modeling

In this chapter we describe the different steps of physical modeling leading in the next section to a boundary value problem. These steps are:

- 1) writing the equation expressing the solid thermal equilibrium,
- 2) introducing the Fourier law connecting the heat flux to the temperature gradient,
- 3) formulating boundary conditions.

1.1.1. *Thermal equilibrium equation*

Figure 1.1 illustrates a homogenous solid Ω . In order to write that this solid is in thermal equilibrium, consider any portion Ω_A related to this solid and write that the heat produced on that portion is equal to the heat flux coming out of it, i.e.:

$$\forall \Omega_A, \quad \int_{\Omega_A} Q \, dv = \int_{\partial\Omega_A} \vec{\phi} \cdot \vec{n} \, ds \quad [1.1]$$

In this equation, $\vec{\phi}$ is a vector characterizing the heat flux surface density (in W/m^2) coming out of Ω_A through its boundary $\partial\Omega_A$, \vec{n} is the outward unit normal to this surface, and Q is a scalar representing an internal heat volumetric source (in W/m^3) in Ω_A . Among the physical phenomena represented by this volumetric term, we can include Joule effect heating (conduction or induction), heat dissipation by plastic deformation, etc.

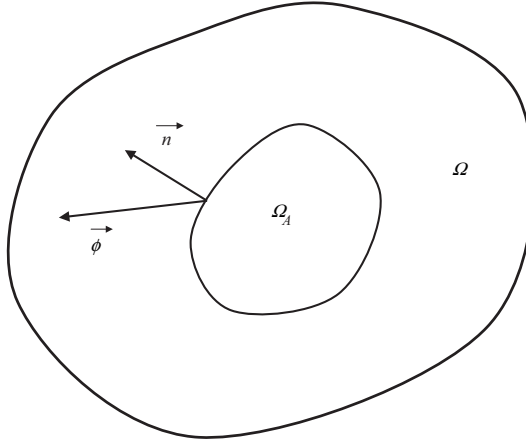


Figure 1.1. *Solid thermal equilibrium*

If the divergence theorem is now applied (integration by parts) to the surface integral of equation [1.1], i.e.:

$$\int_{\partial\Omega_A} \vec{\phi} \cdot \vec{n} ds = \int_{\Omega_A} \text{div}(\vec{\phi}) dv$$

it is possible to write the thermal equilibrium of any portion Ω_A of the solid Ω , in Figure 1.1 as follows:

$$\forall \Omega_A, \quad \int_{\Omega_A} (Q - \text{div}(\vec{\phi})) dv = 0$$

If this equation is verified in any domain Ω_A of the solid Ω we obtain the partial differential equation governing the thermal equilibrium:

$$Q - \text{div}(\vec{\phi}) = 0 \quad \text{at any point of } \Omega \quad [1.2]$$

1.1.2. Fourier law

The law connecting the heat flux density to the temperature field must satisfy the second principle of thermodynamics. In this book, we will not deal with the details of continuum thermodynamics. For that purpose, the reader should consult books related to this discipline. Here we will deal with the simplest law, the Fourier law, connecting linearly the heat flux density to the local temperature gradient. This is

written $\vec{\phi} = -\lambda \overrightarrow{\text{grad}}(T)$, where λ is a positive scalar called thermal conductivity. A temperature variation according to a space direction (i.e. a component of $\overrightarrow{\text{grad}}(T)$) will generate a heat flux in the same direction but inverted (i.e. from hot to cold).

The thermal conductivity of a material may be a function of the local temperature and the space direction in which thermal transfer occurs. In this book we will therefore use a law which is more general than the one described above:

$$\vec{\phi} = -\underline{\lambda} \cdot \overrightarrow{\text{grad}}(T) \quad [1.3]$$

In this equation, $\underline{\lambda}$ is a symmetric second-order tensor in which each component λ_{ij} , possibly as a function of temperature, produces the influence of a thermal gradient in the direction j on the heat flux density in the direction i . These components make up the thermal conductivity matrix. In the case of isotropic materials (similar thermal conductivity in all space directions), this matrix comes down to $\underline{\lambda} = \lambda \underline{I}$, where \underline{I} is the identity matrix, and we go back to our initial expression. If we now insert Fourier law [1.3] into equilibrium equation [1.2], this becomes:

$$\text{div} (\underline{\lambda} \cdot \overrightarrow{\text{grad}}(T)) + Q = 0 \quad \text{at any point of } \Omega \quad [1.4]$$

The equation obtained is a partial differential equation in T . Indeed, the temperature field sought is a function $T(\vec{x})$, where \vec{x} characterizes the position of a point in the solid Ω . For example, in Cartesian coordinates (x, y, z) , and in the case of an isotropic material whose thermal conductivity λ is independent of temperature, this equation can be written as follows:

$$\lambda \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} \right) + Q = 0 \quad \text{at any point of } \Omega$$

1.1.3. Boundary conditions

In order to solve partial differential equation [1.4] expressing the thermal equilibrium of the solid Ω , it is necessary to add boundary conditions. Those conditions express the connection of Ω with the external environment. This connection is generally represented on Ω boundary either by a known flux density $\vec{\phi} \cdot \vec{n}$, or by a known temperature T .

A known flux density condition q on the portion $\partial\Omega_q$ of Ω boundary is generally written as follows:

$$\vec{\phi} \cdot \vec{n} = (-\underline{\lambda} \cdot \overrightarrow{\text{grad}}(T)) \cdot \vec{n} = -q(T) \quad \text{at any point of } \partial\Omega_q \quad [1.5]$$

where \vec{n} is the outward unit normal at the point considered and q is the flux density entering Ω at this point, this density possibly being a function of the temperature T at the same point. This type of condition is used to represent, for instance:

- convection exchanges of the form $q(T) = h(T_{\text{ext}} - T)$, where h is a local heat exchange coefficient (in $\text{W/m}^2/\text{K}$) and T_{ext} is the external environment temperature (Fourier condition),

- radiation exchanges in an infinite environment $q(T) = \sigma\epsilon(T_{\infty}^4 - T^4)$ where σ is Stefan constant ($\sigma = 5,67 \cdot 10^{-8} \text{ W/m}^2/\text{K}^4$), ϵ is the emissivity of the surface assumed to be “gray”, T is the Kelvin temperature at the point considered, and T_{∞} is the Kelvin temperature of the infinite environment.

When the surface flux density q is not a function of the temperature, this boundary condition is called the Neuman condition. It is used for instance to represent:

- an adiabaticity condition ($q = 0$),
- a symmetry condition ($q = 0$ in the symmetry plane).

A temperature condition T_d known on a portion $\partial\Omega_T$ of Ω boundary is written as follows:

$$T = T_d \quad \text{at any point of } \partial\Omega_T \quad [1.6]$$

This type of condition, known as the Dirichlet condition, is in fact a convection exchange limit case, with a very high heat exchange coefficient h and T_d equal to the temperature of the external environment. It is generally used when local experimental temperature measurements are available. It can also be used to represent an anti-symmetry condition, but this is not very frequent in practice.

1.2. Mathematical analysis

For a thermal problem to accept a unique solution in a steady state (problem correctly posed), it is necessary:

- to prescribe at any point of the boundary $\partial\Omega$ of Ω either a known flux density condition (equation [1.5]), or a known temperature condition (equation [1.6]),
- to prescribe at one point at least of $\partial\Omega$ either a known temperature condition or a Fourier condition.

The boundary $\partial\Omega$ can then be broken down into a portion $\partial\Omega_q$ on which a flux condition is applied, and a portion $\partial\Omega_T$ on which the temperature is known. These two portions must satisfy the following conditions:

$$\partial\Omega_T \cap \partial\Omega_q = \emptyset \quad \text{and} \quad \partial\Omega_T \cup \partial\Omega_q = \partial\Omega \quad [1.7]$$

The boundary value problem expressing the equilibrium of the solid Ω can then be written as:

$$\left\{ \begin{array}{l} \text{Find } T(\vec{x}) \text{ at any point } \vec{x} \in \Omega \text{ such that:} \\ R(T) = \text{div}(\underline{\lambda} \cdot \overrightarrow{\text{grad}}(T)) + Q = 0 \quad \text{in } \Omega \\ (\underline{\lambda} \cdot \overrightarrow{\text{grad}}(T)) \cdot \vec{n} = q \quad \text{on } \partial\Omega_q \\ T = T_d \quad \text{on } \partial\Omega_T \end{array} \right. \quad [1.8]$$

The term $R(T)$ involved in this problem is said to be residual. It is a quantity to cancel in the solid Ω while satisfying the boundary conditions representing the connections of the solid with the external environment.

It is possible to work directly on problem [1.8] to determine the temperature field for $T(\vec{x})$ satisfying the thermal equilibrium of the solid. This can be done in a few simple cases in an analytical way, but in most cases, only an approximate solution can be determined. The finite difference method is the first possibility to obtain an approximation of problem [1.8]. In this case, the first- and second-order partial derivatives appearing are evaluated from the first terms of a Taylor sequence development of the temperature field.

1.2.1. *Weighted residual method*

In problem [1.8] we have to cancel a scalar residual $R(T)$ at any point \vec{x} of the solid Ω . First, note that if a temperature field $T(\vec{x})$ cancels this residual, then for any function $\psi(\vec{x})$:

$$\int_{\Omega} \psi R(T) dv = 0 \quad [1.9]$$

The functions $\psi(\vec{x})$ introduced by this method are called weighting functions. The only condition required for these functions is to give a sense to equation [1.9]. These functions will then have to be integrable on Ω . The set of weighting functions $\psi(\vec{x})$ will be written E_{ψ} .

With regard to the temperature field $T(\vec{x})$ sought, it also has to give a sense to equation [1.9]. Due to the expression of the residual $R(T)$ these functions, called admissible, are to have second-order partial derivatives integrable over Ω . The set of admissible functions $T(\vec{x})$ will be written E_T .

The weighted residual method therefore produces the following problem formulation:

$$\left| \begin{array}{l} \text{Find } T \in E_T \text{ such that for any } \psi \in E_\psi: \\ \int_{\Omega} \psi (\operatorname{div} (\underline{\lambda} \cdot \overrightarrow{\operatorname{grad}}(T)) + Q) dv = 0 \\ (\underline{\lambda} \cdot \overrightarrow{\operatorname{grad}}(T)) \cdot \vec{n} = q \quad \text{on } \partial\Omega_q \\ T = T_d \quad \text{on } \partial\Omega_T \end{array} \right. \quad [1.10]$$

It is obvious that, if a temperature field is the solution of [1.8], then it is also the solution of [1.10], whatever E_ψ is. On the contrary, a temperature field which is the solution of problem [1.10] depends on the choice of the sets of functions E_T and E_ψ .

Problem [1.10] is ideal for seeking an approximation of the temperature field $T(\vec{x})$. For this purpose, it is sufficient to choose particular spaces E_ψ and E_T . Partial differential equation [1.4] is then no longer satisfied at any point of Ω ; it is only satisfied in the sense of equation [1.10]. Point collocation and sub-domain collocation methods are based upon formulation [1.10].

Point collocation consists of defining a finite number n of points in the solid Ω with coordinates $\vec{x}_1, \dots, \vec{x}_n$, and choosing for E_ψ the set of Dirac distributions related to them:

$$E_\psi = \{\delta_{\vec{x}_i}, i = 1, \dots, n\}$$

The resolution of problem [1.10] then comes down to seeking a function $T(\vec{x})$ which can be differentiated twice, satisfying the boundary conditions, and canceling the residual at each point \vec{x}_i of Ω :

$$\int_{\Omega} \delta_{\vec{x}_i} R(T) dv = R(T)|_{\vec{x}=\vec{x}_i} = 0 \quad \text{for } i = 1, \dots, n$$

For example, if $T(\vec{x})$ is chosen in a polynomial form, the resolution of these equations allows us to determine the polynomial coefficients. The quality of the solution obtained with this method depends directly on the form selected for the temperature field (set E_T) and the choice of collocation points (number and positions).

Sub-domain collocation is the basis of the finite volume method. It consists of dividing the solid Ω into element volumes Ω_i with no intersection, and in such a way that their union forms Ω . Then, the following weighting functions are selected:

$$E_\psi = \left\{ \psi_i / \psi_i(\vec{x}) = \begin{cases} 1 & \text{if } \vec{x} \in \Omega_i \\ 0 & \text{otherwise} \end{cases} \right\}$$

The resolution of problem [1.10] then comes down to seeking a solution $T(\vec{x})$ satisfying the thermal equilibrium on each element volume or control volume Ω_i . As a matter of fact, using this type of weighting function makes it possible to write on each control volume:

$$\int_{\Omega} \psi_i R(T) dv = \int_{\Omega_i} R(T) dv = 0$$

With the divergence theorem and Fourier law [1.3], this equation can be expressed as follows:

$$\forall i, \quad \int_{\Omega_i} Q dv = \int_{\partial\Omega_i} \vec{\phi} \cdot \vec{n} ds$$

Then we go back to equation [1.1], which must be satisfied for a particular set of sub-domains Ω_i . This equilibrium equation is the basis of the finite volume method. The temperature is then often assumed to be homogenous in an element volume. The heat fluxes coming out of the volume are then calculated, in each direction, from local laws using the temperatures of adjacent volumes.

1.2.2. Weak integral formulation

The weak integral formulation of problem [1.10] is introduced when the volume integral equation involved has a specific form. If the divergence theorem (integration by parts) is applied to this integral equation, we can write:

$$\begin{aligned} & \int_{\Omega} \psi (\operatorname{div} (\underline{\lambda} \cdot \overrightarrow{\operatorname{grad}}(T))) dv \\ &= \int_{\Omega} \operatorname{div} (\psi (\underline{\lambda} \cdot \overrightarrow{\operatorname{grad}}(T))) dv - \int_{\Omega} \overrightarrow{\operatorname{grad}}^T(\psi) \cdot \underline{\lambda} \cdot \overrightarrow{\operatorname{grad}}(T) dv \\ &= \int_{\partial\Omega} \psi (\underline{\lambda} \cdot \overrightarrow{\operatorname{grad}}(T)) \cdot \vec{n} ds - \int_{\Omega} \overrightarrow{\operatorname{grad}}^T(\psi) \cdot \underline{\lambda} \cdot \overrightarrow{\operatorname{grad}}(T) dv \end{aligned}$$

The temperature field sought must now be such that, for any weighting function ψ of E_ψ , the following equations must be satisfied:

$$\begin{cases} \int_{\Omega} \psi Q \, dv + \int_{\partial\Omega} \psi (\underline{\lambda} \cdot \overrightarrow{\text{grad}}(T)) \cdot \overrightarrow{n} \, ds - \int_{\Omega} \overrightarrow{\text{grad}}^T(\psi) \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(T) \, dv = 0 \\ (\underline{\lambda} \cdot \overrightarrow{\text{grad}}(T)) \cdot \overrightarrow{n} = q \quad \text{on } \partial\Omega_q \\ T = T_d \quad \text{on } \partial\Omega_T \end{cases}$$

We see that no more second derivatives of the temperature field appear in the problem formulation. The regularity required on T to give a sense to the problem has therefore been lowered; that is the reason why it is called a weak integral formulation. On the contrary, the regularity required for functions ψ has increased since first derivatives of these functions have appeared. The same regularity is required for functions of E_T and E_ψ . A detailed mathematical analysis, which will not be performed here, makes it possible to demonstrate that these regularity conditions correspond to Sobolev's first space $H^1(\Omega)$, which gathers the functions square integrable on Ω and whose first derivatives are also square integrable on Ω .

An important step toward the weak integral formulation of the problem consists of inserting the boundary conditions on $\partial\Omega_T$ into the space of admissible functions E_T . This is done by defining:

$$E_T = \{T \in H^1(\Omega) / T = T_d \text{ on } \partial\Omega_T\} \quad [1.11]$$

Dirichlet boundary conditions (prescribed temperature) are therefore directly inserted into the solution space. These conditions are said to be essential for the problem formulation.

Besides, we select weighting functions canceling one another on $\partial\Omega_T$. Thus, these functions ψ can be considered to be admissible variations of functions of E_T . Indeed, if $T \in E_T$, then for any ψ of E_ψ , $T + \psi \in E_T$. The following space E_ψ is then defined:

$$E_\psi = \{\psi \in H^1(\Omega) / \psi = 0 \text{ on } \partial\Omega_T\} \quad [1.12]$$

Taking into consideration the flux boundary conditions (on $\partial\Omega_q$) and the previous definition of E_ψ , the surface integral term appearing when integration by parts is applied becomes:

$$\begin{aligned} \int_{\partial\Omega} \psi (\underline{\lambda} \cdot \overrightarrow{\text{grad}}(T)) \cdot \overrightarrow{n} \, ds &= \int_{\partial\Omega_T} \psi (\underline{\lambda} \cdot \overrightarrow{\text{grad}}(T)) \cdot \overrightarrow{n} \, ds \\ &+ \int_{\partial\Omega_q} \psi (\underline{\lambda} \cdot \overrightarrow{\text{grad}}(T)) \cdot \overrightarrow{n} \, ds = \int_{\partial\Omega_q} \psi q \, ds \end{aligned}$$

Now problem [1.10] to be solved can be written as follows:

$$\left| \begin{array}{l} \text{Find } T \in E_T \text{ (defined by equation [1.11]) such that} \\ \text{for any } \psi \in E_\psi \text{ (defined by equation [1.12]):} \\ \int_{\Omega} \psi Q \, dv + \int_{\partial\Omega_q} \psi q \, ds - \int_{\Omega} \overrightarrow{\text{grad}}^T(\psi) \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(T) \, dv = 0 \end{array} \right. \quad [1.13]$$

The flux boundary conditions are included in the function to cancel in problem [1.13]. They are said to be natural boundary conditions for the formulation.

Problem [1.13] defined in this way now offers a continuous mathematical framework for seeking a solution. It can be demonstrated that it is equivalent to boundary value problem [1.8] but is far more appropriate to seek approximate solutions. Indeed the only obligation required for the solution $T(\vec{x})$ is to belong to the space E_T , i.e. to belong to $H^1(\Omega)$ (square integrable function whose first derivatives are square integrable), and to satisfy the essential boundary conditions $T = T_d$ on $\partial\Omega_T$.

As weighting functions ψ of E_ψ can be assimilated to admissible variations of functions E_T , it is easy to demonstrate that, in linear cases ($\underline{\lambda}$ independent of T , Q and q linear according to T), problem [1.13] is equivalent to the minimization problem of a functional $\Pi(T)$. In particular, when Q and q are independent of T , this functional equation is written as follows:

$$\left| \begin{array}{l} \text{Find } T \in E_T \text{ minimizing the functional equation:} \\ \Pi(T) = \frac{1}{2} \int_{\Omega} \overrightarrow{\text{grad}}^T(T) \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(T) \, dv - \int_{\Omega} T Q \, dv - \int_{\partial\Omega_q} T q \, ds \end{array} \right. \quad [1.14]$$

Minimizing Π then comes down to finding T of E_T such that, for any admissible variation δT in E_ψ , the corresponding variation $\delta\Pi$ of Π vanishes. This is expressed as follows:

$$\left| \begin{array}{l} \text{For any } \delta T \in E_\psi: \\ \delta\Pi = \int_{\Omega} \overrightarrow{\text{grad}}^T(\delta T) \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(T) \, dv - \int_{\Omega} \delta T Q \, dv - \int_{\partial\Omega_q} \delta T q \, ds = 0 \end{array} \right.$$

We then go back to problem [1.13].

Approximation methods consist of constructing an approximate solution of problem [1.13] by working in the sub-spaces E_T^n and E_ψ^n of E_T and E_ψ of finite

dimension n . By constructing E_ψ , a particular element T^* of E_T makes it possible to express all the elements of this space in the form $T = T^* + \psi$, where ψ is an element of E_ψ . The Galerkin method [GAL 15] makes use of this property when it considers a space E_T^n defined from the space E_ψ^n as follows:

$$E_T^n = \{T/T = T^* + \psi \text{ with } \psi \in E_\psi^n\}$$

The Galerkin method then consists of replacing continuous problem [1.13] with the following discrete problem (of dimension n):

$$\left| \begin{array}{l} \text{Find } T \in E_T^n \text{ such that for any } \psi \in E_\psi^n: \\ \int_{\Omega} \psi Q dv + \int_{\partial\Omega_q} \psi q(T) ds - \int_{\Omega} \overrightarrow{\text{grad}}^T(\psi) \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(T) dv = 0 \end{array} \right. \quad [1.15]$$

It is obvious that if we increase the dimension n of spaces E_T^n and E_ψ^n , it is possible to construct a series of approximate solutions. It is demonstrated that this series converges to the solution of continuous problem [1.13] when n tends to infinity.

1.3. Working example

In this section, we will deal with the example of a metal plate heated by induction. The aim is to illustrate the application of physical modeling and mathematical analysis processes used in the previous sections to a real case. Physical modeling will allow us to specify the assumptions made and the equations used. Then we will examine the various solution methods:

- direct methods based upon partial differential equations and boundary conditions resulting from physical modeling (problem [1.8]). Here we will restrict our study to the analytical integration model and its approximation by the finite element method;
- collocation methods based upon problem [1.10] obtained by the weighted residual method. We will restrict our study to point and sub-domain collocation methods;
- the Galerkin method resulting from the weak integral form of problem [1.13]. We will study a polynomial approximation of the temperature field then a piecewise linear approximation of this field (finite element method).

1.3.1. Physical modeling

The work example considered in this chapter is that of an induction heated plate. It is illustrated in Figure 1.2. Its physical modeling can be performed at different levels. In this chapter, we will restrict our study to a 1D thermal analysis of the plate thickness. A line of length E (the plate thickness) represents the solid.

Heat conduction is only considered in the direction x . The corresponding thermal conductivity is written λ and assumed to be constant.

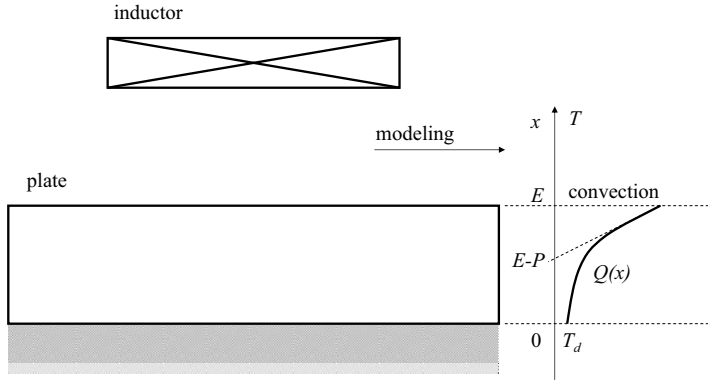


Figure 1.2. Work example: induction heated plate

On the bottom part of the plate ($x = 0$), the temperature is assumed to be known and equal to a constant value T_d . On the top part of the plate, a constant heat exchange coefficient H is used to represent the plate convection exchanges with the external environment, whose temperature is equal to $T_{\text{ext}} = T_d$. Induction heating is represented by the following heat volumetric density:

$$Q = Q(x) = Q_0 e^{\frac{x-E}{P}}$$

where Q_0 corresponds to the heating volumetric power (in W/m^3) at $x = E$, and in which P is a characteristic heating depth (depending in particular on the inductor frequency). Low frequency heating for instance will be represented by a high value of P , and will therefore tend to heat the solid “in its volume”. On the contrary, high frequency heating will concentrate the power on the plate surface (skin effect). This effect is used by the induction hardening process. A graphic representation of the temperature field obtained by various methods is given in this section. To do so, we have used the following values for the parameters:

- plate thickness: $E = 0.1 \text{ m}$;
- material thermal conductivity: $\lambda = 30 \text{ W/m/K}$;
- heat exchange coefficient at $x = E$: $H = 2,000 \text{ W/m}^2/\text{K}$;
- temperature at $x = 0$ and of the external environment: $T_d = 20^\circ\text{C}$;
- heating parameters: $Q_0 = 5 \cdot 10^7 \text{ W/m}^3$ and $P = 0.02 \text{ m}$.

1.3.2. Direct methods

The thermal equilibrium in the plate is governed by problem [1.8] which, in this case, is written as follows:

$$\begin{cases} \lambda \frac{d^2 T}{dx^2} + Q(x) = 0 & \text{for } 0 \leq x \leq E \\ \lambda \frac{dT}{dx} = H(T_d - T(x)) & \text{at } x = E \\ T(x) = T_d & \text{at } x = 0 \end{cases}$$

1.3.2.1. Analytical integration

It is possible to integrate this differential equation analytically. We then obtain the exact solution of the problem, i.e. the temperature field $T(x)$ which satisfies the thermal equilibrium and the boundary conditions:

$$T(x) = T_d + \frac{Q_0 P}{\lambda + EH} \left(1 + \frac{PH}{\lambda} (1 - e^{-\frac{E}{P}}) \right) x + \frac{Q_0 P^2}{\lambda} e^{-\frac{E}{P}} (1 - e^{-\frac{x}{P}})$$

The temperature field obtained is illustrated in Figure 1.3. It is obvious that, since an analytical solution exists in this working example, it is preferable to any other. However, in order to illustrate approximation methods, we are going to apply them to this example, then compare the solutions obtained with this analytical solution.

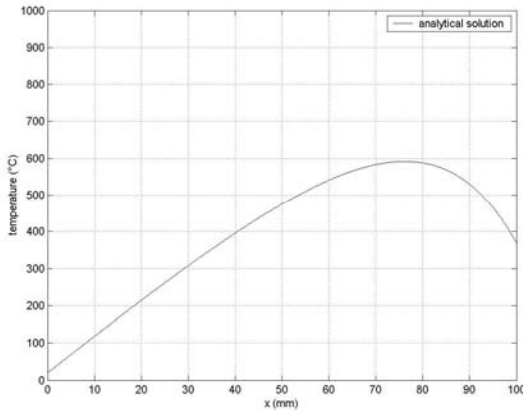


Figure 1.3. Working example: analytical solution

1.3.2.2. The finite difference method

To obtain an approximate solution of the initial differential equation, we discretize axis Ox between $x = 0$ and $x = E$ by defining $n + 1$ equidistant points with x-axis $x_i = ih$ (with $h = E/n$ and $i = 0, 1, \dots, n$). Then, with T_i being the temperature at the x-axis point x_i , an approximation of T second derivatives at the solid internal points $i = 1, \dots, n - 1$ can be obtained as follows:

$$\left. \frac{d^2 T}{dx^2} \right|_{x=x_i} = \frac{T_{i+1} - 2T_i + T_{i-1}}{h^2}$$

Finally, the temperatures at the boundary points with x-axes $x = x_0 = 0$ and $x = x_n = E$ are obtained by applying the boundary conditions. At $x = 0$, the temperature T_0 is known, its value is T_d . At $x = E$, the incoming flux can be expressed as $\lambda(T_n - T_{n-1})/e$, which makes it possible to write:

$$\left(1 + \frac{hH}{\lambda}\right)T_n - T_{n-1} = \frac{hH}{\lambda}T_d$$

The problem to solve is thus expressed in the form of a linear system whose unknowns are the temperatures T_i :

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 0 \\ 0 & 0 & 0 & \cdots & 2 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 1 + \frac{hH}{\lambda} \end{bmatrix} \begin{pmatrix} T_0 \\ T_1 \\ T_2 \\ \vdots \\ T_{n-2} \\ T_{n-1} \\ T_n \end{pmatrix} = \begin{pmatrix} T_d \\ \frac{h^2}{\lambda}Q(x_1) \\ \frac{h^2}{\lambda}Q(x_2) \\ \vdots \\ \frac{h^2}{\lambda}Q(x_{n-2}) \\ \frac{h^2}{\lambda}Q(x_{n-1}) \\ \frac{hH}{\lambda}T_d \end{pmatrix}$$

The solution of this linear system produces an approximate solution of problem [1.8] in our working example. This approximate solution is given by a discrete set of values T_i . Each T_i corresponds to an approximate value of the temperature at the point with x-axis x_i . It should be noted that the finite difference method gives a solution only for points with x-axis x_i . It gives no *a priori* indication of the form of the solution beyond these points. In practice, it is advised to construct an approximate solution on the whole domain with a linear interpolation of the solution on each interval.

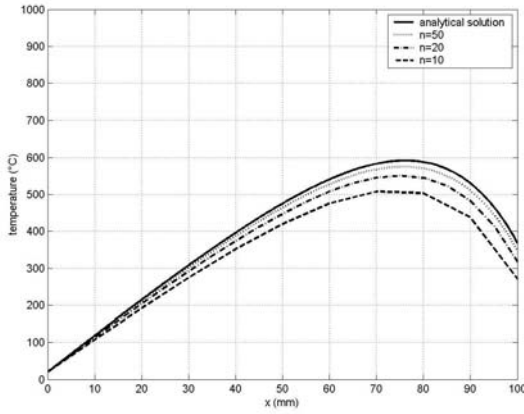


Figure 1.4. Working example: the finite difference method

Figure 1.4 gives the approximate solutions obtained by finite differences with a number n of discretization segments equal to 10, 20 and 50. It is established that, when n increases, the approximate solution tends slowly to the analytical solution.

The finite difference method is relatively easy to implement in cases with simple geometries (1D, 2D or 3D). For instance, in the 3D case, the calculation points are distributed on a Cartesian grid whose main directions are axes Ox , Oy and Oz . However, the problem is far more complicated when the grid does not make it possible to give a good representation of the geometry or write the boundary conditions correctly. In that case, curvilinear finite difference methods can be used.

1.3.3. Collocation methods

Collocation methods are based upon problem [1.10] resulting from problem [1.8] by applying the weighted residual method. In this working example, problem [1.10] is expressed as follows:

$$\left\{ \begin{array}{l} \text{Find } T \in E_T \text{ such that for any } \psi \in E_\psi: \\ \int_0^E \psi \left(\lambda \frac{d^2 T}{dx^2} + Q(x) \right) dx = 0 \\ \lambda \frac{dT}{dx} = H(T_d - T) \quad \text{at } x = E \\ T = T_d \quad \text{at } x = 0 \end{array} \right.$$

1.3.3.1. Point collocation

In order to obtain an approximate solution of the initial differential problem, we discretize axis Ox between $x = 0$ and $x = E$ by defining $n + 1$ equidistant points with x-axis $x_i = ih$ (with $h = E/n$ and $i = 0, \dots, n$), and for E_ψ , we select the set of Dirac distributions corresponding to those points. The solution to the problem then comes down to seeking a function $T(x)$ which can be differentiated twice, satisfying the boundary conditions and canceling the residual at each point x_i . This produces the $n + 3$ equations:

$$\begin{cases} \lambda \frac{d^2 T}{dx^2} + Q(x) = 0 & \text{at } x = x_0, \dots, x_n \\ \lambda \frac{dT}{dx} = H(T_d - T) & \text{at } x = x_n \\ T = T_d & \text{at } x = x_0 \end{cases}$$

For instance, choose for E_T the space of $n + 2$ degree polynomials. The function T will then be expressed as $T(x) = a_0 + a_1x + a_2x^2 + \dots + a_{n+2}x^{n+2}$, with $n + 3$ coefficients to determine. We obtain a linear system in a_i , of $n + 3$ equations with $n + 3$ unknowns:

$$\begin{bmatrix} 0 & 0 & \dots & (n+2)(n+1)x_0^n \\ 0 & 0 & \dots & (n+2)(n+1)x_1^n \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & (n+2)(n+1)x_n^n \\ \frac{H}{\lambda} & 1 + \frac{H}{\lambda}x_n & \dots & \left(n+2 + \frac{H}{\lambda}x_n\right)x_n^{n+1} \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \\ a_{n+1} \\ a_{n+2} \end{pmatrix} = \begin{pmatrix} -\frac{Q(x_0)}{\lambda} \\ -\frac{Q(x_1)}{\lambda} \\ \vdots \\ -\frac{Q(x_n)}{\lambda} \\ \frac{H}{\lambda}T_d \\ T_d \end{pmatrix}$$

The solution of this linear system yields coefficients a_i of the polynomial used as the approximation of the function $T(x)$ sought. Figure 1.5 gives the approximations obtained with a number of collocation points equal to 3 ($n = 2$), 4 ($n = 3$) and 6 ($n = 5$). We can see that with only 6 points, the solution obtained practically coincides with the analytical solution in this case. However, in this case a 7-degree polynomial is necessary. In 2D or 3D, the number of coefficients to determine would soon become very high, making this situation ill-suited to the linear system solution.

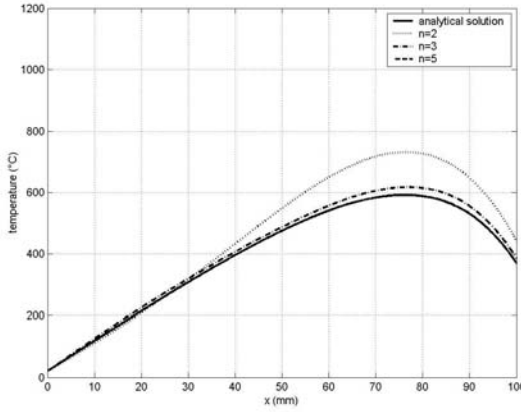


Figure 1.5. Working example: point collocation

1.3.3.2. Sub-domain collocation

Axis Ox is divided into n control volumes. These are segments of length $h = E/n$. They are limited by points $x_{i-1} = (i-1)h$ and $x_i = ih$ (with $i = 1, \dots, n$). If we make T_i the temperature (assumed to be constant) in the segment $[x_{i-1}, x_i]$ it is possible to calculate the first and second derivatives with a finite difference scheme between the segment centers. We then obtain a linear system on the temperatures T_i of each of the n element segments:

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 2 & \cdots & 0 & 0 \\ 0 & -1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 2 & -1 \\ 0 & 0 & \cdots & -1 & 1 + \frac{hH}{\lambda} \end{bmatrix} \begin{pmatrix} T_1 \\ T_2 \\ T_3 \\ \cdots \\ T_{n-1} \\ T_n \end{pmatrix} = \begin{pmatrix} T_d \\ \frac{h}{\lambda} \int_{x_1}^{x_2} Q(x) dx \\ \frac{h}{\lambda} \int_{x_2}^{x_3} Q(x) dx \\ \cdots \\ \frac{h}{\lambda} \int_{x_{n-2}}^{x_{n-1}} Q(x) dx \\ \frac{hH}{\lambda} T_d \end{pmatrix}$$

Figure 1.6 illustrates the sub-domain collocation method applied to our working example, with respectively 20, 50 and 100 segments. It should be noted that the approximate solution tends to the analytical solution when the number of segments

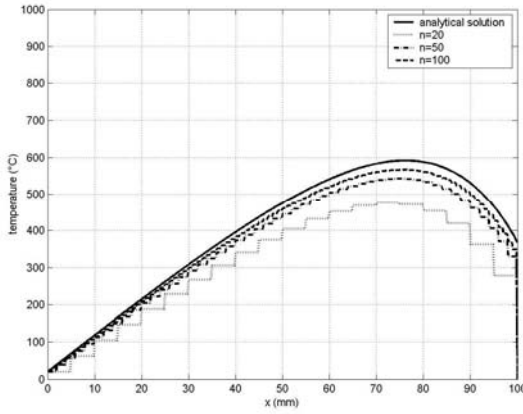


Figure 1.6. Working example: sub-domain collocation

increases. This method is the basis of the finite volume method, which is widely used in thermal science. It combines a simple formulation, like finite differences, with the very physical notion of thermal equilibrium per element volume.

1.3.4. Galerkin method

The Galerkin method is based upon the weak integral form [1.13] of the problem to solve. In this work example, problem [1.13] is expressed as follows:

$$\left| \begin{array}{l} \text{Find } T \in E_T \text{ such that for any } \psi \in E_\psi: \\ \int_0^E \psi(x)Q(x)dx + H\psi(E)(T_d - T(E)) - \lambda \int_0^E \frac{d\psi}{dx} \frac{dT}{dx} dx = 0 \end{array} \right.$$

Spaces E_T and E_ψ are defined as follows:

$$\begin{aligned} E_\psi &= \{ \psi \in H^1([0, E]) / \psi(0) = 0 \} \\ E_T &= \{ T \in H^1([0, E]) / T(0) = T_d \} \end{aligned}$$

The Galerkin method consists of selecting finite spaces E_ψ^n and E_T^n , of size n , instead of spaces E_ψ and E_T .

1.3.4.1. Polynomial functions

Here we will consider space E_ψ^n generated by the n monomials x^i ($i = 1, \dots, n$). In order to satisfy the condition $\psi = 0$ on the boundary $\partial\Omega_T$ (here $x = 0$), the

elements ψ of E_ψ^n will be expressed with n parameters as follows:

$$\psi(x) = \sum_{i=1}^{i=n} b_i x^i$$

The quantity to cancel in the discrete problem then becomes:

$$\sum_{i=1}^{i=n} b_i \left(\int_0^E x^i Q(x) dx + H E^i (T_d - T(E)) - \lambda \int_0^E i x^{i-1} \frac{dT}{dx} dx \right) = 0$$

As this quantity must be equal to zero for any function ψ of E_ψ^n , i.e. for any set of components b_i , the n quantities between brackets, functions of T , must be equal to zero. This produces n equations in T .

To generate the space E_T^n of admissible functions, the temperature field $T(x)$ is expressed as the sum of a particular admissible function $T_0(x) = T_d$ and an element E_ψ^n . Therefore, the temperature field is sought in the following form, where a_i represents the n parameters to determine:

$$T(x) = T_d + \sum_{j=1}^n a_j x^j$$

The n equations finally become the linear functions of the n unknowns a_i in the form:

$$\sum_{j=1}^n \left(H E^{i+j} + \lambda \int_0^E i j x^{i+j-2} dx \right) a_j = \int_0^E x^i Q(x) dx$$

Figure 1.7 illustrates the temperature field obtained with this method for 2-, 3- and 4-degree polynomials. We can see that the approximate solution converges toward the analytical solution when the polynomial degree increases.

The major advantage of this method is that the matrix and the load vector of the system to solve have general expressions, added to which, the matrix is symmetric. Conversely, these quantities include integral equations whose analytical estimation may be complex. Their numerical estimation is certainly possible, but generates new approximations. In our example, the integral equations have been estimated from their analytical forms. Moreover, the matrix is full. Applying this method to 2D or 3D cases presupposes that one is capable of constructing a polynomial approximation satisfying the essential boundary conditions, which does not happen very frequently in practice.

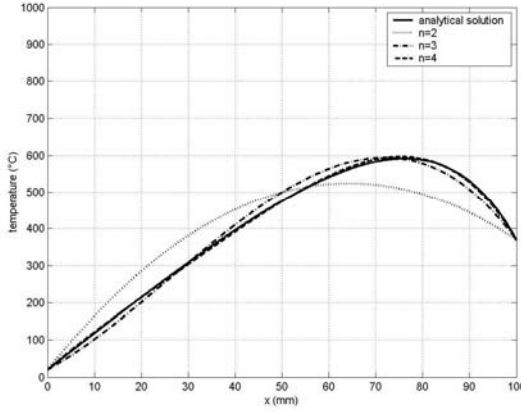


Figure 1.7. Working example: polynomial approximation

1.3.4.2. Piecewise linear functions

Starting with the discretization of segment $[0, E]$ performed from $n + 1$ points, with x-axis $x_i = ih$ with $h = E/n$ and $i = 0, \dots, n$, we can consider for E_ψ^n the space of continuous functions varying linearly on each interval $[x_i, x_{i+1}]$. It is easy to demonstrate that this space is generated by the $n + 1$ following functions $N_i(x)$ (this is expressed as $x_{-1} = -h$ and $x_{n+1} = E + h$):

$$N_i(x) = \begin{cases} \frac{x - x_{i-1}}{h} & \text{if } x_{i-1} \leq x \leq x_i \\ \frac{x_{i+1} - x}{h} & \text{if } x_i \leq x \leq x_{i+1} \\ 0 & \text{if not} \end{cases}$$

In order to satisfy the condition $\psi = 0$, at $x = 0$, the functions of space E_ψ^n will then be expressed as follows:

$$\psi(x) = \sum_{i=1}^n N_i(x) \psi_i$$

Space E_ψ^n is therefore of the dimension n . A function ψ of this space will be characterized by its n parameters $\psi_1, \psi_2, \dots, \psi_n$, corresponding to the values taken by the function respectively at the points x_1, x_2, \dots, x_n . Between these points, the function ψ is obtained by linear interpolation. The quantity to cancel in the discrete

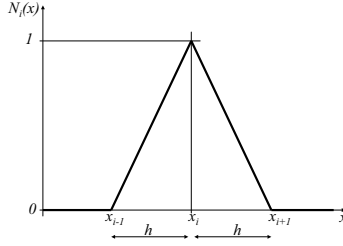


Figure 1.8. *Generating piecewise linear functions*

problem now becomes:

$$\sum_{i=1}^{i=n} \psi_i \left(\int_0^E N_i(x) Q(x) dx + H N_i(E) (T_d - T(E)) - \lambda \int_0^E \frac{dN_i}{dx} \frac{dT}{dx} dx \right) = 0$$

As this quantity must be equal to zero for any function ψ , and hence for any set of components ψ_i , the term between brackets in this equation must be equal to zero for any index i .

We now introduce a particular admissible solution $T^*(x) = N_0(x)T_d$ to express any function of space E_T^n as follows:

$$T(x) = N_0(x)T_d + \sum_{i=1}^n N_i(x)T_i$$

The approximate solution selected is then a piecewise linear function. Its value is T_d at $x = x_0 = 0$, T_i at $x = x_i$ (for $i = 1, \dots, n$), and is obtained by linear interpolation between two consecutive points. When such a function is inserted into the quantity to cancel, we finally obtain a linear system of the following type:

$$\sum_{j=1}^{j=n} K_{ij} T_j = F_i$$

whose unknowns are temperatures T_1, \dots, T_n . The terms K_{ij} of the matrix and F_i of the load vector of this linear system are expressed analytically as follows:

$$\begin{cases} K_{ij} = \lambda \int_0^E \frac{dN_i}{dx} \frac{dN_j}{dx} dx + H N_i(E) N_j(E) \\ F_i = \int_0^E \left(N_i(x) Q(x) - \lambda \frac{dN_i}{dx} \frac{dN_0}{dx} T_d \right) dx + H N_i(x) N_0(x) T_d \end{cases}$$

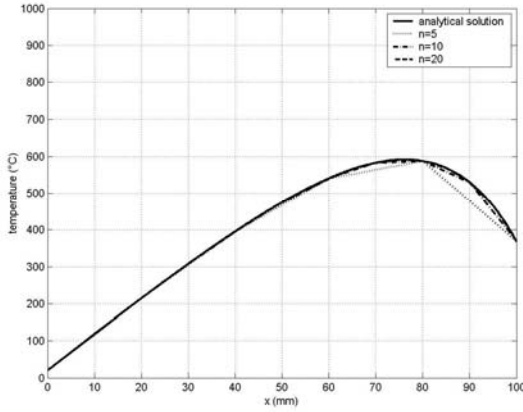


Figure 1.9. Working example: piecewise linear approximation

Figure 1.9 illustrates the approximate solution in the case of 5, 10 and 20 segments distributed on axis Ox (between $x = 0$ and $x = E$). We can see that this solution tends rapidly toward the analytical solution when the number of discretization segments increases.

The approximation used here is a finite element approximation. It has the same advantages as a polynomial approximation (symmetric matrix, general and specific expression of the matrix and load vector), but has none of the drawbacks mentioned before. Indeed the calculation of integral equations included in the linear system is much easier. Specifically, the integrations involved in the expression of the terms K_{ij} and F_i can be restricted to the only fields in which the function to integrate is not equal to zero. For instance, for the diagonal terms of the matrix $[K_{ij}]$ we obtain:

$$K_{ii} = \lambda \int_{x_{i-1}}^{x_{i+1}} \frac{dN_i}{dx}^2 dx + HN_i^2(E)$$

Moreover, it should be noted here that matrix $[K_{ij}]$ is sparse.

In this example, the integral equations have been calculated in an analytical way. We will see later on in this book that effective numerical methods make it possible to estimate these integral equations accurately and rapidly.

This page intentionally left blank

Chapter 2

The Finite Element Method

2.1. Finite element approximation

Finite element approximation consists of dividing (or meshing) the geometric domain of study into element domains or finite elements. The approximation is built upon each element domain from the function values at certain specific points, called nodes, while ensuring certain regularity conditions of the global approximation on the whole domain. This type of approximation is also called sub-domain nodal approximation. If the division, or meshing, performed includes a total of n_t nodes, then the problem includes n_t unknowns representing the values taken by the function $T(\vec{x})$ at each node.

2.1.1. *Mesh*

Divide (discretize) a solid Ω into m element domains, termed Ω^e , called finite elements and satisfying the following two conditions:

- the union of the m element domains Ω^e makes up the solid Ω ;
- the intersection of two different element domains is void.

The major advantage of this type of discretization is that each element domain Ω^e can be expressed in a simple geometric form (2D triangle or quadrangle, 3D tetrahedron, hexahedron, etc.). These domains are interconnected by nodes. The set of these element domains is called a mesh of the domain Ω . In this book, the total number of mesh nodes will be written n_t .

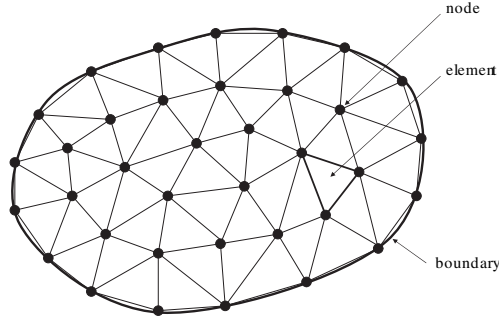


Figure 2.1. Mesh example of a domain Ω

Figure 2.1 illustrates the example of a mesh constructed on a 2D domain Ω . In that figure we can see that the triangular elements used give a polyhedric approximation of the domain boundary. To improve that approximation, we can, for example, increase the number of elements. In the next chapter, we will see that there are also elements with curved boundaries, making it possible to obtain a very good approximation of the boundary of the domain studied.

In a mesh, the nodes and elements are numbered. The n_t nodes included in the mesh are numbered sequentially from 1 to n_t , following their order of creation, during discretization for instance. This numbering process is called global numbering. Then we will designate:

- $\{T\}$ the vector including the n_t temperatures at the mesh nodes in a column;
- $\langle\psi\rangle$ the vector including the n_t weighting function values at the mesh nodes on a line.

For each element e we now write n^e the number of nodes corresponding to it. These nodes are numbered locally from 1 to n^e . Thus:

- $\{T^e\}$ will be the vector including in a column the n^e node temperatures of the element e ,
- $\langle\psi^e\rangle$ will be the vector including on a line the n^e weighting function values at the nodes of the element e .

The global and local numbering methods of the mesh nodes are illustrated in Figure 2.2. They must now be interconnected. This is done by means of a matrix, $[A^e]$, which makes it possible to move from the vectors $\{T\}$ or $\langle\psi\rangle$ (of dimension n_t) to the vectors $\{T^e\}$ or $\langle\psi^e\rangle$ (of dimension n^e):

$$\{T^e\} = [A^e] \cdot \{T\} \quad \text{and} \quad \langle\psi^e\rangle = \langle\psi\rangle \cdot [A^e]^T \quad [2.1]$$

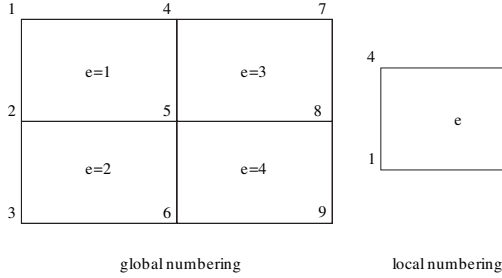


Figure 2.2. *Numbering example in a mesh*

The matrix $[A^e]$ has n^e lines and n_t columns. Each term A_{ij}^e equals 1 if the node numbered (globally) j on the whole structure coincides with the node numbered (locally) i on element e and 0 otherwise. In Figure 2.2, the previous equation is written as follows for the node temperatures of element 1:

$$\underbrace{\begin{Bmatrix} T_1^1 \\ T_2^1 \\ T_3^1 \\ T_4^1 \end{Bmatrix}}_{\{T^1\}} = \underbrace{\begin{bmatrix} 010000000 \\ 000010000 \\ 000100000 \\ 100000000 \end{bmatrix}}_{[A^1]} \cdot \underbrace{\begin{Bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \\ T_6 \\ T_7 \\ T_8 \\ T_9 \end{Bmatrix}}_{\{T\}}$$

Likewise, the element node weighting functions are expressed as follows:

$$\underbrace{\langle \psi_1^1 \quad \psi_2^1 \quad \psi_3^1 \quad \psi_4^1 \rangle}_{\langle \psi^1 \rangle} = \underbrace{\langle \psi_1 \quad \psi_2 \quad \cdots \quad \psi_8 \quad \psi_9 \rangle}_{\langle \psi \rangle} \cdot \underbrace{\begin{bmatrix} 0001 \\ 1000 \\ 0000 \\ 0010 \\ 0100 \\ 0000 \\ 0000 \\ 0000 \\ 0000 \end{bmatrix}}_{[A^1]^T}$$

2.1.2. Nodal approximation

The quantities involved in the problem to be solved are now estimated in each element e from the values they take at the nodes corresponding to that element. If we define $\overline{\Omega}^e = \Omega^e \cup \partial\Omega^e$ (inside of the element plus its boundary), we now write for each element e , $T(\vec{x})$ and $\psi(\vec{x})$ in the form:

$$\begin{aligned} \forall \vec{x} \in \overline{\Omega}^e, \quad T(\vec{x}) &= \sum_{i=1}^{n^e} N_i^e(\vec{x}) T_i^e = \langle N_1^e(\vec{x}), \dots, N_{n^e}^e(\vec{x}) \rangle \cdot \begin{Bmatrix} T_1^e \\ \vdots \\ T_{n^e}^e \end{Bmatrix} \\ &= \langle N^e(\vec{x}) \rangle \cdot \{T^e\} \end{aligned} \quad [2.2]$$

$$\begin{aligned} \forall \vec{x} \in \overline{\Omega}^e, \quad \psi(\vec{x}) &= \sum_{i=1}^{n^e} \psi_i^e N_i^e(\vec{x}) = \langle \psi_1^e, \dots, \psi_{n^e}^e \rangle \cdot \begin{Bmatrix} N_1^e(\vec{x}) \\ \vdots \\ N_{n^e}^e(\vec{x}) \end{Bmatrix} \\ &= \langle \psi^e \rangle \cdot \{N^e(\vec{x})\} \end{aligned} \quad [2.3]$$

In equations [2.2] and [2.3], T_i^e and ψ_i^e respectively represent the temperature and the value of the weighting function at node number i (local numbering) of element e .

The functions $N_i^e(\vec{x})$ introduced are known functions, making it possible to interpolate the temperature in element e from the node values. N_i^e is the interpolation function, or basis function, or shape function corresponding to the node i (local numbering) of element e , $\{N^e\}$ is the vector of dimension n^e including in a column all the functions corresponding to element e , and $\langle N^e \rangle$ is the vector including them on a line.

Note that it is equally possible to write $\psi(\vec{x}) = \langle N^e(\vec{x}) \rangle \cdot \{\psi^e\}$ or $\psi(\vec{x}) = \langle \psi^e \rangle \cdot \{N^e(\vec{x})\}$. The choice of the second form is prompted by the fact that the parameters corresponding to the weighting functions will then be written as a line vector $\langle \psi \rangle$.

For instance, in the case of element 1 in Figure 2.2, it will be possible to use the following expressions to approximate the temperature and weighting functions in element 1:

$$T(x, y) = \underbrace{\langle N_1^1(x, y) \quad N_2^1(x, y) \quad N_3^1(x, y) \quad N_4^1(x, y) \rangle}_{\langle N^1 \rangle} \cdot \underbrace{\begin{Bmatrix} T_1^1 \\ T_2^1 \\ T_3^1 \\ T_4^1 \end{Bmatrix}}_{\{T^1\}}$$

$$\psi(x, y) = \underbrace{\langle \psi_1^1 \quad \psi_2^1 \quad \psi_3^1 \quad \psi_4^1 \rangle}_{\langle \psi^1 \rangle} \cdot \underbrace{\begin{Bmatrix} N_1^1(x, y) \\ N_2^1(x, y) \\ N_3^1(x, y) \\ N_4^1(x, y) \end{Bmatrix}}_{\{N^1\}}$$

It is obvious that the nodal approximations defined by equations [2.2] and [2.3] are continuously differentiable inside each element e (as long as the interpolation functions are). In order to generate continuous approximations in the whole domain Ω we make use of shape functions satisfying the following two continuity conditions:

- 1) if \vec{x}_j is the position of the node j of element e , then $N_i^e(\vec{x}_j) = 1$ if $i = j$ and 0 otherwise;
- 2) for any point indexed \vec{x} located on a face or an edge of element e which does not include node i , then $N_i^e(\vec{x}) = 0$.

An approximation with shape functions satisfying these conditions is said to be a conform approximation. It makes it possible to generate continuous functions in the whole domain Ω . We demonstrate that, if the functions N_i^e are differentiable in Ω^e , these conditions are sufficient to allow the approximation to belong to space $H^1(\Omega)$. More regular approximations can be constructed, for example with conditions on the derivative connections between the elements. These will not be dealt with in this book.

It is important to note at this point that nodal approximation makes it possible not only to estimate a function in an element from its node values, but also to calculate its gradient, thus:

$$\begin{aligned} \forall \vec{x} \in \Omega^e, \quad \overrightarrow{\text{grad}}(T(\vec{x})) &= \sum_{i=1}^{n^e} \overrightarrow{\text{grad}}(N_i^e(\vec{x})) T_i^e \\ &= \left\langle \overrightarrow{\text{grad}}(N_1^e), \dots, \overrightarrow{\text{grad}}(N_{n^e}^e) \right\rangle \cdot \begin{Bmatrix} T_1^e \\ \vdots \\ T_{n^e}^e \end{Bmatrix} \quad [2.4] \\ &= \left\langle \overrightarrow{\text{grad}}(N^e(\vec{x})) \right\rangle \cdot \{T^e\} \end{aligned}$$

$$\forall \vec{x} \in \Omega^e, \quad \overrightarrow{\text{grad}}^T(\psi(\vec{x})) = \sum_{i=1}^{n^e} \psi_i^e \overrightarrow{\text{grad}}^T(N_i^e(\vec{x}))$$

$$\begin{aligned}
&= \langle \psi_1^e, \dots, \psi_{n^e}^e \rangle \cdot \begin{Bmatrix} \overrightarrow{\text{grad}}^T(N_1^e) \\ \vdots \\ \overrightarrow{\text{grad}}^T(N_{n^e}^e) \end{Bmatrix} \\
&= \langle \psi^e \rangle \cdot \left\{ \overrightarrow{\text{grad}}^T(N^e(\vec{x})) \right\} \tag{2.5}
\end{aligned}$$

In equation [2.4], $\langle \overrightarrow{\text{grad}}(N^e) \rangle$ is a vector including n^e terms on a line. These terms are gradients $\overrightarrow{\text{grad}}(N_i^e)$, with a number of terms equal to the dimension of the problem dealt with. In equation [2.5], $\left\{ \overrightarrow{\text{grad}}^T(N^e) \right\}$ is a vector including n^e terms in a column. These terms are transposed gradients $\overrightarrow{\text{grad}}^T(N_i^e)$, with a number of terms equal to the dimension of the problem dealt with. For example, with a 2D element in Figure 2.2, if we consider a Cartesian coordinate system (x, y) , it will be possible to estimate the temperature gradient and the weighting function gradient in element 1 as follows:

$$\begin{aligned}
\underbrace{\begin{pmatrix} \frac{\partial T}{\partial x} \\ \frac{\partial T}{\partial y} \end{pmatrix}}_{\overrightarrow{\text{grad}}(T)} &= \underbrace{\left\langle \begin{pmatrix} \frac{\partial N_1^1}{\partial x} \\ \frac{\partial N_1^1}{\partial y} \end{pmatrix} \quad \begin{pmatrix} \frac{\partial N_2^1}{\partial x} \\ \frac{\partial N_2^1}{\partial y} \end{pmatrix} \quad \begin{pmatrix} \frac{\partial N_3^1}{\partial x} \\ \frac{\partial N_3^1}{\partial y} \end{pmatrix} \quad \begin{pmatrix} \frac{\partial N_4^1}{\partial x} \\ \frac{\partial N_4^1}{\partial y} \end{pmatrix} \right\rangle}_{\langle \overrightarrow{\text{grad}}(N^1) \rangle} \cdot \underbrace{\begin{Bmatrix} T_1^1 \\ T_2^1 \\ T_3^1 \\ T_4^1 \end{Bmatrix}}_{\{T^1\}} \\
\underbrace{\begin{pmatrix} \frac{\partial \psi}{\partial x} & \frac{\partial \psi}{\partial y} \end{pmatrix}}_{\overrightarrow{\text{grad}}^T(\psi)} &= \underbrace{\langle \psi_1^1 \quad \psi_2^1 \quad \psi_3^1 \quad \psi_4^1 \rangle}_{\langle \psi^1 \rangle} \cdot \underbrace{\begin{Bmatrix} \begin{pmatrix} \frac{\partial N_1^1}{\partial x} & \frac{\partial N_1^1}{\partial y} \end{pmatrix} \\ \begin{pmatrix} \frac{\partial N_2^1}{\partial x} & \frac{\partial N_2^1}{\partial y} \end{pmatrix} \\ \begin{pmatrix} \frac{\partial N_3^1}{\partial x} & \frac{\partial N_3^1}{\partial y} \end{pmatrix} \\ \begin{pmatrix} \frac{\partial N_4^1}{\partial x} & \frac{\partial N_4^1}{\partial y} \end{pmatrix} \end{Bmatrix}}_{\left\{ \overrightarrow{\text{grad}}^T(N^1) \right\}}
\end{aligned}$$

Finally, note that if the temperature field approximation is performed by continuous construction on the whole domain Ω , that of its gradient, and hence the heat flux density, is not so *a priori*.

2.2. Discrete problem formulation

In the previous chapter, we saw that it was possible to approximate continuous problem [1.13] with a discrete problem (Galerkin method, equation [1.15]). To this

end, we define two spaces E_ψ^n and E_T^n of dimension n , to replace respectively spaces E_ψ (weighting functions) and E_T (admissible functions).

The finite element method is based upon the sub-domain nodal approximation of the functions ψ and T . First, we will decompose the integral equations for the entire solid Ω appearing in [1.15] as a sum of integral functions on the element domains Ω^e , then introduce the nodal approximation of the functions ψ and T inside each element. This will lead us to define element quantities corresponding to each element of the mesh. Then we will perform the assembly operation, the goal of which is to form a global system to solve. This global system will be the same dimension n_t as the number of nodes used for the mesh. The insertion of the essential boundary conditions (on T) into this global system will then be studied in detail.

2.2.1. Element quantities

If we apply nodal approximation [2.3] to the weighting functions $\psi(\vec{x})$ in discrete problem [1.15], the quantities involved can be written as follows:

$$\begin{aligned} \int_{\Omega} \psi Q \, dv &= \sum_{e=1}^m \langle \psi^e \rangle \cdot \int_{\Omega^e} \{N^e\} Q \, dv \\ \int_{\partial\Omega_q} \psi q \, ds &= \sum_{e=1}^m \langle \psi^e \rangle \cdot \int_{\partial\Omega^e \cap \partial\Omega_q} \{N^e\} q \, ds \\ \int_{\Omega} \overrightarrow{\text{grad}}^T(\psi) \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(T) \, dv &= \sum_{e=1}^m \langle \psi^e \rangle \cdot \int_{\Omega^e} \left\{ \overrightarrow{\text{grad}}^T(N^e) \right\} \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(T) \, dv \end{aligned}$$

If these expressions are transferred into equation [1.15], the quantity to cancel then becomes:

$$\int_{\Omega} \psi Q \, dv + \int_{\partial\Omega_q} \psi q \, ds - \int_{\Omega} \overrightarrow{\text{grad}}^T(\psi) \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(T) \, dv = \sum_{e=1}^m \langle \psi^e \rangle \cdot \{R^e(T)\}$$

where the vector $\{R^e(T)\}$, specific to each element e , is called element residual. If we apply nodal approximation [2.2] to the temperature field, this element residual is expressed according to the nodal temperatures $\{T^e\}$, as follows:

$$\begin{aligned} \{R^e(T)\} &= \int_{\Omega^e} \{N^e\} Q \, dv + \int_{\partial\Omega^e \cap \partial\Omega_q} \{N^e\} q \, ds \\ &\quad - \int_{\Omega^e} \left\{ \overrightarrow{\text{grad}}^T(N^e) \right\} \cdot \underline{\lambda} \cdot \left\langle \overrightarrow{\text{grad}}(N^e) \right\rangle \{T^e\} \, dv \end{aligned} \tag{2.6}$$

The element residual is a fundamental quantity resulting from the finite element method. In this section, we have restricted our study to the so-called linear cases. A linear case is characterized by the fact that the element residual can be expressed linearly according to the node temperatures of the corresponding element, i.e. according to the vector $\{T^e\}$. Equation [2.6] shows that the element residual is a linear function of the node temperatures of the corresponding element if:

- the thermal conductivity of the material (matrix $\underline{\lambda}$) is not temperature-dependent;
- the term Q , the volumetric heat source, is a temperature-independent function (as a matter of fact, it could vary linearly with the temperature, but as far as we know, this type of variation does not correspond to any physical example);
- the heat flux q applied to the boundary $\partial\Omega_q$ of the solid is a linear function of the temperature (for instance $q = h(T_{\text{ext}} - T) + q_0$, where h and q_0 are temperature-independent).

Then it is possible to write:

$$\{R^e(T)\} = \{F^e\} - [K^e] \cdot \{T^e\}$$

In this equation, $[K^e]$ is a square matrix of dimension $n^e \times n^e$. This is called an element matrix. It is also called an element conductance matrix, because we will see that the main contribution comes from thermal conductivity. It is also called an element stiffness matrix, an irritating term for thermal scientists but revealing the origins of the finite element method in mechanical engineering. The vector $\{F^e\}$ in this equation includes n^e terms in a column. It is called an element load vector.

The element matrix $[K^e]$ and the element load vector $\{F^e\}$ are written as follows:

$$[K^e] = [\Lambda^e] + [H^e] \quad \text{with:} \quad \begin{cases} [\Lambda^e] = \int_{\Omega^e} \left\{ \overrightarrow{\text{grad}}^T(N^e) \cdot \underline{\lambda} \cdot \left\langle \overrightarrow{\text{grad}}(N^e) \right\rangle dv \right\} \\ [H^e] = \int_{\partial\Omega^e \cap \partial\Omega_q} \{N^e\} h \langle N^e \rangle ds \end{cases} \quad [2.7]$$

$$\{F^e\} = \int_{\Omega^e} \{N^e\} Q dv + \int_{\partial\Omega^e \cap \partial\Omega_q} \{N^e\} (hT_{\text{ext}} + q_0) ds \quad [2.8]$$

In these general expressions, note that:

- the matrix and element load vector are obtained by integration on the corresponding domain Ω^e . In Chapter 3, we will introduce an efficient numerical method (Gauss method) which makes it possible to evaluate these integral equations accurately;

– the element matrix $[K^e]$ is symmetric. As a matter of fact, the general term K_{ij}^e of this matrix is expressed as follows:

$$K_{ij}^e = \int_{\Omega^e} \left(\overrightarrow{\text{grad}}^T(N_i^e) \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(N_j^e) \right) dv + \int_{\partial\Omega^e \cap \partial\Omega_q} N_i^e h N_j^e ds$$

2.2.2. Assembly

One of the major advantages of the finite element method is that it highlights element quantities, which are therefore calculated in the corresponding element only. However, the temperature field sought is related to the whole environment studied. Thus, the assembly operation consists of turning these element quantities into global quantities so as to construct the equation system to solve.

We saw that the quantity to cancel in the discrete problem resulting from [1.13] was determined according to an element residual, $\{R^e\}$, whose general expression is given by equation [2.6]. If we now use equation [2.1], applied to the weighting functions, this quantity to cancel can be expressed as follows:

$$\sum_{e=1}^m \langle \psi^e \rangle \cdot \{R^e(T)\} = \langle \psi \rangle \cdot \sum_{e=1}^m [A^e]^T \cdot \{R^e(T)\} = \langle \psi \rangle \cdot \{R(T)\} = 0 \quad [2.9]$$

In this equation, $\langle \psi \rangle$ is a vector of dimension n_t , including on a line the weighting function values ψ at the mesh nodes. With regard to the vector $\{R(T)\}$ it is also of dimension n_t and includes in a column n_t functions of T resulting from the element residual assembly at each node. This vector is said to be residual. Since the above expression must be equal to zero whatever the weighting functions ψ , the n_t components of the residual vector must be equal to zero. This is written as follows:

$$\{R(T)\} = \sum_{e=1}^m [A^e]^T \cdot \{R^e(T)\} = \{0\}$$

In fact, we will see later that the insertion of the essential boundary conditions may lead to either not canceling certain components of that residual vector or modifying its expression.

In the linear case, the element residual is written linearly according to the temperatures. Indeed if we use equation [2.1] once more, but this time applied to the nodal temperatures, we obtain:

$$\{R(T)\} = \sum_{e=1}^m [A^e]^T \cdot \{R^e(T)\} = \{F\} - [K] \cdot \{T\} \quad [2.10]$$

In this equation, $[K]$ is a square matrix of dimension $n_t \times n_t$, which is obtained by assembling the element matrices $[K^e]$. It is called a global matrix or a conductance matrix or a stiffness matrix (after mechanical engineering once again):

$$[K] = \sum_{e=1}^m [A^e]^T \cdot [K^e] \cdot [A^e] \quad [2.11]$$

Likewise, $\{F\}$ is a vector of dimension n_t , which is obtained by assembling the element quantities $\{F^e\}$. It is called a load vector or stress global vector:

$$\{F\} = \sum_{e=1}^m [A^e]^T \cdot \{F^e\} \quad [2.12]$$

Canceling the global residual components leads us to solve the following linear system (of dimension n_t):

$$[K] \cdot \{T\} = \{F\}$$

In Chapter 3, we will see that inserting essential boundary conditions modifies either the linear system dimension or the matrix and load vector of that system.

Figure 2.3 illustrates the construction of $[K]$ and $\{F\}$ by assembling the matrices and element load vectors as in the case of the mesh in Figure 2.2:

- the component K_{ij} of $[K]$ corresponds to nodes i and j . It alone incorporates the contribution of elements including nodes i and j . For instance, the component K_{55} receives contributions from the element matrices of elements 1, 2, 3 and 4 because all these elements include node 5. On the contrary, only element 3 which includes nodes 5 and 7 brings a contribution to the component K_{75} . Finally many components K_{ij} are equal to zero because no element incorporates the corresponding nodes (for example $K_{37} = 0$);

- the component F_i of $\{F\}$ corresponds to node i . It includes the contributions of all the elements including node i . For instance, F_6 receives the contributions of the element load vectors of elements 2 and 4 which contain node 6.

It can be noted that the solution of the previous linear system minimizes the following quadratic form:

$$\Pi^* = \frac{1}{2} \langle T \rangle \cdot [K] \cdot \{T\} - \langle T \rangle \cdot \{F\}$$

This quadratic form corresponds to the discretization of the functional Π introduced in formulation [1.14].

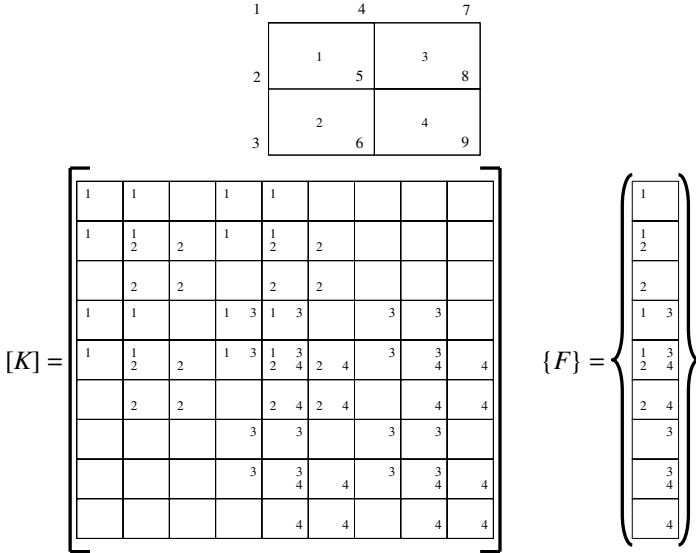


Figure 2.3. Example of an assembly on a simple mesh (contributions of the elements to the global matrix and load vector)

2.3. Solution

In the previous section, we obtained a discrete formulation of the quantity to cancel in continuous problem [1.13]. In the linear case, this formulation leads to the solution of a linear system related to the temperatures at the mesh nodes. In this section, we list the main methods used to insert the essential (temperature) boundary conditions into the system to solve, then detail a few techniques for storing and solving the linear system obtained and end with an analysis of results.

2.3.1. Application of temperature boundary conditions

Reconsider equation [2.9], with, in the linear case, expression [2.10] for the residual $\{R(T)\}$. We obtain:

$$\langle \psi \rangle \cdot (\{F\} - [K] \cdot \{T\}) = 0 \quad [2.13]$$

In this equation, the nodal temperature vector $\{T\}$ is of dimension n_t . Likewise, $\langle \psi \rangle$ is a vector of dimension n_t , including all the nodal parameters of the weighting functions.

It is important to note at this point that, among the n_t nodes of the mesh used, some are located on the boundary $\partial\Omega_T$ of the solid, in which the temperature T is known and in which the weighting functions ψ are therefore equal to zero. If we designate n^* the number of mesh nodes located on that boundary, it is convenient to partition the nodal vectors $\langle\psi\rangle$ and $\{T\}$ as follows:

$$\langle\psi\rangle = \langle\langle\psi_c\rangle \quad \langle\psi_l\rangle\rangle, \quad \{T\} = \left\{ \begin{Bmatrix} \{T_c\} \\ \{T_l\} \end{Bmatrix} \right\}$$

We have gathered from the first components of the vectors $\langle\psi\rangle$ and $\{T\}$ the n^* nodal quantities known *a priori* (vectors $\langle\psi_c\rangle$ and $\{T_c\}$). In order to cancel the weighting functions ψ on the domain boundary $\partial\Omega_T$, it is sufficient to cancel their values at the nodes of the boundary $\partial\Omega_T$, values gathered in vector $\langle\psi_c\rangle$. Likewise, for a function T to be admissible, it is sufficient to give it the temperature values known at the nodes of the boundary $\partial\Omega_T$ which can be expressed as $\{T_c\} = \{T_d\}$. In fact this natural course of action is not the one generally used in finite element software, while linear vectors and systems are never partitioned in this way (with the n^* components numbered first). This expression is introduced to simplify the presentation of the different methods used to apply the temperature boundary conditions.

If the same partitioning method is used for the matrix $[K]$ and the load vector $\{F\}$ of the problem to solve, equation [2.13] can eventually be rewritten as follows:

$$\langle\langle\psi_c\rangle \quad \langle\psi_l\rangle\rangle \cdot \left(\left\{ \begin{Bmatrix} \{F_c\} \\ \{F_l\} \end{Bmatrix} \right\} - \begin{bmatrix} [K_{cc}] & [K_{cl}] \\ [K_{cl}]^T & [K_{ll}] \end{bmatrix} \cdot \left\{ \begin{Bmatrix} \{T_c\} \\ \{T_l\} \end{Bmatrix} \right\} \right) = 0 \quad [2.14]$$

As this equation must be satisfied whatever the vector $\langle\psi_l\rangle$ and with $\langle\psi_c\rangle = \langle 0 \rangle$, this leads to the following linear system, of $n_t - n^*$ equations, for $n_t - n^*$ unknowns (nodal vector $\{T_l\}$):

$$[K_{ll}] \cdot \{T_l\} = \{F_l\} - [K_{cl}]^T \cdot \{T_d\} \quad [2.15]$$

The resolution of this system yields the unknown nodal vector $\{T_l\}$ and hence the temperature at any point of Ω by making use of the shape functions. This specific method, called the elimination method, is rarely used in computer codes.

We might think that, as we know n^* nodal temperatures (at the nodes of $\partial\Omega_T$), the physical problem contains only $n_t - n^*$ unknowns (vector $\{T_l\}$). As a matter of fact, this is not the case. We introduce concurrently to n^* known nodal temperatures n^* additional unknowns corresponding to the heat flux prescribed by the external

environment in order to satisfy the condition $\{T_c\} = \{T_d\}$. Those n^* additional unknowns correspond to the reactions of the external environment at the nodes where the temperature is prescribed; they are gathered in the vector $\{r_c\}$.

The introduction of the vector $\{r_c\}$ now allows us to constitute the linear system to solve, in which the number of unknowns is n_t (the components of the vectors $\{T_l\}$ and $\{r_c\}$) as follows:

$$\begin{Bmatrix} \{F_c\} + \{r_c\} \\ \{F_l\} \end{Bmatrix} - \begin{bmatrix} [K_{cc}] & [K_{cl}] \\ [K_{cl}]^T & [K_{ll}] \end{bmatrix} \cdot \begin{Bmatrix} \{T_c\} \\ \{T_l\} \end{Bmatrix} = 0 \quad [2.16]$$

Whatever the method used to take into consideration the (temperature) essential boundary conditions, the resolution of the problem yields the unknown nodal vector $\{T_l\}$ by satisfying the condition $\{T_c\} = \{T_d\}$. This makes it possible to obtain the temperature field $T(\vec{x})$ by nodal approximation.

The penalty method is the most widely used in computer codes to determine the essential conditions $\{T_c\} = \{T_d\}$. Applying those conditions to certain nodes of a boundary comes down to assuming that at this point the solid is plunged in a fluid at the corresponding temperature, with a very high heat transfer coefficient which will be written P here. This comes down to replacing the (temperature) essential boundary conditions with (node) flux conditions, by using the following expression for the external environment reactions at these points:

$$\{r_c\} = P(\{T_d\} - \{T_c\})$$

With this expression, use equation [2.16], with the vectors $\{T_l\}$ and $\{T_c\}$ (i.e. all the components of the unknown nodal vector $\{T\}$) as unknowns:

$$\begin{bmatrix} [K_{cc}] + P[I] & [K_{cl}] \\ [K_{cl}]^T & [K_{ll}] \end{bmatrix} \cdot \begin{Bmatrix} \{T_c\} \\ \{T_l\} \end{Bmatrix} = \begin{Bmatrix} \{F_c\} + P\{T_d\} \\ \{F_l\} \end{Bmatrix} \quad [2.17]$$

The penalty method therefore consists of adding P on the diagonal of the matrix $[K_{cc}]$ and $P\{T_c\}$ to the vector $\{F_c\}$. In equation [2.17], $[I]$ represents the identity matrix and, in practice, a value $P \gg \text{Max}(K_{ij})$ is selected. This leads to a linear system with predominant terms on the diagonal. In addition to its good numerical stability, this linear system has the advantage of solving a system of similar dimension, whatever the temperature boundary conditions used. The dimension of this system is always equal to the number of mesh nodes.

Another method is commonly used in computer codes. It is called Lagrange multipliers and consists of introducing an unknown vector $\{\lambda_c\}$, of dimension

n^* , into the function to cancel. This unknown vector will be compensated by n^* additional equations $\{T_c\} = \{T_d\}$. The system to solve therefore includes $n_t + n^*$ equations, for $n_t + n^*$ unknowns (vectors $\{T_c\}$, $\{T_l\}$ and $\{\lambda_c\}$):

$$\begin{bmatrix} [K_{cc}] & -k[I] & [K_{cl}] \\ -k[I] & 0 & 0 \\ [K_{cl}]^T & 0 & [K_{ll}] \end{bmatrix} \cdot \begin{Bmatrix} \{T_c\} \\ \frac{1}{k}\{\lambda_c\} \\ \{T_l\} \end{Bmatrix} = \begin{Bmatrix} \{F_c\} \\ -k\{T_d\} \\ \{F_l\} \end{Bmatrix} \quad [2.18]$$

A factor k is introduced into equation [2.18]. This factor is homogenous with a thermal conductivity and allows us to condition the system properly (for its numerical solution). Each node value included in the vector $\{\lambda_c\}$ plays the role of a Lagrange multiplier corresponding to a prescribed temperature condition. System [2.18] can then be obtained directly (with $k = 1$) by solving the following saddle point problem:

$$\text{Max}_{\lambda \geq 0} \left(\text{Min}_T \left(\frac{1}{2} \langle T \rangle \cdot [K] \cdot \{T\} - \langle T \rangle \cdot \{F\} + (\langle T \rangle - \langle T_d \rangle) \cdot \{\lambda_c\} \right) \right)$$

It can be noted that the system to solve in equation [2.18] is always symmetric (as it is with other methods). One difficulty of this type of method, combined with pivot methods used to solve the linear system, is that it reveals diagonal terms equal to zero in the first member matrix. This difficulty can be overcome by numbering the equations to solve. This method also has the drawback of increasing the size of the linear system to solve.

In the case of Lagrange multipliers, the external environment reactions required to satisfy the condition $\{T_c\} = \{T_d\}$ i.e. the vector $\{r_c\}$ components, coincide with Lagrange multipliers $\{\lambda_c\}$.

2.3.2. Linear system solution

The matrices involved in linear systems [2.15], [2.17] and [2.18] are generally large. Indeed, it is common today to process models including several tens and even hundreds of thousands of nodes.

In order to solve the linear system, it is necessary to use the method best adapted to the problem being dealt with. Most of the time, the size of the problem and the accuracy desired in its resolution will be the key factors in the choice of the solution method to use. The problem size is now a major factor, all the more so as a matrix storage method is related to each solution method (see next section) and the problem size a computer is capable of processing is directly limited by its memory space.

A comprehensive description of all existing methods would go beyond the scope of this book. We will therefore restrict our study to the most common methods used in academic and industrial computer codes. Interested readers are invited to consult specialized books such as [CIA 82].

Consider therefore a linear equation system similar to that obtained by the finite element technique. This system is as follows:

$$[A] \cdot \{x\} = \{b\} \quad [2.19]$$

where $\{x\}$ represents the unknowns vector, $[A]$ the matrix and $\{b\}$ the load vector. The system size (number of unknowns and number of equations) will be written n . This size, as well as the expression of $\{x\}$, $[A]$ and $\{b\}$ according to the quantities defined in this book, depends on the method used to apply the temperature boundary conditions:

– with an elimination method, we have a system of dimension $n = n_t - n^*$, that is, total number of nodes minus the number of nodes at which the temperature is known, where the components are expressed according to equation [2.15]:

$$\left| \begin{array}{l} [A] = [K_{ll}] \\ \{x\} = \{T_l\} \\ \{b\} = \{F_l\} - [K_{cl}]^T \cdot \{T_d\} \end{array} \right.$$

– with a penalty method, the system is of dimension $n = n_t$, the total number of nodes, and equation [2.17] allows us to express the system components as follows:

$$\left| \begin{array}{l} [A] = \begin{bmatrix} [K_{cc}] + P[I] & [K_{cl}] \\ [K_{cl}]^T & [K_{ll}] \end{bmatrix} \\ \{x\} = \begin{Bmatrix} \{T_c\} \\ \{T_l\} \end{Bmatrix} \\ \{b\} = \begin{Bmatrix} \{F_c\} + P\{T_d\} \\ \{F_l\} \end{Bmatrix} \end{array} \right.$$

– with a Lagrange multiplier method, the linear system dimension becomes $n = n_t + n^*$, total number of nodes plus the number of nodes at which the temperature is known, and the components are written according to equation [2.18]:

$$\left| \begin{array}{l}
 [A] = \begin{bmatrix} [K_{cc}] & -k[I] & [K_{cl}] \\ -k[I] & 0 & 0 \\ [K_{cl}]^T & 0 & [K_{ll}] \end{bmatrix} \\
 \{x\} = \begin{Bmatrix} \{T_c\} \\ \frac{1}{k} \{\lambda_c\} \\ \{T_l\} \end{Bmatrix} \\
 \{b\} = \begin{Bmatrix} \{F_c\} \\ -k \{T_d\} \\ \{F_l\} \end{Bmatrix}
 \end{array} \right.$$

The solution of system [2.19] is simply written $\{x\} = [A]^{-1} \cdot \{b\}$. However, contrary to what this expression might suggest, computing $\{x\}$ never requires us to calculate the inverted matrix $[A]$. As a matter of fact, computing an inverted matrix of dimension $n \times n$ requires the solution of n linear systems.

Globally, linear system solution methods are arranged into 2 classes:

- direct methods (Gauss, Cholesky, etc.),
- iterative methods (Jacobi, Gauss-Seidel, conjugate gradients, etc.).

2.3.2.1. Direct methods

The general principle of direct methods is to transform the initial equation system into a system with the same solution but whose resolution is simpler. Among the whole set of direct methods, the Gauss method is probably the most popular one. It is a general method which is applied to symmetric or non-symmetric matrices, providing that these matrices can be inverted, of course. It consists of two steps:

- an elimination step which triangulates system [2.19],
- a second step to solve the triangular system obtained.

The triangulation of system [2.19], of dimension n , is performed in $n - 1$ steps. At each step, an equivalent linear system is constructed by eliminating the terms under the diagonal of an additional column of the matrix. Let $[A]^{(1)} = [A]$, $\{b\}^{(1)} = \{b\}$, and let $[A]^{(i)}$ and $\{b\}^{(i)}$, respectively be the matrix and load vector of the equivalent linear system obtained after $i - 1$ steps; $[A]^{(i)}$ is written as follows:

$$[A]^{(i)} = \begin{bmatrix} A_{11}^{(i)} & \cdots & A_{1i}^{(i)} & \cdots & A_{1n}^{(i)} \\ & \ddots & \vdots & & \vdots \\ 0 & & A_{ii}^{(i)} & \cdots & A_{in}^{(i)} \\ & & \vdots & & \vdots \\ & & A_{ni}^{(i)} & \cdots & A_{nn}^{(i)} \end{bmatrix}$$

Matrix $[A]^{(i+1)}$ and load vector $\{b\}^{(i+1)}$ in the following step are then obtained by:

$$\begin{aligned} [A]^{(i+1)} &= [E]^{(i)} \cdot [A]^{(i)} \\ \{b\}^{(i+1)} &= [E]^{(i)} \cdot \{b\}^{(i)} \end{aligned} \quad \text{with} \quad [E]^{(i)} = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & 0 \\ & & -\frac{A_{i+1,i}^{(i)}}{A_{ii}^{(i)}} & 1 & \\ & 0 & \vdots & 0 & \ddots \\ & & -\frac{A_{ni}^{(i)}}{A_{ii}^{(i)}} & & 1 \end{bmatrix}$$

It is clear that this operation is possible only if the terms $A_{ii}^{(i)}$ (called pivots) are not equal to zero. If matrix $[K]$ is singular, we will necessarily have a pivot equal to zero even if $[K]$ can be inverted. As a demonstration, examine the following system of 2 equations with 2 unknowns:

$$\begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix} \cdot \begin{Bmatrix} T_1 \\ T_2 \end{Bmatrix} = \begin{Bmatrix} 100 \\ 100 \end{Bmatrix}$$

It is obvious that the first pivot encountered with the Gauss method is equal to zero, whereas the system admits the unique solution $T_1 = 50$, $T_2 = 100$. To perform the triangulation, it is then necessary to replace the equation considered with one of the following. (Generally speaking, to obtain the highest possible solution accuracy, it will be sensible to process at each step the equation with the pivot whose absolute value is highest.)

As a matter of fact, the Gauss method comes down to factorizing the matrix of system [2.19] as follows:

$$[A] = [L] \cdot [U]$$

where $[L]$ is a lower triangular matrix and $[U]$ an upper triangular matrix. The matrix $[U]$ coincides with the triangular matrix obtained by the elimination process described previously and $[L] = [E^{(1)}]^{-1} \cdot [E^{(2)}]^{-1} \dots [E^{(n)}]^{-1}$. The solution of a linear system with the Gauss method requires about $\frac{2n^3}{3}$ arithmetic operations. When the matrix of system [2.19] is symmetric positive definite as is the case in heat conduction linear problems, it will be wise to choose the Cholesky method which requires only about $\frac{n^3}{3}$ arithmetic operations. This method consists of factorizing $[A]$ as follows:

$$[A] = [L] \cdot [L]^T$$

Matrix $[L]$ is a lower triangular matrix whose elements are computed gradually. We first calculate $L_{11} = \sqrt{A_{11}}$ then, for $i = 2, \dots, n, L_{i1} = \frac{A_{i1}}{L_{11}}$. The L_{ip} being then assumed to be calculated for $p = 1, \dots, j-1$ and $i = p+1, \dots, n$, we calculate consecutively:

$$L_{jj} = \sqrt{A_{jj} - \sum_{p=1}^{j-1} L_{jp}^2}$$

$$\forall i = j+1, \dots, n, \quad L_{ij} = \frac{\sqrt{A_{ij} - \sum_{p=1}^{j-1} L_{ip}L_{jp}}}{L_{jj}}$$

2.3.2.2. Iterative methods

Iterative methods give the solution of system [2.19] as the limit of a series of vectors $\{x\}^{(i)}$. They are very efficient at processing problems of a large size insofar as they require less data than direct methods, as will be seen in the next section.

The general principle of these methods is to factorize the matrix $[A]$ of system [2.19] as follows:

$$[A] = [M] - [N]$$

then to approximate the solution iteratively by solving, at each iteration, the system:

$$[M] \cdot \{x\}^{(i+1)} = [N] \cdot \{x\}^{(i)} + \{b\}$$

Matrix $[M]$ will be chosen so that the solution is simple. The Jacobi method, the Gauss-Seidel method and the over-relaxation method operate on this principle and correspond to specific choices of matrices $[M]$ and $[N]$.

Now assume $[A]$ to be symmetric positive definite and spend some time on the conjugate gradient method. In this method, we consider the function f which relates the following scalar to any vector $\{x\}$:

$$f(\{x\}) = \frac{1}{2} \langle x \rangle \cdot [A] \cdot \{x\} - \langle x \rangle \cdot \{b\}$$

It is easy to note that the solution sought for linear system [2.19] coincides with the minimizing vector f . The conjugate gradient method consists of approximating step by step the solution by minimizing the function f at each iteration in a particular direction, called descent direction. Consider $\{u\}^{(i)}$, an arbitrary vector for the moment, as the direction of the iteration i . The iterated $\{x\}^{(i+1)}$ is chosen as $\{x\}^{(i+1)} = \{x\}^{(i)} + \alpha^{(i)} \{u\}^{(i)}$, and $\alpha^{(i)}$ is determined in such a way that $\{x\}^{(i+1)}$ minimizes f in the direction $\{u\}^{(i)}$. We find:

$$\alpha^{(i)} = \frac{\langle u \rangle^{(i)} \cdot \{r\}^{(i)}}{\langle u \rangle^{(i)} \cdot [A] \cdot \{u\}^{(i)}}$$

where $\{r\}^{(i)} = \{b\} - [A] \cdot \{x^{(i)}\}$ is the residual vector of iteration i .

Starting with a test vector $\{x\}^{(0)}$ and a descent direction $\{u\}^{(0)} = \{r\}^{(0)}$, the vector $\{u\}^{(i)}$ is constructed at each iteration so that $\langle u \rangle^{(i)} \cdot [A] \cdot \{u\}^{(i-1)} = 0$. For this purpose, we write $\{u\}^{(i)} = \{r\}^{(i)} + \beta^{(i)} \{u\}^{(i-1)}$. We then demonstrate by recursion that we have:

$$\forall k = 1, \dots, i-1, \quad \begin{cases} \langle r \rangle^{(i)} \cdot \{u\}^{(k)} = 0 \\ \langle r \rangle^{(i)} \cdot \{r\}^{(k)} = 0 \\ \langle u \rangle^{(i)} \cdot [A] \cdot \{u\}^{(k)} = 0 \end{cases}$$

According to the latter relationship, we say that the consecutive descent directions are $[A]$ -combined. We eventually obtain:

$$\begin{aligned} \{u\}^{(i)} &= \{r\}^{(i)} - \frac{\langle r \rangle^{(i)} \cdot [A] \cdot \{u\}^{(i-1)}}{\langle u \rangle^{(i-1)} \cdot [A] \cdot \{u\}^{(i-1)}} \{u\}^{(i-1)} \\ &= \{r\}^{(i)} + \frac{\langle r \rangle^{(i)} \cdot \{r\}^{(i)}}{\langle r \rangle^{(i-1)} \cdot \{r\}^{(i-1)}} \{u\}^{(i-1)} \end{aligned}$$

The conjugate gradient method converges to the solution in, at most, n iterations. In this sense it can be considered to be a direct method. Various pre-conditioning

methods make it possible to obtain an accurate solution with a relatively small number of iterations. This method is extended to the case of matrices which are not positive definite symmetric.

2.3.3. *Storing the linear system matrix*

In practice, the matrix of the linear system to solve is stored by means of a method using its properties. For example, if it is symmetric, it will be sufficient to store only the upper triangular part. If, in addition, it is “sparse” (i.e. it includes a great number of terms equal to zero), which is the case with the finite element method, the following methods are generally used:

- the skyline storage of a matrix is the most widely used method in computer codes. The matrix is stored in columns (or lines) variable in length. It is particularly well-adapted to linear system direct solution methods (Gauss, Choleski, etc.), which consist of triangulating the matrix. Figure 2.4 illustrates the storage of the matrix $[K]$ resulting from the assembly of Figure 2.3. It should be noted that only the terms of the upper part of the matrix have been stored. The maximum number of terms stored per line on a symmetric matrix is half the bandwidth. In Figure 2.4, the half bandwidth equals 5, whereas the number of nodes is 9. We can show that generally speaking the half bandwidth is equal to the maximum difference between the node numbers of a same element, increased by 1. In Figure 2.4, we obtain for instance $(5 - 1) + 1 = 5$ considering element 1. If we renumber the mesh nodes in a sensible way, it is possible to replace certain lines and columns in the matrix so as to reduce the bandwidth of the matrix stored. This renumbering operation is very important from a practical point of view. Indeed it allows us to substantially reduce the number of terms to store and to accelerate the solution process. If b is the system half bandwidth, then this method involves the storage of a number of values equal to about b times the matrix dimension to which it is necessary to add a pointer per line (or column) giving its length. As b is globally proportional to the matrix dimension, the number of terms to store is proportional to the square of the problem dimension;

- the sparse or compact storage of a matrix is used mainly when an iterative solution method (Jacobi, Gauss-Seidel, conjugate gradients, etc.) of the linear system is employed. This method stores only the matrix terms different from zero. To illustrate this storage method, consider the mesh in Figure 2.4. We see that the maximum number of non-vanishing terms on a line of the matrix is equal to 9. This number corresponds to the maximum number of nodes connected to one node, which is reached here at node 5. If we refine the discretization, by dividing our rectangle into 10×10 elements, we see that the maximum number of nodes connected to a same node, and hence of terms different from zero on a matrix line, remains equal to 9. Therefore, the number of terms different from zero to store is proportional to

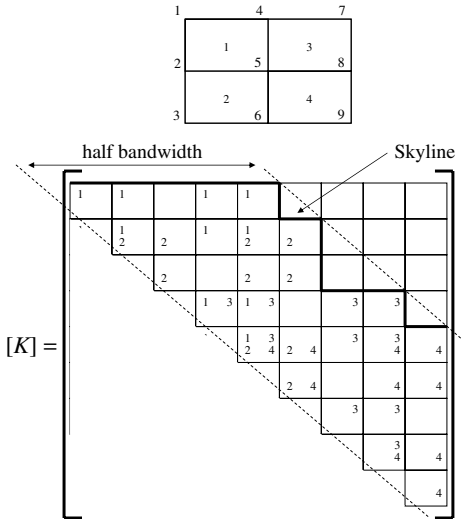


Figure 2.4. Example of skyline storage

the number of nodes (equal to kn_t). The proportionality factor k is given by the mean number of nodes connected to a given node (in the order of 9 in our example). According to the type of elements, the value of this factor ranges from 8 to 20 on a 2D mesh and 20 to 80 in 3D cases. However, for each term different from zero, it is necessary to store its value and position in the matrix (line and column numbers). Finally the number of terms to store is proportional to n_t and equal to $3kn_t$, which makes this storage method particularly efficient for large equation systems.

2.3.4. Analysis of results

Solving a steady state heat conduction problem by means of the finite element method leads to a vector $\{T\}$ including the temperatures calculated at the mesh nodes. This vector makes it possible to calculate a temperature field at any point of the domain of study by using the nodal approximation and therefore the shape functions. It also enables us to calculate the temperature gradient field with the shape function gradients.

In order to obtain that result, the method involves the calculation of element quantities obtained by numerical integration with the Gauss method (see Chapter 3). As a consequence, the variables involved in element quantities are calculated in each element of the mesh, at the integration points (Gauss points). The results obtained during analysis can then be arranged into two categories and they are not used in the same way:

– the values calculated at the nodes (nodal quantities), which in our case are the temperatures. If the temperature field nodal approximation (equation [2.2]) is used, the temperature at each point of the solid in thermal equilibrium can be calculated. This can be visualized by drawing the temperature variation along a curve crossing the solid or temperature contour lines or surfaces. Figure 2.12 illustrates the drawing of temperature contour lines;

– the values calculated at the element integration points (see section 3.2) are those involved in element quantities. In our case, it is essentially the temperature gradient and heat flux density $\vec{\phi}$ which can be derived from Fourier's law [1.3]:

$$\begin{aligned}\overrightarrow{\text{grad}}(T) &= \langle \overrightarrow{\text{grad}}(N^e) \rangle \cdot \{T^e\} \\ \vec{\phi} &= -\lambda \overrightarrow{\text{grad}}(T)\end{aligned}$$

As seen in section 2.1, the heat flux density approximation, which is proportional to the thermal gradient according to Fourier's law [1.3], is not continuous on all the domain Ω . From a physical point of view, the heat flux density is continuous in a homogenous environment. Therefore, it may be interesting to construct an approximation of the heat flux density on the whole domain from the values calculated at the element integration points. To do so, it is traditional to use the finite element approximation from nodal values of that density. These nodal values can be obtained in several different ways, for instance:

- by extrapolating at the nodes of each element the values calculated at the integration points, then computing the mean node values obtained;
- by smoothing according to the method described hereafter.

In addition to obtaining a more accurate approximation, this method allows us to use the same isovalue drawing tools as those used for the temperature field.

2.3.4.1. Smoothing the heat flux density

Among the quantities estimated at the mesh integration points, we have selected the heat flux density $\vec{\phi}$ to illustrate a smoothing method.

As mentioned above, a fairly natural way of smoothing these quantities consists of using the finite element approximation used for the analysis. We then have to seek a field $\vec{\Phi}(\vec{x})$ in the form of a nodal approximation:

$$\forall \vec{x} \in \Omega^e, \quad \begin{cases} \vec{\Phi}(\vec{x}) = \sum_{i=1}^{n^e} N_i^e(\vec{x}) \vec{\Phi}_i^e = \langle N^e \rangle \cdot \{ \vec{\Phi}^e \} \\ \{ \vec{\Phi}^e \} = [A^e] \cdot \{ \vec{\Phi} \} \end{cases}$$

In this equation, Ω^e is an element domain (a mesh element), n^e is the number of nodes of this element, $\langle N^e \rangle$ is a vector including its n^e shape functions on a line, and $\{\vec{\Phi}^e\}$ is a vector including in a column the n^e node values of Ω^e of the heat flux density vector (the number of components of that vector depends on the geometric size of the problem dealt with). Besides, $[A^e]$ is the matrix connecting the vector of local unknowns, $\{\vec{\Phi}^e\}$, to the vector $\{\vec{\Phi}\}$ including in a column the nodal values of all the heat flux density components.

Therefore, the objective is to determine the nodal values included in the vector $\{\vec{\Phi}\}$ from the field $\vec{\phi}$ calculated from the nodal temperatures. These values are calculated in order to minimize the following error function:

$$F = \int_{\Omega} (\vec{\phi} - \vec{\Phi}) \cdot (\vec{\phi} - \vec{\Phi}) dv$$

It should be noted at this point that the error functional F does not necessarily result from an integration on the whole domain studied. It is possible to restrict the integration to a portion of the domain here. This is necessary when, for instance, the environment is composed of the discrete assembly of several materials (for example, coated metal sheet). In this case, the thermal gradient is no longer continuous at the interface between the two materials.

If we simply state that the variation δF of F implied by any variation $\delta \vec{\Phi}$ of $\vec{\Phi}$, at $\vec{\phi}$ fixed, is equal to zero, we obtain:

$$\delta F = -2 \int_{\Omega} \delta \vec{\Phi} \cdot (\vec{\phi} - \vec{\Phi}) dv = 0$$

If we use the nodal approximation to express $\vec{\Phi}$ and $\delta \vec{\Phi}$, we finally obtain the following linear system to solve:

$$[L] \cdot \{\vec{\Phi}\} = \{\vec{S}\} \quad [2.20]$$

In that system only the load vector $\{\vec{S}\}$ involves the heat flux density estimated at the integration points. That load vector is expressed in the form of the assembly of element quantities $\{\vec{S}^e\}$:

$$\{\vec{S}\} = \sum_{e=1}^m [A^e]^T \cdot \{\vec{S}^e\} \quad \text{with} \quad \{\vec{S}^e\} = \int_{\Omega^e} \{N^e\} \vec{\phi} dv$$

It will then be possible to calculate the element quantities $\{\vec{S}^e\}$ by means of the Gauss integration scheme, which directly uses the heat flux densities known at the integration points.

Matrix $[L]$ of the system to solve for smoothing is in fact independent of the quantities to smooth. It is also expressed as the assembly of element quantities $[L^e]$:

$$[L] = \sum_{e=1}^m [A^e]^T \cdot [L^e] \cdot [A^e] \quad \text{with} \quad [L^e] = \int_{\Omega^e} \{N^e\} \cdot \langle N^e \rangle dv$$

Note in equation [2.20] that smoothing heat flux densities requires, in a 3D analysis the solution of a linear system including $3n_t$ equations for $3n_t$ unknowns (n_t is the number of mesh nodes). This is *a priori* more costly than the original analysis. However, it can be noted that matrix $[L]$ has a very special form, similar to the thermal capacitance matrix encountered in transient analyses (see Part 2). In the case of transient analyses, it is common to use approximate capacitance matrices in a lumped diagonal form because it dramatically simplifies the calculations while producing acceptable results.

By operating here in the same way, matrix $[L]$ is replaced with a lumped diagonal matrix $[L^d]$ obtained in this instance by transferring on the diagonal the sum of the terms on each line (δ_{ij} is the Kronecker symbol):

$$L_{ij}^d = \delta_{ij} \sum_{k=1}^{n_t} L_{ik}$$

It is obvious that the solution of system [2.20] is then immediate.

This method is applicable to any value calculated at the integration points (thermal gradient, flux, phase proportions, etc.). It may lead to node values greater than those actually calculated at the integration points. This may be a problem when the physical values involved are limited (phase proportions for instance are to range from 0 to 1). It is then necessary to introduce physical boundaries into the minimization problem.

2.3.4.2. Result accuracy

It is very important to underline at this point that the finite element method is a technique for the approximate solution of the problem encountered. Among the many sources of error appearing during the numerical simulation of a physical problem, we will focus in this book on those always present during finite element analysis:

- errors in physical data input (material properties, boundary conditions and stresses);

- geometric errors due to mesh representation;
- numerical errors (element quantity integration and linear system solution);
- discretization errors (nodal approximation).

The most important errors are probably related to input data. Indeed, the material's physical properties or boundary conditions are very often known to an approximation of 10%. It is then futile to try and refine a mesh to gain 5% on the numerical result accuracy.

Geometric errors result from the inability of the mesh to accurately represent the geometry of the object considered. It is therefore necessary for the user of finite element calculation software to create the mesh best adapted to the physical problem and geometry studied. Only geometric details superfluous to the physical model considered can be eliminated.

Numerical errors related to the finite element method are mainly due to the integration diagram used to calculate the element quantities and to the method used to solve the linear system obtained. Numerical integration errors occurring when calculating element quantities are generally negligible when one does not depart from the traditional integration scheme (see Chapter 3). Numerical errors in the solution can easily be evaluated by calculating, after resolution, the residual:

$$\{R_l\} = \{F_l\} - [K_{ll}] \cdot \{T_l\} - [K_{cl}]^T \cdot \{T_d\}$$

Strictly speaking, the norm of $\{R_l\}$ must be equal to zero. In practice it must be very small compared with the values of the loads applied or the norm of the vector $\{r_c\}$ giving the heat fluxes calculated at the nodes where the temperature was prescribed.

Before beginning a discretization error calculation, it is possible to check a few element items, in particular by ensuring that the flux density calculated at the domain boundary satisfies accurately enough the flux boundary conditions (i.e. $\vec{\phi} \cdot \vec{n} = -q$ on $\partial\Omega_q$).

It is difficult to evaluate with accuracy the discretization error made. There are methods to do so, however. The finite elements presented in this book guarantee the temperature field continuity on the whole domain of study Ω but not that of its derivatives (in particular the temperature gradient and heat flux density). The methods to evaluate the approximation error are generally based upon the quantification of these discontinuities (for example, the heat flux density) on the element boundaries.

Applied to thermal science, the *a posteriori* most widespread error evaluation method in computer codes [ZIE 87] consists of:

- 1) constructing a continuous heat flux density field on the whole domain from the densities calculated directly at the integration points;
- 2) quantifying the deviation between these continuous fields and the fields actually calculated.

For example, it is possible to use the above method to construct a continuous heat flux density field. Then it is possible to define an error E^e on each element as follows (V^e is the volume occupied by Ω^e and V that occupied by Ω):

$$E^e = \sqrt{\frac{\frac{1}{V^e} \int_{\Omega^e} \|\vec{\phi} - \vec{\Phi}\|^2 dv}{\frac{1}{V} \int_{\Omega} \|\vec{\Phi}\|^2 dv}} \quad [2.21]$$

Generally speaking, in order to reduce the discretization error, it will be necessary to finely mesh an area with high temperature gradient or heat flux density variations. There are two main methods to increase accuracy in the solution obtained:

- method h (the most common), which consists of increasing the number of nodes by reducing the mesh size in high gradient variation areas;
- method p , which consists of increasing gradually the approximation degree while adding degrees of freedom to the problem.

2.4. Working example

In this section, the objective is to illustrate the notions reviewed in this chapter on a 2D example. To do so, we consider the solid in Figure 2.5, inside which runs a heat

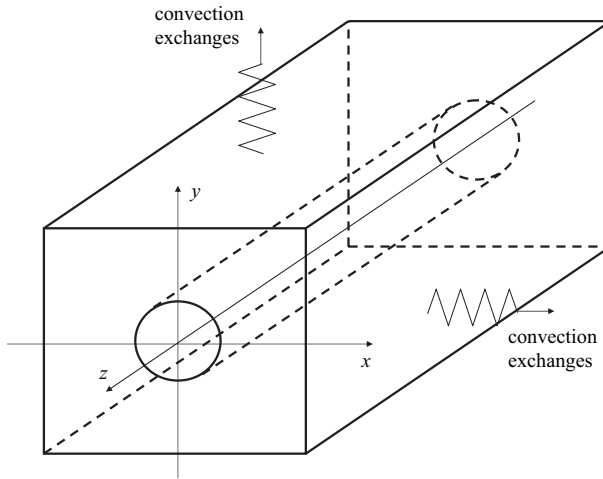


Figure 2.5. Working example

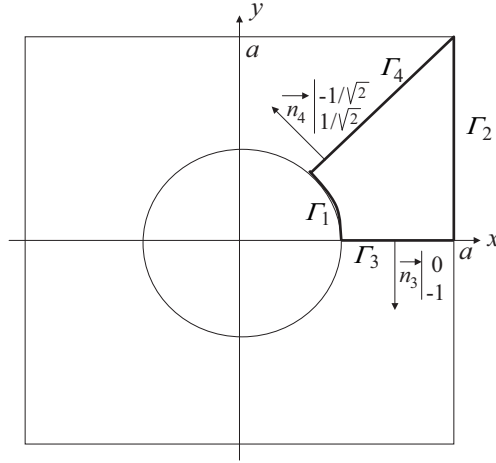


Figure 2.6. Working example: 2D diagram

carrying fluid at temperature T_{int} . The exchange coefficient between the heat carrying fluid and the solid is assumed to be high enough to consider the temperature to be constant and equal to T_{int} inside the solid. On its external wall, an exchange with the environment at temperature T_{ext} is represented by a convection flux. The exchange coefficient is written h and assumed to be constant. Now it is necessary to determine the temperature field in the solid according to T_{int} , T_{ext} , h , and the heat conductivity λ of the material assumed to be isotropic and constant.

The problem is represented by a 2D section (with unit thickness) of the solid along axis Oz . The problem symmetries make it possible to restrict the study to one-eighth of that section. Figure 2.6 illustrates this representation. The physical process to model is therefore the thermal equilibrium section. By taking the symmetric conditions into consideration, we obtain:

$$\left\{ \begin{array}{ll} \lambda \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right) = 0 & \text{in } \Omega \\ T = T_{\text{int}} & \text{on } \Gamma_1 \\ \lambda \frac{\partial T}{\partial x} = h(T_{\text{ext}} - T) & \text{on } \Gamma_2 \\ \vec{\text{grad}}(T) \cdot \vec{n}_3 = -\frac{\partial T}{\partial y} = 0 & \text{on } \Gamma_3 \\ \vec{\text{grad}}(T) \cdot \vec{n}_4 = -\frac{\sqrt{2}}{2} \left(\frac{\partial T}{\partial x} - \frac{\partial T}{\partial y} \right) = 0 & \text{on } \Gamma_4 \end{array} \right.$$

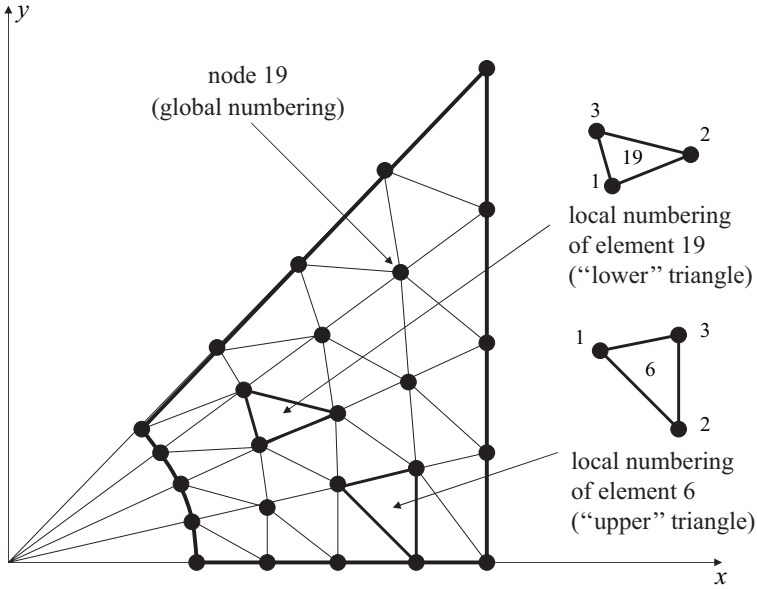


Figure 2.7. Working example: mesh

This problem will be addressed with the finite element method. After constructing the mesh, we will calculate the element quantities, make the assembly, apply the temperature boundary conditions (on Γ_1) and solve the problem.

2.4.1. Finite element approximation

2.4.1.1. Mesh

The mesh is constructed by means of first order triangular elements (with three nodes). To do so, we use in this instance:

- an angle division in p angle sector $\pi/4p$;
- a division of the radiuses obtained into p segments of equal length.

Thus, we obtain $n = (p + 1)^2$ nodes and $m = 2p^2$ triangular finite elements. This is described in Figure 2.7 in the case $p = 4$, i.e. 25 nodes and 32 elements. The local numbering of the elements is also described in this figure. If we now number globally the nodes “radius after radius”, then “from left to right”, a node will be characterized by an index I referring to the angular sector, varying from 1 to $p + 1$, and an index J referring to its position compared with the origin, varying from 1 to $p + 1$. It will always be possible to determine the global number i of a node (varying from 1 to

$n = (p + 1)^2$) by means of an angular index I and a distance index J in the form $i = (I - 1)(p + 1) + J$. The coordinates of node i will then be:

$$\begin{cases} x_i = R \cos \alpha_I + \frac{J-1}{p}(a - R \cos \alpha_I) \\ y_i = R \sin \alpha_I + \frac{J-1}{p}(a - R \sin \alpha_I) \end{cases} \quad \text{with} \quad \alpha_I = \frac{\pi}{4p}(I-1)$$

If we number the elements in the same way, an element will be characterized by an index K referring to the angular sector, varying from 1 to p , and an index L referring to its position compared with the origin, varying from 1 to p . If we differentiate the “lower” triangles (with the odd index e) from the “upper” triangles (with the even index e), the connectivities are written as follows:

- for the elements $e = 2p(K - 1) + 2L - 1$ (“lower” triangles), the connected nodes have numbers $(K - 1)(p + 1) + L$, $K(p + 1) + L + 1$ and $K(p + 1) + L$;
- for the elements $e = 2p(K - 1) + 2L$ (“upper” triangles), the connected nodes have numbers $(K - 1)(p + 1) + L$, $(K - 1)(p + 1) + L + 1$ and $K(p + 1) + L + 1$.

This numbering method makes it possible to index the nodes in the mesh and construct for each element e the matrix $[A^e]$ locating it. For example, for element 19, in Figure 2.7, that matrix has 3 lines and 25 columns. Its only terms different from zero are $A_{1,12}^{19} = 1$, $A_{2,13}^{19} = 1$ and $A_{3,17}^{19} = 1$.

2.4.1.2. Nodal approximation

The nodal approximation will now be performed with the shape functions $N^e(x, y)$ satisfying the conditions mentioned in section 2.1. If i , j and k are the local numbers of the nodes making element e , then the triplet (i, j, k) is an even permutation of $(1, 2, 3)$, and, for instance, $N_i^e(x, y)$ must equal 1 at nodes i and 0 on the segment connecting nodes j and k . This is illustrated by Figure 2.8. From three points with coordinates (x_i, y_i) , (x_j, y_j) and (x_k, y_k) , it is possible to define a linear function equal to 1 at the first node and 0 on the segment connecting the other two. That linear function is the shape function N_i^e of node i (local numbering) of element e . It is expressed as $N_i^e(x, y) = a_i^e x + b_i^e y + c_i^e$, where the coefficients a_i^e , b_i^e and c_i^e , depend on the coordinates of the nodes forming element e in the following form:

$$\begin{cases} a_i^e = \frac{y_j - y_k}{(y_i - y_j)(x_k - x_j) - (x_i - x_j)(y_k - y_j)} \\ b_i^e = \frac{x_k - x_j}{(y_i - y_j)(x_k - x_j) - (x_i - x_j)(y_k - y_j)} \\ c_i^e = \frac{x_j(y_k - y_j) - y_j(x_k - x_j)}{(y_i - y_j)(x_k - x_j) - (x_i - x_j)(y_k - y_j)} \end{cases}$$

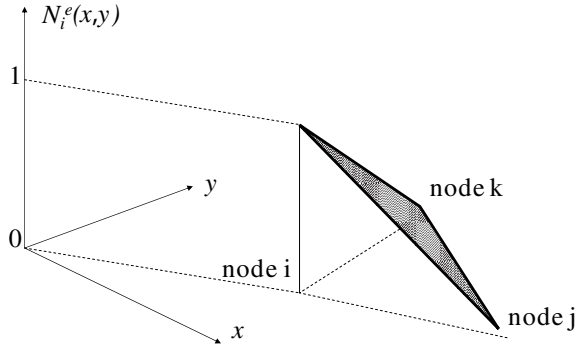


Figure 2.8. Working example: shape function $N_i^e(x, y)$

2.4.2. Discrete problem formulation

2.4.2.1. Element quantities

In order to estimate the element quantities, we select an element e . Its surface is termed S^e . It is obtained from the coordinates of nodes i , j and k , belonging to the element by the following expression:

$$S^e = \frac{1}{2} |(x_j - x_i)(y_k - y_i) - (y_j - y_i)(x_k - x_i)|$$

The components of the element matrix $[K^e]$ and of the load vector $\{F^e\}$ of element e are given by the general formulae previously obtained. In this working example, the zero flux conditions (on Γ_3 and Γ_4) only yield zero terms in the load vector. Besides, as we considered a unit thickness for the representation in Figure 2.6, the volume and surface integrals become respectively surface and contour integrals. We finally obtain:

$$\begin{cases} K_{ij}^e = \lambda \int_{S^e} \left(\frac{\partial N_i^e}{\partial x} \frac{\partial N_j^e}{\partial x} + \frac{\partial N_i^e}{\partial y} \frac{\partial N_j^e}{\partial y} \right) ds + h \int_{\partial S^e \cap \Gamma_2} N_i^e N_j^e dl \\ F_i^e = h T_{\text{ext}} \int_{\partial S^e \cap \Gamma_2} N_i^e dl \end{cases}$$

In this equation, note that the first term of $[K^e]$, related to the thermal conductivity, can be expressed directly for any element e in the form $\lambda S^e (a_i a_j + b_i b_j)$. On the contrary, the second term and the vector $\{F^e\}$, which are the contributions of convection exchanges (coefficient h), are more difficult to determine.

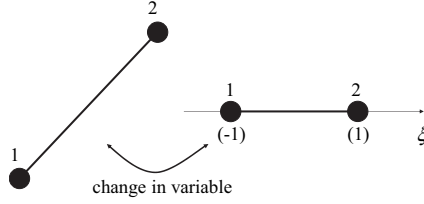


Figure 2.9. Working example: skin elements on Γ_2

An intelligent means of estimating the terms due to convection exchanges (terms of K_{ij}^e and F_i^e involving an integration on $S^e \cap \Gamma_2$) is to introduce an additional mesh on the boundary Γ_2 on which these conditions are applied. That additional mesh is often referred to as a skin mesh. In this 2D example, it would be made of two-noded linear elements, with coordinates (x_1, y_1) and (x_2, y_2) . Then by changing the variables, as described in Figure 2.9, it is possible to characterize the position of a point on this element with a parameter ξ varying from -1 to 1 as follows:

$$\begin{cases} x(\xi) = N_1(\xi)x_1 + N_2(\xi)x_2 \\ y(\xi) = N_1(\xi)y_1 + N_2(\xi)y_2 \end{cases} \quad \text{with} \quad \begin{cases} N_1(\xi) = \frac{1-\xi}{2} \\ N_2(\xi) = \frac{1+\xi}{2} \end{cases}$$

Now if we term L^e the length of the skin element corresponding to the surface element e , we note that $dl^2 = dx^2 + dy^2 = \frac{L^{e2}}{4} d\xi^2$, i.e. $dl = \frac{L^e}{2} d\xi$, and it is possible to write the contribution of each skin element as follows:

$$\begin{aligned} h \int_0^{L^e} N_i^e N_j^e dl &= \frac{hL^e}{2} \int_{-1}^1 N_i(\xi) N_j(\xi) d\xi \\ hT_{\text{ext}} \int_0^{L^e} N_i^e dl &= \frac{hL^e T_{\text{ext}}}{2} \int_{-1}^1 N_i(\xi) d\xi \end{aligned}$$

This makes it possible to fully define the element matrices and load vectors for each element e as follows:

– if element e has no common edge with Γ_2 , then:

$$[K^e] = \lambda S^e \begin{bmatrix} a_1^2 + b_1^2 & a_1 a_2 + b_1 b_2 & a_1 a_3 + b_1 b_3 \\ a_1 a_2 + b_1 b_2 & a_2^2 + b_2^2 & a_2 a_3 + b_2 b_3 \\ a_1 a_3 + b_1 b_3 & a_2 a_3 + b_2 b_3 & a_3^2 + b_3^2 \end{bmatrix}$$

$$\{F^e\} = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}$$

– if element e has a common edge with Γ_2 , then this edge is carried by the locally numbered nodes 2 and 3, and we obtain:

$$[K^e] = \lambda S^e \begin{bmatrix} a_1^2 + b_1^2 & a_1 a_2 + b_1 b_2 & a_1 a_3 + b_1 b_3 \\ a_1 a_2 + b_1 b_2 & a_2^2 + b_2^2 + \frac{hL^e}{3\lambda S^e} & a_2 a_3 + b_2 b_3 + \frac{hL^e}{6\lambda S^e} \\ a_1 a_3 + b_1 b_3 & a_2 a_3 + b_2 b_3 + \frac{hL^e}{6\lambda S^e} & a_3^2 + b_3^2 + \frac{hL^e}{3\lambda S^e} \end{bmatrix}$$

$$\{F^e\} = hL^e T_{\text{ext}} \begin{Bmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{Bmatrix}$$

2.4.2.2. Assembly

Figure 2.10 illustrates the assembly operation for the calculation of the components $K_{13,13}$ and $K_{3,8}$ of the global matrix. The local node numbering method described in Figure 2.7 was used.

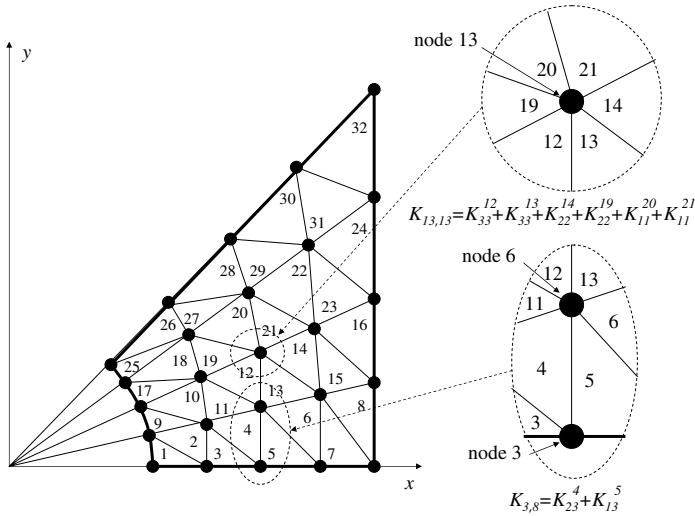


Figure 2.10. Working example: assembling the terms, $K_{13,13}$ and $K_{3,8}$, of the global matrix

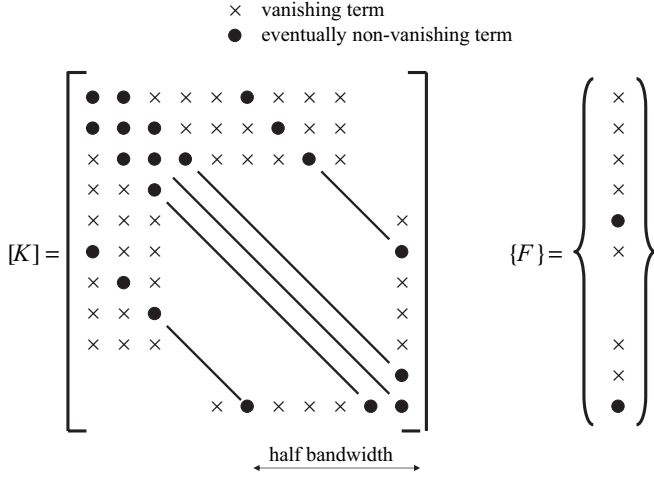


Figure 2.11. Working example: assembling the global matrix and load vector

By performing the assembly with $p = 4$ (as in Figure 2.7), we finally obtain a global matrix and load vector whose terms different from zero are given in Figure 2.11. In this figure, we can note that the half bandwidth is equal to 6.

2.4.3. Solution

2.4.3.1. Application of boundary conditions

In order to apply the boundary conditions, we will only use the penalty method here. According to this method, a high constant P is added to the diagonal components of the matrix $[K]$ corresponding to the nodes located on the boundary Γ_1 . Likewise, a constant PT_{int} is added to the components of the load vector $\{F\}$ referring to the same nodes. In the case $p = 4$ these nodes have numbers 1, 6, 11, 16 and 21. The application of the essential boundary conditions then involves the following modifications:

$$\begin{aligned}
 K_{1,1} &\longrightarrow K_{1,1} + P & F_1 &\longrightarrow F_1 + PT_{\text{int}} \\
 K_{6,6} &\longrightarrow K_{6,6} + P & F_6 &\longrightarrow F_6 + PT_{\text{int}} \\
 K_{11,11} &\longrightarrow K_{11,11} + P & F_{11} &\longrightarrow F_{11} + PT_{\text{int}} \\
 K_{16,16} &\longrightarrow K_{16,16} + P & F_{16} &\longrightarrow F_{16} + PT_{\text{int}} \\
 K_{21,21} &\longrightarrow K_{21,21} + P & F_{21} &\longrightarrow F_{21} + PT_{\text{int}}
 \end{aligned}$$

It may be noted that the global matrix $[K]$ described in Figure 2.11 cannot be inverted. In fact, without any temperature boundary conditions, there is an infinite

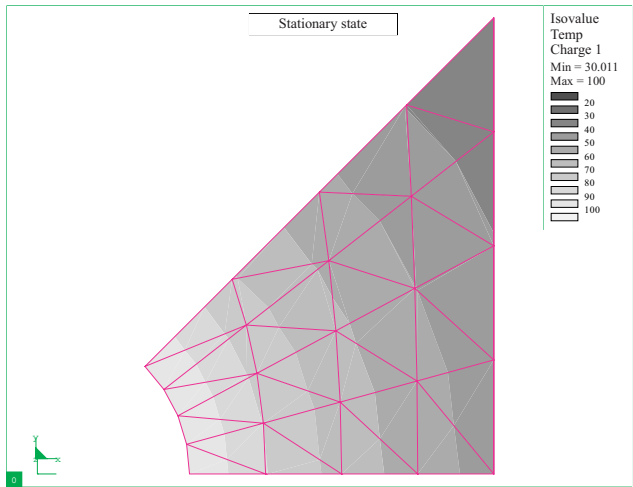


Figure 2.12. Drawing isothermatures

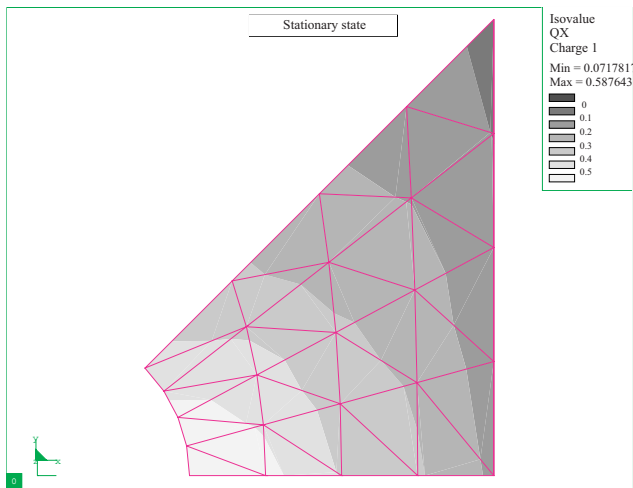


Figure 2.13. Drawing isoflux values in the direction x

number of solutions, all of them being separate from each other with a uniform temperature field (T constant in Ω). In this example, the temperature boundary conditions are therefore necessary to obtain a single solution.

2.4.3.2. *Solution*

This section requires the creation of a short computer program incorporating the methods described in this chapter. This program is easy to write. Therefore, we invite the reader to do this, as an exercise, in order to check the temperature field obtained. We have used existing software to perform this thermal analysis. Figures 2.12 and 2.13 have been made with the following data:

- geometry: $R = 2 \text{ mm}$, $a = 6 \text{ mm}$;
- thermal conductivity: $\lambda = 0.03 \text{ W/mm/K}$;
- boundary conditions: $T_{\text{int}} = 100^\circ\text{C}$, $T_{\text{ext}} = 20^\circ\text{C}$, $H = 0.01 \text{ W/m}^2/\text{K}$.

Figure 2.12 gives the temperature field, while Figure 2.13 gives the thermal flux in direction x .

This page intentionally left blank

Chapter 3

Isoparametric Finite Elements

3.1. Definitions

In Chapter 2, we introduced the major principles of finite element approximation leading to an approximate solution of continuous problem [1.13] by solving an equation system. This approximate solution is defined by a set of nodal values gathered in the vector $\{T\}$. The temperature at any point of the mesh is then obtained by interpolation inside each element (equation [2.2]).

This section concerns the description of isoparametric elements which are by far the most widely used elements in most computer codes. In order to introduce them, the concept of a reference element is presented. This notion makes it possible to define any element by means of a simple geometric transformation. The isoparametric elements use a geometric transformation defined by the element node location and the reference element's shape functions. Finally there will be a presentation of the properties required by the interpolation functions so that the approximate solution obtained by the finite element method leads to the continuous solution when the mesh is refined.

3.1.1. *Reference element*

The construction of the shape functions of complex geometric elements is not always straightforward. For convenience, the notion of a reference element is used. A reference element is a very simple element, indexed in a reference space, which can be transformed into an actual element by means of a simple geometric transformation (Figure 3.1).

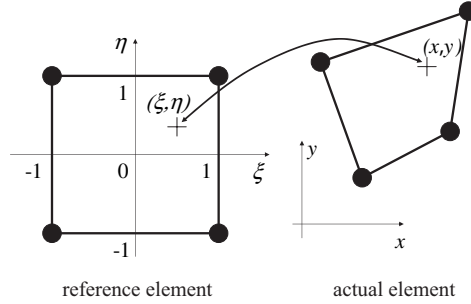


Figure 3.1. Reference element and actual element

The geometric transformation of a reference element is defined by means of the actual element's geometric nodes. It is *a priori* appropriate to differentiate the geometric nodes from the interpolation nodes of the element. Indeed, the geometric nodes of an element characterize its shape and location in the mesh while the interpolation nodes carry the problem unknowns (the temperatures in the present instance).

In practice, it is common to use a linear geometric transformation in relation to the coordinates of the geometric nodes. Thus, if element e includes p^e geometric nodes with coordinates \vec{x}_i^e (with $i = 1, \dots, p^e$), then the position \vec{x} of any point of the element will be obtained from the position $\vec{\xi}$ of the corresponding point in the reference element as follows:

$$\vec{x}(\vec{\xi}) = \sum_{i=1}^{p^e} G_i^e(\vec{\xi}) \vec{x}_i^e \quad [3.1]$$

The functions G_i^e are called geometric transformation functions. They take a scalar value and act on the coordinates $\vec{\xi}$ of a point on the reference element.

The geometric transformation serves the following purposes:

- to be bijective; which means that the Jacobian of the transformation must keep a constant sign and never be equal to zero inside an element;
- match the geometric nodes of the reference and actual elements. This is true with the condition $G_i^e(\vec{\xi}_j) = 1$ if $i = j$ and 0 if $i \neq j$, where $\vec{\xi}_j$ represents the coordinates of the geometric node j in the reference element;
- match the boundaries of the reference and actual elements.

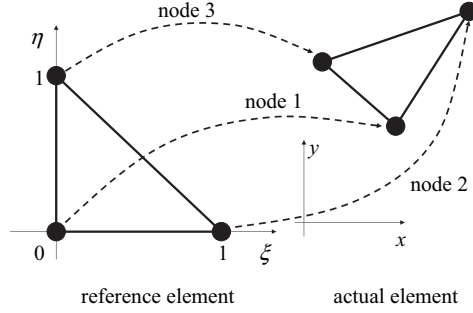


Figure 3.2. “Linear triangle” element

The similarity of relationship [3.1] with relationship [2.2] is obvious. Relationship [3.1] can thus be considered to be a nodal approximation of the geometry.

Now examine in detail three examples of elements constructed from a reference element. For simplification purposes, we will examine the 2D case only. The reader will find the characteristics of other elements at the end of this chapter.

3.1.1.1. *Triangular element with linear transformation functions*

Figure 3.2 represents the case of a triangle element. Three geometric nodes are used. In this case, the coordinates of a point are written $\vec{\xi} = (\xi, \eta)$ in the reference space and $\vec{x} = (x, y)$ in the actual space. Consider the geometric transformation functions defined by the three top nodes:

$$G_1^e(\xi, \eta) = 1 - \xi - \eta$$

$$G_2^e(\xi, \eta) = \xi$$

$$G_3^e(\xi, \eta) = \eta$$

Then the coordinates (x, y) of any point in the actual element can be obtained from the nodes 1 (x_1, y_1) , 2 (x_2, y_2) and 3 (x_3, y_3) in the form:

$$x(\xi, \eta) = (1 - \xi - \eta)x_1 + \xi x_2 + \eta x_3$$

$$y(\xi, \eta) = (1 - \xi - \eta)y_1 + \xi y_2 + \eta y_3$$

In these expressions, ξ and η represent the coordinates of the corresponding point in the reference element (Figure 3.1). It is easy to see that this transformation

matches the nodes and boundaries of the reference and actual elements. Furthermore, the Jacobian J is written:

$$J = \begin{vmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} \end{vmatrix} = \begin{vmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{vmatrix}$$

$$= (x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1)$$

Note that the Jacobian transformation is constant. It is the component according to Oz of the cross product between $\vec{x}_2 - \vec{x}_1$ and $\vec{x}_3 - \vec{x}_1$. Therefore, this Jacobian will not be equal to zero as long as the three geometric nodes of the triangle are not aligned.

3.1.1.2. *Quadrangle element with linear transformation functions*

Figure 3.3 represents the case of a quadrangle using four geometric nodes. The geometric transformation functions selected are:

$$G_1^e(\xi, \eta) = \frac{1}{4}(1 - \xi)(1 - \eta)$$

$$G_2^e(\xi, \eta) = \frac{1}{4}(1 + \xi)(1 - \eta)$$

$$G_3^e(\xi, \eta) = \frac{1}{4}(1 + \xi)(1 + \eta)$$

$$G_4^e(\xi, \eta) = \frac{1}{4}(1 - \xi)(1 + \eta)$$

Then the coordinates (x, y) of any point in the actual element are obtained from the nodes 1 (x_1, y_1) , 2 (x_2, y_2) , 3 (x_3, y_3) and 4 (x_4, y_4) in the form:

$$x(\xi, \eta) = \frac{1}{4}(1 - \xi)(1 - \eta)x_1 + \frac{1}{4}(1 + \xi)(1 - \eta)x_2$$

$$+ \frac{1}{4}(1 + \xi)(1 + \eta)x_3 + \frac{1}{4}(1 - \xi)(1 + \eta)x_4$$

$$y(\xi, \eta) = \frac{1}{4}(1 - \xi)(1 - \eta)y_1 + \frac{1}{4}(1 + \xi)(1 - \eta)y_2$$

$$+ \frac{1}{4}(1 + \xi)(1 + \eta)y_3 + \frac{1}{4}(1 - \xi)(1 + \eta)y_4$$

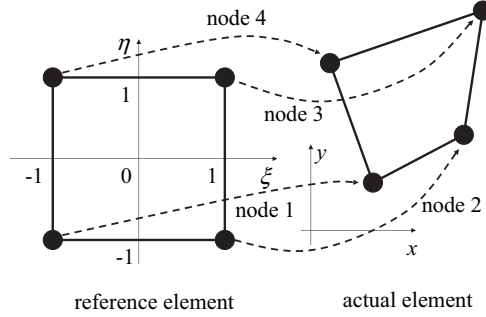


Figure 3.3. “Linear quadrangle” element

In these expressions, ξ and η represent the coordinates of the corresponding point in the reference element (Figure 3.1). Here again, note that this transformation matches the geometric nodes and sides of the reference and actual elements.

Following on from the results obtained for the linear triangle, it is easy to demonstrate that, if V_i is the component according to Oz of the cross product between $\vec{x}_j - \vec{x}_i$ and $\vec{x}_k - \vec{x}_i$, where nodes j and k are next to node i obtained by anticlockwise and clockwise rotation respectively in the reference element, then the Jacobian transformation is written as follows:

$$J(\xi, \eta) = \sum_{i=1}^4 G_i^e(\xi, \eta) V_i$$

Thus, given the definition of the geometric transformation functions, the Jacobian will keep a constant sign in the whole actual element if the V_i all have the same sign. In Figure 3.3, the geometric nodes of the reference and actual elements are numbered in the trigonometric sense. All the V_i are therefore positive since all the actual quadrangle angles are strictly between 0° and 180° . The corresponding Jacobian is therefore positive in all of the actual element. It is never equal to zero and the transformation is therefore bijective.

Conversely, if one of the quadrangle angle values had been greater than 180° , then the corresponding term V_i would have been negative, and the Jacobian sign would have changed inside the actual element. This is illustrated by Figure 3.4. In this case the transformation is no longer bijective and the element cannot be used. The angles of a linear quadrangle must therefore range strictly from 0° to 180° .

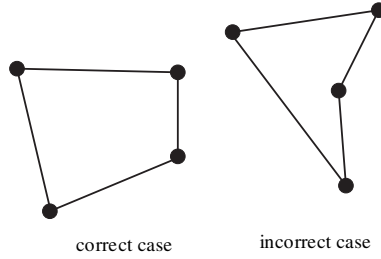


Figure 3.4. Condition of non-singularity of the quadrangle element

3.1.1.3. Quadrangle element with quadratic transformation functions

Figure 3.5 represents the case of a quadrangle using eight geometric nodes. The geometric transformation functions selected are:

$$G_1^e(\xi, \eta) = -\frac{1}{4}(1 - \xi)(1 - \eta)(1 + \xi + \eta)$$

$$G_2^e(\xi, \eta) = \frac{1}{2}(1 - \xi^2)(1 - \eta)$$

$$G_3^e(\xi, \eta) = -\frac{1}{4}(1 + \xi)(1 - \eta)(1 - \xi + \eta)$$

$$G_4^e(\xi, \eta) = \frac{1}{2}(1 + \xi)(1 - \eta^2)$$

$$G_5^e(\xi, \eta) = -\frac{1}{4}(1 + \xi)(1 + \eta)(1 - \xi - \eta)$$

$$G_6^e(\xi, \eta) = \frac{1}{2}(1 - \xi^2)(1 + \eta)$$

$$G_7^e(\xi, \eta) = -\frac{1}{4}(1 - \xi)(1 + \eta)(1 + \xi - \eta)$$

$$G_8^e(\xi, \eta) = \frac{1}{2}(1 - \xi)(1 - \eta^2)$$

This geometric transformation does match the nodes and sides of the reference and actual elements. Indeed, consider for instance in Figure 3.5 the straight segment of the reference element formed by nodes 1, 2 and 3. Any point on this segment is defined by $-1 \leq \xi \leq 1$ and $\eta = -1$. The only functions different from zero are G_1^e , G_2^e and G_3^e . Restricted to this segment the geometric transformation becomes:

$$x(\xi) = -\frac{1}{2}\xi(1 - \xi)x_1 + (1 - \xi^2)x_2 + \frac{1}{2}\xi(1 + \xi)x_3$$

$$y(\xi) = -\frac{1}{2}\xi(1 - \xi)y_1 + (1 - \xi^2)y_2 + \frac{1}{2}\xi(1 + \xi)y_3$$

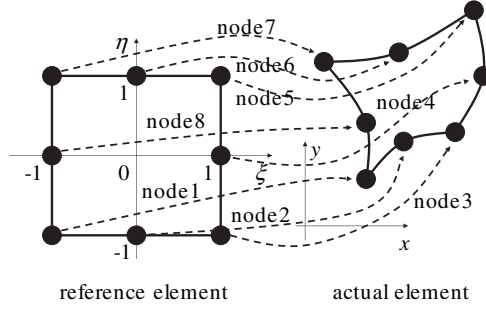


Figure 3.5. *Quadrangle element with quadratic transformation functions*

In the equations, (x_1, y_1) , (x_2, y_2) and (x_3, y_3) are the coordinates of node numbers 1, 2 and 3 of the actual element. The straight segment of the reference element is transformed into a curve parameterized by ξ giving the coordinates (x, y) of the points of this curve in the actual element.

3.1.2. Isoparametric elements

We have just examined the geometric construction of a finite element from a reference element. This construction is based upon the definition of a bijective transformation between the reference element and the actual element from the location of geometric nodes. This transformation is defined by means of geometric transformation functions. Through this geometric transformation, it is now possible to substitute the approximation on the actual element for the approximation on the reference element. By calling $N_i^e(\vec{\xi})$ the shape functions of the reference element (i varying from 1 to the number of interpolation nodes), the approximation inside the actual element is completely defined by:

– the bijective geometric transformation connecting any point \vec{x} of the actual element to a single point $\vec{\xi}$ of the reference element as follows:

$$\vec{x}(\vec{\xi}) = \sum_{i=1}^{p^e} G_i^e(\vec{\xi}) \vec{x}_i^e$$

– the nodal approximation of the temperature in the reference element, which is written as follows:

$$T(\vec{\xi}) = \sum_{i=1}^{n^e} N_i^e(\vec{\xi}) T_i^e$$

This approach is obviously interesting. As the reference elements are geometrically very simple (triangle, square, cube unit, etc.), it is easy to determine the shape functions $N_i^e(\vec{\xi})$ as well as the geometric transformation functions $G_i^e(\vec{\xi})$.

The elements most commonly used in computer codes use geometric nodes and geometric transformation functions respectively similar to the reference element interpolation nodes and shape functions. They are termed isoparametric elements. Other types of elements are sometimes used. Subparametric elements have fewer geometric nodes than interpolation nodes. On the contrary, superparametric elements have more geometric nodes than interpolation nodes.

Later on in this chapter as well as in the whole book, we will deal with isoparametric elements only. In this case, the number of geometric and interpolation nodes in an element e are equal, $p^e = n^e$, and the geometric transformation functions of the reference element are such that $G_i^e(\vec{\xi}) = N_i^e(\vec{\xi})$. The nodal temperature approximation (equation [2.2]) can then be written for any mesh element Ω^e as follows:

$$\begin{aligned} \forall \vec{x} \in \Omega^e, \quad T(\vec{x}) &= \sum_{i=1}^{n^e} N_i^e(\vec{x}(\vec{\xi})) T_i^e \\ &= \left\langle N_i^e(\vec{x}(\vec{\xi})) \right\rangle \cdot \{T^e\} \end{aligned}$$

Actually, the $N_i^e(\vec{x}(\vec{\xi}))$ can be formulated $N_i^e(\vec{\xi})$. However, it is appropriate to bear in mind the fact that the interpolation functions on the reference element and on the actual element are different. They can be deduced from one another by changing the variables between the coordinates \vec{x} and $\vec{\xi}$.

For instance, in the case of an eight-noded quadrangle, the interpolation functions N_i^e will be written for the reference element as follows:

$$\begin{aligned} N_1^e(\xi, \eta) &= -\frac{1}{4}(1 - \xi)(1 - \eta)(1 + \xi + \eta) \\ N_2^e(\xi, \eta) &= \frac{1}{2}(1 - \xi^2)(1 - \eta) \\ N_3^e(\xi, \eta) &= -\frac{1}{4}(1 + \xi)(1 - \eta)(1 - \xi + \eta) \\ N_4^e(\xi, \eta) &= \frac{1}{2}(1 + \xi)(1 - \eta^2) \\ N_5^e(\xi, \eta) &= -\frac{1}{4}(1 + \xi)(1 + \eta)(1 - \xi - \eta) \end{aligned}$$

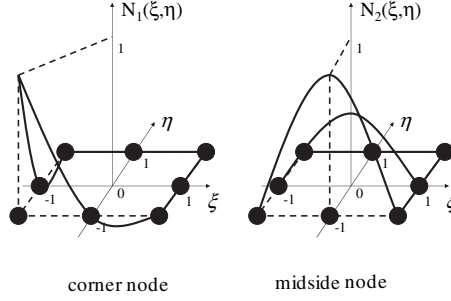


Figure 3.6. Interpolation functions of a reference eight-noded quadrangle

$$N_6^e(\xi, \eta) = \frac{1}{2}(1 - \xi^2)(1 + \eta)$$

$$N_7^e(\xi, \eta) = -\frac{1}{4}(1 - \xi)(1 + \eta)(1 + \xi - \eta)$$

$$N_8^e(\xi, \eta) = \frac{1}{2}(1 - \xi)(1 - \eta^2)$$

These functions are represented in Figure 3.6 for both types of node present in the element. Thus, a corner node (node 1 for instance) is distinguished from an midside node located in the middle of a segment (node 2 for instance). Note that interpolation functions vanish for nodes different from the one they are attached to, and more generally, on all the sides of the element which does not include this node.

It is important to note at this point that the location of midside nodes in a second order (or upper) isoparametric element may substantially modify the temperature approximation on the actual element.

For instance, in the case of a 1D second-order element, the three nodes are defined in the reference element by ξ equal to -1 , 0 and 1 . Given the shape functions of this element (see Table 3.3), the temperature approximation in the reference element is written in the form of a second degree polynomial in ξ :

$$\forall \xi \in [-1, 1], \quad T(\xi) = T_2 + \xi \frac{T_3 - T_1}{2} + \xi^2 \frac{T_1 - 2T_2 + T_3}{2}$$

By choosing an index centered on the actual element and writing $x_1 = -a$ and $x_3 = a$, the geometric transformation from the reference element to the actual element is written (Figure 3.7):

$$\forall \xi \in [-1, 1], \quad x(\xi) = x_2(1 - \xi^2) + a\xi$$

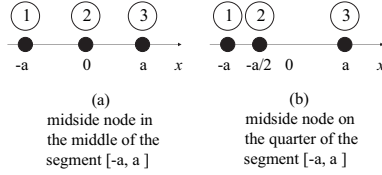


Figure 3.7. Locating the midside node of a 1D three-noded element in the middle (a) or on the quarter (b) of the segment

Now examine the following situations:

- the midside node (number 2) is located in the middle of the segment (Figure 3.7a), i.e. $x_2 = 0$,
- the midside node (number 2) is located on the quarter of the segment (Figure 3.7b), i.e. $x_2 = -a/2$.

In the first instance ($x_2 = 0$), the geometric transformation from the reference element to the actual element is defined by $x = \xi a$. Proportionality between x and ξ then leads to approximating the temperature on the actual element in the form of a second degree polynomial in x :

$$\forall x \in [-a, a], \quad T(x) = T_2 + x \frac{T_3 - T_1}{2a} + x^2 \frac{T_1 - 2T_2 + T_3}{2a^2}$$

In the second instance, the geometric transformation becomes $x = (\xi^2 + 2\xi - 1)\frac{a}{2}$. The inversion of this relationship gives two possible solutions for ξ according to x . We retain the formula with values of ξ going from -1 to 1 , i.e. $\xi = -1 + \sqrt{2 + 2x/a}$. This expression leads to an approximation of the temperature on the actual element as follows:

$$\begin{aligned} \forall x \in [-a, a], \quad T(x) = T_1 + (-3T_1 + 4T_2 - T_3) \sqrt{\frac{x+a}{2a}} \\ + (2T_1 - 4T_2 + 2T_3) \frac{x+a}{2a} \end{aligned}$$

This expression now involves such terms as $\sqrt{(x+a)/2a}$. This temperature approximation is now very different from that obtained when the midside node is located in the middle of the segment. In this case, the temperature gradient obtained by derivation with respect to x , and hence the heat flux density, tends to infinity as one gets near node 1 (x tending to $-a$).

Therefore, it is necessary to be very careful when locating midside nodes on the sides of elements. Especially with second-order elements, we will try to locate the middle node on the perpendicular bisector of the segment connecting the corner nodes. Another location for the midside node must be restricted to very specific cases. For example, in linear elastic fracture mechanics, it is possible to reproduce the crack tip singular stress field by using a mesh radiating from this tip and moving the midside nodes to the quarter of the segments connecting the corner nodes.

3.1.3. *Interpolation function properties*

The finite element method makes it possible to obtain an approximate solution of problem [1.13] in the form of a vector $\{T\}$ including the temperatures on the mesh nodes. Then the temperature at any point is obtained by nodal approximation. A few convergence conditions of the approximate solution to the continuous solution will be given hereafter.

For the approximate solution to converge toward the solution of continuous problem [1.13] when the number of nodes increases, it is necessary for the nodal temperature approximation defined by equation [2.2] to belong to the set E_T of admissible solutions of the continuous problem, which means that:

- inside each element, the temperature field nodal approximation, hence the interpolation functions, have to be sufficiently regular. In our case, this approximation must be class at least C^1 , that is to say such that its first derivatives are continuous;

- on the boundary between two elements, the nodal temperature field approximation must be continuous. The finite elements satisfying this condition are said to conform or compatible. Some elements, said to be incompatible, contravene this condition. The integral form of the problem then includes additional terms showing the function jumps on the element boundaries. The convergences of the approximate solution toward the continuous solution can only be ensured if these additional terms are bound and tend to zero when the element size itself tends to zero. This type of element is fairly often found in mechanics (some shell elements in particular), but very rarely in thermal science. Finally note that the continuous condition for compatible elements is valid for the temperature field only, not for its gradient. As a general rule, the temperature gradient is discontinuous on the boundary between two elements;

- as the mesh is refined, the approximate temperature solution and its first derivatives must tend toward constants (the values taken by the continuous solution). As in practice polynomial interpolation functions are used, this condition means that the approximate polynomial basis inside each element has to be complete at order 1,

that is to say it includes terms in x in a 1D problem, in x and y in 2D, and in x , y and z in 3D. In fact, using a complete polynomial basis of the order $p + 1$ ensures the convergence of the temperature field derivatives up to the order p .

The first condition is easily satisfied by choosing polynomial interpolation functions while the second condition will be satisfied by choosing interpolation functions N_i^e equal to zero on any side or face of the reference element not including node i .

The third condition forces the approximate polynomial basis to be complete at the order 1, which means that inside each element Ω^e the nodal approximation can be written at the first order next to a point in the form $T(\vec{x}) = T_0 + \vec{a} \cdot \vec{x}$. If this development is applied to the node with coordinates \vec{x}_i^e , we obtain $T_i^e = T_0 + \vec{a} \cdot \vec{x}_i^e$. If in the nodal approximation the node temperatures T_i^e are substituted for the expression above, it is possible to write simultaneously:

$$\left\{ \begin{array}{l} T(\vec{x}(\vec{\xi})) = T_0 + \vec{a} \cdot \vec{x}(\vec{\xi}) \\ T(\vec{x}(\vec{\xi})) = \sum_{i=1}^{n^e} N_i^e(\vec{\xi}) T_i^e \\ \quad = T_0 \sum_{i=1}^{n^e} N_i^e(\vec{\xi}) + \vec{a} \cdot \sum_{i=1}^{n^e} N_i^e(\vec{\xi}) \vec{x}_i^e \\ \quad = T_0 \sum_{i=1}^{n^e} N_i^e(\vec{\xi}) + \vec{a} \cdot \vec{x}(\vec{\xi}) \end{array} \right.$$

Therefore, the third condition will be satisfied as soon as:

$$\sum_{i=1}^{n^e} N_i^e(\vec{\xi}) = 1$$

It is possible to mix several types of finite elements in the same analysis. However, these elements must be compatible to satisfy the temperature continuity condition.

For instance, it is not advised to combine isoparametric elements whose faces or segments have different node numbers, as illustrated in Figure 3.8.

3.2. Calculation of element quantities

The matrices and element vectors obtained by equations [2.6], [2.7] and [2.8] are written in the form of integrals over the volume Ω^e or surface $\partial\Omega^e$ of an actual

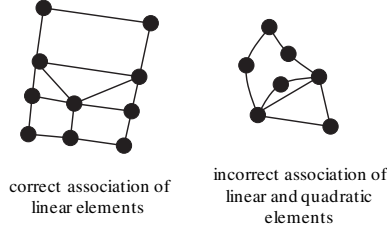


Figure 3.8. *Mixing isoparametric elements*

element e . In order to calculate these integrals, it is convenient to work on the reference element with the geometric transformation making it possible to move from the actual coordinates \vec{x} to the reference coordinates $\vec{\xi}$. For this purpose the Jacobian matrix of this transformation \underline{J} is used; it is defined from its tangent linear application to a given point:

$$\vec{x} = \vec{x}(\vec{\xi}) \implies d\vec{x} = \underline{J} \cdot d\vec{\xi} \quad \text{with} \quad \underline{J} = \underline{\text{grad}}_{\vec{\xi}}(\vec{x})$$

In this equation, the operator $\underline{\text{grad}}_{\vec{\xi}}()$ designates the gradient in the coordinate system $\vec{\xi}$. For example, in a Cartesian 3D case, written as follows: $\vec{x} = (x, y, z)$ and $\vec{\xi} = (\xi, \eta, \zeta)$, we obtain:

$$\underline{J} = \begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} & \frac{\partial x}{\partial \zeta} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} & \frac{\partial y}{\partial \zeta} \\ \frac{\partial z}{\partial \xi} & \frac{\partial z}{\partial \eta} & \frac{\partial z}{\partial \zeta} \end{bmatrix}$$

Firstly, we will express the element quantities in the reference frame related to the element e under consideration. Then we will note that these quantities can be numerically evaluated in a very accurate way by means of the Gauss method. Finally we will illustrate this method on 1D, 2D and 3D reference elements.

3.2.1. *Expression in the reference frame*

The fundamental quantity to calculate during a finite element analysis is the element residual $\{R^e\}$ (equation [2.6]). This quantity is the result of the integration on element Ω^e or on its boundary $\partial\Omega^e$ of functions involving:

- the thermal conductivity matrix $\underline{\lambda}$;
- the heat production volumetric term Q ;

- the flux q entering through the boundary $\partial\Omega_q$;
- the geometric characteristics of the element through its interpolation functions N_i^e and their gradients $\overrightarrow{\text{grad}}(N_i^e)$.

In order to turn an integral on Ω^e into an integral on the reference element, it is appropriate to transport these quantities by means of the (bijective) geometric transformation connecting the coordinates \overrightarrow{x} of a point on Ω^e to the coordinates $\overrightarrow{\xi}$ of the corresponding point in the reference element.

The transport of those quantities which are explicit functions of the location \overrightarrow{x} of the point on Ω^e is easy to perform. It is sufficient to replace \overrightarrow{x} with its expression according to $\overrightarrow{\xi}$ (equation [3.1] applied to an isoparametric element). It is, however, more difficult to transport the interpolation function gradients. It is necessary to use the Jacobian transformation matrix. In fact the derivation rule of composed functions allows us to write, in the case of Cartesian coordinates, by designating $\overrightarrow{x} = (x, y, z)$ and $\overrightarrow{\xi} = (\xi, \eta, \zeta)$:

$$\underbrace{\left(\frac{\partial N_i^e}{\partial \xi} \quad \frac{\partial N_i^e}{\partial \eta} \quad \frac{\partial N_i^e}{\partial \zeta} \right)}_{\overrightarrow{\text{grad}}_{\xi}^T(N_i^e)} = \underbrace{\left(\frac{\partial N_i^e}{\partial x} \quad \frac{\partial N_i^e}{\partial y} \quad \frac{\partial N_i^e}{\partial z} \right)}_{\overrightarrow{\text{grad}}^T(N_i^e)} \cdot \underbrace{\begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} & \frac{\partial x}{\partial \zeta} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} & \frac{\partial y}{\partial \zeta} \\ \frac{\partial z}{\partial \xi} & \frac{\partial z}{\partial \eta} & \frac{\partial z}{\partial \zeta} \end{bmatrix}}_J$$

In general, the terms involved in the expression of element quantities and including the gradient operator can be written as follows:

$$\begin{cases} \overrightarrow{\text{grad}}^T(N_i^e) = \overrightarrow{\text{grad}}_{\xi}^T(N_i^e) \cdot \underline{J}^{-1} \\ \overrightarrow{\text{grad}}(N_i^e) = \underline{J}^{-T} \cdot \overrightarrow{\text{grad}}_{\xi}(N_i^e) \end{cases} \quad [3.2]$$

In this equation, \underline{J}^{-1} is the inverse Jacobian matrix, and \underline{J}^{-T} is the transposed inverse Jacobian matrix.

The last step for transporting element quantities toward the reference element is the transformation of the volume element dv . If dv_0 designates the volume element in the reference element coordinate system, then we obtain:

$$dv = J dv_0$$

where J is the Jacobian matrix determinant, also called the Jacobian.

With these results, the element residual $\{R^e\}$ can be expressed for each element e , in the form of integrals on the volume or surfaces of the reference element Ω^0 corresponding to it. For instance, the linear case in the previous chapter, in which the element residual was expressed linearly according to the element node temperatures (equation [2.10]), corresponded to a heat conductance matrix $\underline{\lambda}$ and a volumetric term Q independent of the temperature, and a flux q appears as follows: $q = q_0 + h(T_{\text{ext}} - T)$. This residual becomes:

$$\{R^e\} = \left[\underbrace{\left(\int_{\Omega^0} \{N^e\} Q J dv_0 + \int_{\partial\Omega_q^0} \{N^e\} (q_0 + hT_{\text{ext}}) J_s ds_0 \right)}_{\{F^e\}} - \underbrace{\left(\int_{\Omega^0} \left\{ \overrightarrow{\text{grad}}_{\xi}^T(N^e) \right\} \cdot \underline{J}^{-1} \cdot \underline{\lambda} \cdot \underline{J}^{-T} \cdot \left\langle \overrightarrow{\text{grad}}_{\xi}(N^e) \right\rangle J dv_0 + \int_{\partial\Omega_q^0} \{N^e\} h \langle N^e \rangle J_s ds_0 \right)}_{[K^e]=[A^e]+[H^e]} \right] \cdot \{T^e\}$$

In this expression, Ω^0 represents the reference element corresponding to the actual element e , $\partial\Omega_q^0$ is the face of the reference element corresponding to that of the actual element on which the flux q is applied, ds_0 is the surface element corresponding to that face, and J_s is the determinant of that face transport Jacobian matrix.

Note therefore that the calculation of element quantities comes down to the calculation of integrals of type $\int_{\Omega^0} f(\vec{\xi}) dv_0$. Generally speaking, these integrals cannot be calculated analytically. Numerical integration techniques are then used.

3.2.2. Gaussian quadrature

Numerical integration consists of replacing an integral with a sum of contributions on a finite number n_g of points with coordinates $\vec{\xi}_k$ (with $k = 1, \dots, n_g$):

$$\int_{\Omega^0} f(\vec{\xi}) dv_0 \equiv \sum_{k=1}^{n_g} w_k f(\vec{\xi}_k)$$

In this equation, $\vec{\xi}_k$ locates the position of the integration point number k , and w_k is the weight of this point contribution to the integral value.

The Gauss method (Gaussian quadrature) is different from other methods essentially because both the integration point coordinates and weights are determined

to obtain maximum accuracy. They are calculated in order to minimize the approximation error in the integration of a polynomial function $p(\vec{\xi})$ interpolating the function $f(\vec{\xi})$ at those points.

3.2.2.1. 1D numerical integration

In the 1D case, the reference element is always characterized by ξ varying from -1 to 1 . The integration of a function $f(\vec{\xi})$ on the reference volume Ω^0 can then be written as:

$$I = \int_{\Omega^0} f(\vec{\xi}) dv_0 = \int_{-1}^1 f(\xi) d\xi$$

The evaluation of the integral I depends on the calculation of an approximating integral \bar{I} :

$$\bar{I} = \int_{-1}^1 p(\xi) d\xi$$

where $p(\xi)$ is a polynomial interpolating function $f(\xi)$ on n_g points. It is obvious that the approximation \bar{I} depends directly on the number and location of interpolation points of $f(\xi)$.

Assume first that the positions of the integration points are known, and write them ξ_1, \dots, ξ_{n_g} . The polynomial $p(\xi)$ interpolating the function $f(\xi)$ on these points can then be expressed in the form of a linear combination of Lagrange polynomials $L_k(\xi)$, satisfying the condition $L_k(\xi_j) = \delta_{kj}$:

$$p(\xi) = \sum_{k=1}^{n_g} L_k(\xi) f(\xi_k)$$

$$L_k(\xi) = \frac{(\xi - \xi_1) \cdots (\xi - \xi_{k-1})(\xi - \xi_{k+1}) \cdots (\xi - \xi_{n_g})}{(\xi_k - \xi_1) \cdots (\xi_k - \xi_{k-1})(\xi_k - \xi_{k+1}) \cdots (\xi_k - \xi_{n_g})}$$

If we now insert this polynomial expression into the approximating integral \bar{I} , we obtain a formula in which the weights w_k of the contribution of each integration point ξ_k to the integral \bar{I} can be calculated analytically according to the locations of these points:

$$\bar{I} = \sum_{k=1}^{n_g} \underbrace{\left(\int_{-1}^1 L_k(\xi) d\xi \right)}_{w_k} f(\xi_k)$$

The *a priori* choice of the integration point locations ξ_k can be made in various ways. For instance, in the Newton-Cotes integration method, the points ξ_k are evenly distributed in the form $\xi_k = -1 + 2 \frac{k-1}{n_g-1}$ on the segment $[-1, 1]$. It is demonstrated that this method, like all those *a priori* setting values ξ_k , allows us to integrate accurately any polynomial function $f(\xi)$ of a degree less than or equal to $n_g - 1$.

In the Gauss method, the position ξ_k of the integration points is also optimized to ensure a more accurate estimation. In this case, we therefore obtain $2n_g$ parameters to optimize (n_g positions ξ_k and n_g weights w_k), which makes it possible to integrate accurately with the same number of points any polynomial function $f(\xi)$ of a degree less than or equal to $2n_g - 1$. For instance, a 3-degree polynomial will be accurately integrated with 2 points by the Gauss method, whereas 4 will be necessary with the Newton-Cotes method. This method makes it possible to obtain the same accuracy by substantially reducing the number of points where the function to integrate has to be estimated, which reduces the calculation time.

In order to calculate the optimal positions of the integration points ξ_k , the expression of the interpolation polynomial $p(\xi)$ is enhanced with n_g additional terms in the form:

$$\begin{cases} p(\xi) = \sum_{k=1}^{n_g} L_k(\xi) f(\xi_k) + \sum_{k=1}^{n_g} \alpha_k \xi^{k-1} q(\xi) \\ q(\xi) = \prod_{l=1}^{n_g} (\xi - \xi_l) \end{cases}$$

Note in this equation that $p(\xi)$ is now a polynomial of degree $2n_g - 1$ with $2n_g$ parameters (the ξ_k and the α_k). Moreover, this polynomial always coincides with function $f(\xi)$ at the points ξ_k , since $q(\xi_k) = 0$.

Now the integral \bar{I} approximating I is written as follows:

$$\bar{I} = \sum_{k=1}^{n_g} \left(\int_{-1}^1 L_k(\xi) d\xi \right) f(\xi_k) + \sum_{k=1}^{n_g} \left(\int_{-1}^1 \xi^{k-1} q(\xi) d\xi \right) \alpha_k$$

Since the integration is to be exact for any polynomial of a degree less than or equal to $2n_g - 1$, the integral \bar{I} must be independent of the coefficients α_k , which leads us to choose the ξ_k so that:

$$\forall l = 1, \dots, n_g, \quad \int_{-1}^1 \xi^{l-1} q(\xi) d\xi = 0$$

For instance, if we decide to choose two integration points ξ_1 and ξ_2 , then the location of these points will be determined with $q(\xi) = (\xi - \xi_1)(\xi - \xi_2)$:

$$\begin{cases} \int_{-1}^1 (\xi - \xi_1)(\xi - \xi_2) d\xi = 0 \\ \int_{-1}^1 \xi(\xi - \xi_1)(\xi - \xi_2) d\xi = 0 \end{cases} \Rightarrow \begin{cases} \frac{2}{3} + 2\xi_1\xi_2 = 0 \\ \frac{2}{3}(\xi_1 + \xi_2) = 0 \end{cases} \Rightarrow \begin{cases} \xi_1 = -\frac{\sqrt{3}}{3} \\ \xi_2 = \frac{\sqrt{3}}{3} \end{cases}$$

The calculation of weights w_1 and w_2 corresponding to the integration points then generates:

$$\begin{cases} w_1 = \int_{-1}^1 L_1(\xi) d\xi = \int_{-1}^1 \frac{\xi - \xi_2}{\xi_1 - \xi_2} d\xi = 1 \\ w_2 = \int_{-1}^1 L_2(\xi) d\xi = \int_{-1}^1 \frac{\xi - \xi_1}{\xi_2 - \xi_1} d\xi = 1 \end{cases}$$

More generally speaking, Table 3.1 gives the integration point locations and corresponding weights for the Gaussian quadratures mostly used in computer codes.





n_g	positions	ξ_k	w_k
1		0.0	2.0
2		$\pm 0.57735\ 02691\ 89626$	1.0
3		-0.77459 66692 41483 0.0 +0.77459 66692 41483	0.55555 55555 55556 0.88888 88888 88889 0.55555 55555 55556
4		$\pm 0.86113\ 63115\ 94053$ $\pm 0.33998\ 10435\ 84856$	0.34785 48451 37454 0.65214 51548 62546

Table 3.1. 1D Gauss integration points and weights

The number of integration points in an element depends on the complexity of the functions to integrate on that element for the calculation of element quantities. For instance, for a 1D two-noded element, the interpolation functions are linear (see Table 3.3) and a single integration point will be sufficient to accurately integrate the element quantities (as well as their gradients, of course). However, the element quantities

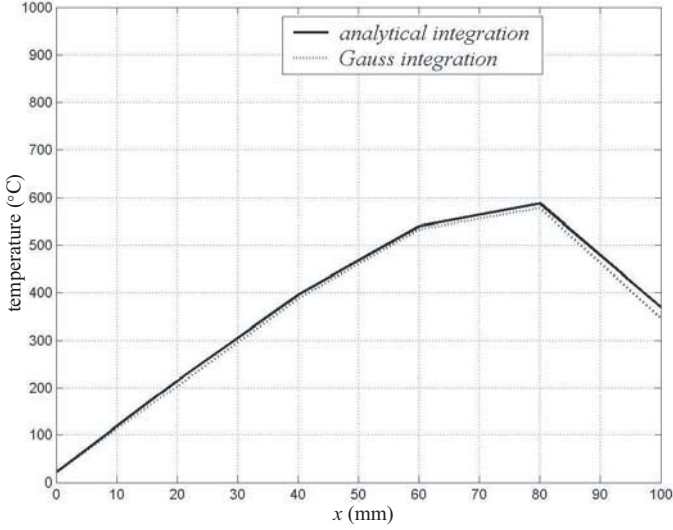


Figure 3.9. Gauss method induced error

require other functions such as the volumetric heat source Q . If these functions do not vary linearly according to the space coordinates, then this method will introduce an error in the estimation of these element quantities. Figure 3.9 illustrates this error in the working example of Chapter 1, with a discretization on five finite elements.

3.2.2.2. 2D and 3D numerical integration

Extending the Gauss integration method to 2D or 3D elements is relatively easy with square or cubic geometries. In such a case, the integration directions are independent and it is possible to write:

$$\int_{-1}^1 \int_{-1}^1 f(\xi, \eta) d\xi d\eta \approx \sum_{k=1}^{n_g^\xi} \sum_{l=1}^{n_g^\eta} w_k^\xi w_l^\eta f(\xi_k, \eta_l)$$

$$\int_{-1}^1 \int_{-1}^1 \int_{-1}^1 f(\xi, \eta, \zeta) d\xi d\eta d\zeta \approx \sum_{k=1}^{n_g^\xi} \sum_{l=1}^{n_g^\eta} \sum_{m=1}^{n_g^\zeta} w_k^\xi w_l^\eta w_m^\zeta f(\xi_k, \eta_l, \zeta_m)$$

The point numbers $(n_g^\xi, n_g^\eta, n_g^\zeta)$, their coordinates (ξ_k, η_l, ζ_m) and their weights $(w_k^\xi, w_l^\eta, w_m^\zeta)$ are obtained from the results of section 3.2.2.1 applied in the corresponding direction.

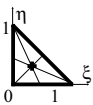
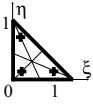
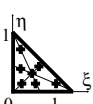
n_g	positions	ξ_k	η_k	w_k
1		$\xi_1=0.33333\ 33333\ 33333$	$\eta_1 = \xi_1$	$w_1=0.5$
3		$\xi_1=0.16666\ 66666\ 667$ $\xi_2=0.66666\ 66666\ 667$ $\xi_3 = \xi_1$	$\eta_1 = \xi_1$ $\eta_2 = \xi_1$ $\eta_3 = \xi_2$	$w_1=0.16666\ 66666\ 667$ $w_2 = w_1$ $w_3 = w_1$
7		$\xi_1=0.10128\ 65073\ 235$ $\xi_2=0.79742\ 69853\ 531$ $\xi_3 = \xi_1$ $\xi_4=0.47014\ 20641\ 051$ $\xi_5 = \xi_4$ $\xi_6=0.05971\ 58717\ 898$ $\xi_7=0.33333\ 33333\ 333$	$\eta_1 = \xi_1$ $\eta_2 = \xi_1$ $\eta_3 = \xi_2$ $\eta_4 = \xi_6$ $\eta_5 = \xi_4$ $\eta_6 = \xi_4$ $\eta_7 = \xi_7$	$w_1=0.06296\ 95902\ 724$ $w_2 = w_1$ $w_3 = w_1$ $w_4=0.066197\ 07639\ 425$ $w_5 = w_4$ $w_6 = w_4$ $w_7=0.1125$

Table 3.2. Gauss integration schemes for triangles

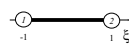
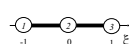
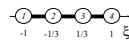
Type	Geometry	Interpolation functions
Linear element with two nodes		$N_1(\xi) = \frac{1 - \xi}{2}$ $N_2(\xi) = \frac{1 + \xi}{2}$
Quadratic element with three nodes		$N_1(\xi) = \frac{1}{2}\xi(\xi - 1)$ $N_2(\xi) = 1 - \xi^2$ $N_3(\xi) = \frac{1}{2}\xi(\xi + 1)$
Cubic element with four nodes		$N_1(\xi) = \frac{9}{16}(\xi - 1)\left(\frac{1}{9} - \xi^2\right)$ $N_2(\xi) = \frac{27}{16}(\xi^2 - 1)\left(\xi - \frac{1}{3}\right)$ $N_3(\xi) = \frac{27}{16}(1 - \xi^2)\left(\xi + \frac{1}{3}\right)$ $N_4(\xi) = \frac{9}{16}(\xi + 1)\left(\xi^2 - \frac{1}{9}\right)$

Table 3.3. Most common 1D finite elements

For other types of elements (triangle, tetrahedron, wedge, etc.), it is impossible to act in the same way. A great deal of research work has enabled us to determine several numerical schemes, the most common of which appear in Table 3.2 for a triangle. These schemes can then be combined to construct new schemes adapted to tetrahedra and wedges (Table 3.5).

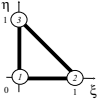
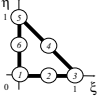
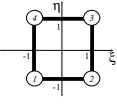
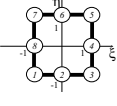
Type	Geometry	Interpolation functions
Linear element with three nodes (triangle)		$N_1(\xi, \eta) = 1 - \xi - \eta$ $N_2(\xi, \eta) = \xi$ $N_3(\xi, \eta) = \eta$
Quadratic element with six nodes (triangle)		$N_1(\xi, \eta) = -\alpha(1 - 2\alpha)$ $N_2(\xi, \eta) = 4\xi\alpha$ $N_3(\xi, \eta) = -\xi(1 - 2\xi)$ $N_4(\xi, \eta) = 4\xi\eta$ $N_5(\xi, \eta) = -\eta(1 - 2\eta)$ $N_6(\xi, \eta) = 4\eta\alpha$ (with $\alpha = 1 - \xi - \eta$)
Linear element with four nodes (quadrangle)		$N_1(\xi, \eta) = \frac{1}{4}(1 - \xi)(1 - \eta)$ $N_2(\xi, \eta) = \frac{1}{4}(1 + \xi)(1 - \eta)$ $N_3(\xi, \eta) = \frac{1}{4}(1 + \xi)(1 + \eta)$ $N_4(\xi, \eta) = \frac{1}{4}(1 - \xi)(1 + \eta)$
Quadratic element with eight nodes (quadrangle)		$N_1(\xi, \eta) = -\frac{1}{4}(1 - \xi)(1 - \eta)(1 + \xi + \eta)$ $N_2(\xi, \eta) = \frac{1}{2}(1 - \xi^2)(1 - \eta)$ $N_3(\xi, \eta) = -\frac{1}{4}(1 + \xi)(1 - \eta)(1 - \xi + \eta)$ $N_4(\xi, \eta) = \frac{1}{2}(1 + \xi)(1 - \eta^2)$ $N_5(\xi, \eta) = -\frac{1}{4}(1 + \xi)(1 + \eta)(1 - \xi - \eta)$ $N_6(\xi, \eta) = \frac{1}{2}(1 - \xi^2)(1 + \eta)$ $N_7(\xi, \eta) = -\frac{1}{4}(1 - \xi)(1 + \eta)(1 + \xi - \eta)$ $N_8(\xi, \eta) = \frac{1}{2}(1 - \xi)(1 - \eta^2)$

Table 3.4. Most common 2D finite elements

3.3. Some finite elements

In this section, we will deal only with the most commonly used reference elements, for which the geometry, local node numbering and interpolation functions are specified. The 1D elements are in Table 3.3. The four-noded cubic element is hardly used. The 2D elements are in Table 3.4. Some of them are widely used in this chapter (the three-noded triangle for instance). Finally, Table 3.5 includes a few 3D elements. There are others, such as the 15-noded wedge or the 20-noded cube (quadratic

interpolation). There are also elements mixing a quadratic interpolation in one or two directions with a linear interpolation in the remaining directions. The interpolation functions are then obtained by applying the relationships presented in this chapter.




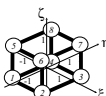
Type	Geometry	Interpolation functions
Linear element with four nodes (tetrahedron)		$N_1(\xi, \eta, \zeta) = 1 - \xi - \eta - \zeta$ $N_2(\xi, \eta, \zeta) = \xi$ $N_3(\xi, \eta, \zeta) = \eta$ $N_4(\xi, \eta, \zeta) = \zeta$
Quadratic element with ten nodes (tetrahedron)		$N_1(\xi, \eta, \zeta) = -\alpha(1 - 2\alpha)$ $N_2(\xi, \eta, \zeta) = 4\xi\alpha$ $N_3(\xi, \eta, \zeta) = -\xi(1 - 2\xi)$ $N_4(\xi, \eta, \zeta) = 4\xi\eta$ $N_5(\xi, \eta, \zeta) = -\eta(1 - 2\eta)$ $N_6(\xi, \eta, \zeta) = 4\eta\alpha$ $N_7(\xi, \eta, \zeta) = 4\zeta\alpha$ $N_8(\xi, \eta, \zeta) = 4\xi\zeta$ $N_9(\xi, \eta, \zeta) = 4\eta\zeta$ $N_{10}(\xi, \eta, \zeta) = -\zeta(1 - 2\zeta)$ (with $\alpha = 1 - \xi - \eta - \zeta$)
Linear element with six nodes (wedge)		$N_1(\xi, \eta, \zeta) = \frac{1}{2}(1 - \xi - \eta)(1 - \zeta)$ $N_2(\xi, \eta, \zeta) = \frac{1}{2}\xi(1 - \zeta)$ $N_3(\xi, \eta, \zeta) = \frac{1}{2}\eta(1 - \zeta)$ $N_4(\xi, \eta, \zeta) = \frac{1}{2}(1 - \xi - \eta)(1 + \zeta)$ $N_5(\xi, \eta, \zeta) = \frac{1}{2}\xi(1 + \zeta)$ $N_6(\xi, \eta, \zeta) = \frac{1}{2}\eta(1 + \zeta)$
Linear element with eight nodes (hexahedron)		$N_1(\xi, \eta, \zeta) = \frac{1}{8}(1 - \xi)(1 - \eta)(1 - \zeta)$ $N_2(\xi, \eta, \zeta) = \frac{1}{8}(1 + \xi)(1 - \eta)(1 - \zeta)$ $N_3(\xi, \eta, \zeta) = \frac{1}{8}(1 + \xi)(1 + \eta)(1 - \zeta)$ $N_4(\xi, \eta, \zeta) = \frac{1}{8}(1 - \xi)(1 + \eta)(1 - \zeta)$ $N_5(\xi, \eta, \zeta) = \frac{1}{8}(1 - \xi)(1 - \eta)(1 + \zeta)$ $N_6(\xi, \eta, \zeta) = \frac{1}{8}(1 + \xi)(1 - \eta)(1 + \zeta)$ $N_7(\xi, \eta, \zeta) = \frac{1}{8}(1 + \xi)(1 + \eta)(1 + \zeta)$ $N_8(\xi, \eta, \zeta) = \frac{1}{8}(1 - \xi)(1 + \eta)(1 + \zeta)$

Table 3.5. Most common 3D finite elements

PART 2

Transient State, Non-linearities, Transport Phenomena

This page intentionally left blank

Introduction

Part 1 of this book dealt with the detailed description of the finite element method using, to this end, the simple case of steady state conduction. This allowed us to introduce the basic physical models used in thermal science (thermal balance, Fourier law, etc.), then show how the finite element method can be used to solve a particular problem.

In Part 2, complementary notions will be introduced and their consequences for the finite element method will be shown. These concepts are:

- the transient state;
- non-linearities;
- transport phenomena.

The study of transient heat conduction requires consideration of the thermal capacitance of materials. Indeed, a solid undergoing a thermal load does not instantly reach a steady state. The temperature in the solid varies over time. In order to obtain this variation, the finite element method must be coupled with a time integration scheme.

The behavior of a material in the steady or transient state is rarely linear. As a matter of fact, its thermal conductivity is often temperature-dependent. Moreover, the boundary conditions and the loads applied to a solid are often non-linear. For example, the heat exchange coefficient between a solid and an external medium often depends on the local temperature of the solid surface. The traditional methods of solving non-linear thermal problems in the steady or transient state will be described.

Transport phenomena are present when the solid is no longer assumed to be motionless as a result of the thermal loads it is subjected to. Such is the case when

a welding operation is simulated, the heat source moving along the weld seal. A solution in a reference frame related to that source will have to take into consideration the relative movement of the solid. Transport phenomena also introduce numerical instabilities and the remedial methods for this will be described.

Chapter 4

Transient Heat Conduction

4.1. Problem formulation

Transient heat conduction involves time. In order to apply the finite element method to the numerical simulation of this phenomenon, we use a course of action similar to that introduced in Chapters 1 and 2. This method leads us to a time differential system for the temperatures at the mesh nodes.

4.1.1. *The continuous problem*

If we perform the energy balance as in [RAP 98] for instance, the left-hand term of the steady state equilibrium equation [1.4], i.e. $\text{div}(\underline{\lambda} \cdot \overrightarrow{\text{grad}}(T)) + Q$, appears to be the heat brought to a volume element of the solid Ω . The first principle of thermodynamics specifies that this heat produces an internal energy change on the domain over time. That internal energy change, equal to zero in the steady state, can be expressed here as $\rho \dot{H}$, where ρ is the material volumetric mass and \dot{H} the variation over time of its specific (i.e. per mass unit) enthalpy H . Equilibrium equation [1.4] thus becomes:

$$\text{div}(\underline{\lambda} \cdot \overrightarrow{\text{grad}}(T)) + Q = \rho \dot{H} \quad \text{at any point of } \Omega \quad [4.1]$$

In this case, the boundary conditions applied to the solid are similar to those mentioned in the steady state, i.e:

- an input flux q known on a portion $\partial\Omega_q$ of Ω boundary;
- a known temperature T_d on the other portion $\partial\Omega_T$ of that boundary.

However, in transient states, these boundary conditions may change over time, as well as the source term Q . It follows from this that the boundary value problem of transient heat conduction is expressed as:

$$\left. \begin{aligned}
 &\text{Find } T(\vec{x}, t) \text{ at any point } \vec{x} \in \Omega \text{ verifying:} \\
 &\text{at } t = 0: \\
 &T(\vec{x}, 0) = T_0(\vec{x}) \\
 &\text{at any instant } t \in]0, t_f[: \\
 &R(T, t) = \text{div}(\underline{\lambda} \cdot \overrightarrow{\text{grad}}(T)) + Q(T, t) - \rho \dot{H}(T, t) = 0 \text{ in } \Omega \\
 &\left. \begin{aligned}
 &(\underline{\lambda} \cdot \overrightarrow{\text{grad}}(T)) \cdot \vec{n} = q \text{ on } \partial\Omega_q \\
 &T = T_d \text{ on } \partial\Omega_T
 \end{aligned} \right\} \text{ with } \partial\Omega_q \cup \partial\Omega_T = \partial\Omega
 \end{aligned} \right\} \quad [4.2]$$

The term $R(T, t)$ present in this problem is called residual, as in the steady state. It is a quantity to cancel at any point of the solid Ω , and at any instant of the study, i.e. between instants $t = 0$ and $t = t_f$, while satisfying at these instants the boundary conditions illustrating the exchanges of the solid with the external environment. These boundary conditions are applied on the boundary $\partial\Omega$ of the solid Ω , and we must check that $\partial\Omega_q \cup \partial\Omega_T = \partial\Omega$.

Conversely, in this boundary value problem an initial condition appears, the initial temperature field $T_0(\vec{x})$ in the solid, as well as the enthalpy derivative with respect to time. We could therefore think of applying a time finite element approximation by using the initial conditions as boundary conditions. However, time variations are treated differently from space variations in most computer codes. The finite element approximation is only used for space variations, time integration requiring a finite difference technique like the one described in section 4.2.

If a method similar to that of Chapter 1 is applied to space variations, the following weak integral formulation is obtained, where the spaces E_T and E_ψ are those defined by equations [1.11] and [1.12]:

$$\left. \begin{aligned}
 &\text{Find } T: t \in [0, t_f[\longrightarrow T(\vec{x}, t) \in E_T \text{ such that:} \\
 &\text{at } t = 0: \\
 &T(\vec{x}, 0) = T_0(\vec{x}) \\
 &\text{for } t \in]0, t_f[\text{ and for any } \psi \in E_\psi: \\
 &\int_{\Omega} \psi(Q - \rho \dot{H}) dv + \int_{\partial\Omega_q} \psi q ds - \int_{\Omega} \overrightarrow{\text{grad}}^T(\psi) \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(T) dv = 0
 \end{aligned} \right\} \quad [4.3]$$

The Galerkin method constructs at any instant a temperature approximation by reasoning on n -dimensional spaces E_T^n and E_ψ^n . Continuous problem [4.3] is then approximated by the following discrete problem (of dimension n):

$$\begin{aligned}
 & \text{Find } T: t \in [0, t_f[\longrightarrow T(\vec{x}, t) \in E_T^n \text{ such that:} \\
 & \text{at } t = 0: \\
 & T(\vec{x}, 0) = T_0(\vec{x}) \\
 & \text{for } t \in]0, t_f[\text{ and for any } \psi \in E_\psi^n: \\
 & \int_{\Omega} \psi(Q - \rho \dot{H}) dv + \int_{\partial\Omega_q} \psi q ds - \int_{\Omega} \overrightarrow{\text{grad}}^T(\psi) \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(T) dv = 0
 \end{aligned} \tag{4.4}$$

4.1.2. Finite element approximation

In discrete problem [4.4], the temperature field $T(\vec{x}, t)$ is estimated inside each element e of a mesh, and at any instant t , from the values it takes at the nodes corresponding to that element, at that same instant, while the weighting functions $\psi(\vec{x})$ are estimated as in the steady state:

$$\begin{aligned}
 \forall \vec{x} \in \Omega^e, \forall t \in [0, t_f[, \quad T(\vec{x}, t) &= \sum_{i=1}^{n^e} N_i^e(\vec{x}) T_i^e(t) \\
 &= \langle N_1^e(\vec{x}), \dots, N_{n^e}^e(\vec{x}) \rangle \cdot \begin{Bmatrix} T_1^e(t) \\ \vdots \\ T_{n^e}^e(t) \end{Bmatrix} \\
 &= \langle N^e(\vec{x}) \rangle \cdot \{T^e(t)\}
 \end{aligned} \tag{4.5}$$

$$\begin{aligned}
 \forall \vec{x} \in \Omega^e, \forall t \in [0, t_f[, \quad \psi(\vec{x}) &= \sum_{i=1}^{n^e} \psi_i^e N_i^e(\vec{x}) \\
 &= \langle \psi_1^e, \dots, \psi_{n^e}^e \rangle \cdot \begin{Bmatrix} N_1^e(\vec{x}) \\ \vdots \\ N_{n^e}^e(\vec{x}) \end{Bmatrix} \\
 &= \langle \psi^e \rangle \cdot \{N^e(\vec{x})\}
 \end{aligned} \tag{4.6}$$

In equations [4.5] and [4.6], $T_i^e(t)$ and ψ_i^e respectively represent the temperature at instant t , and the weighting function value at the node number i (local numbering) of element e . The shape functions introduced, $N_i^e(\vec{x})$, are estimated similarly to those described in the previous chapters.

If nodal approximation [4.6] is applied to the weighting functions $\psi(\vec{x})$ in the discrete problem [4.4], the quantities involved in that problem can be written as follows:

$$\begin{aligned}\int_{\Omega} \psi(Q - \rho\dot{H})dv &= \sum_{e=1}^m \langle \psi^e \rangle \cdot \int_{\Omega^e} \{N^e\} (Q - \rho\dot{H})dv \\ \int_{\partial\Omega_q} \psi q ds &= \sum_{e=1}^m \langle \psi^e \rangle \cdot \int_{\partial\Omega^e \cap \partial\Omega_q} \{N^e\} q ds \\ \int_{\Omega} \overrightarrow{\text{grad}}^T(\psi) \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(T) dv &= \sum_{e=1}^m \langle \psi^e \rangle \cdot \int_{\Omega^e} \left\{ \overrightarrow{\text{grad}}^T(N^e) \right\} \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(T) dv\end{aligned}$$

If these expressions are transferred into equation [4.4], the quantity to cancel becomes:

$$\left| \begin{aligned} & \int_{\Omega} \psi(Q - \rho\dot{H})dv + \int_{\partial\Omega_q} \psi q ds \\ & - \int_{\Omega} \overrightarrow{\text{grad}}^T(\psi) \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(T) dv \end{aligned} \right. = \sum_{e=1}^m \langle \psi^e \rangle \cdot \{R^e(T, t)\}$$

where the vector $\{R^e(T, t)\}$, specific to each element e , is called the element residual, and is written according to the temperature field $T(\vec{x}, t)$ and time t as:

$$\{R^e(T, t)\} = \left| \begin{aligned} & \int_{\Omega^e} \{N^e\} (Q - \rho\dot{H})dv + \int_{\partial\Omega^e \cap \partial\Omega_q} \{N^e\} q ds \\ & - \int_{\Omega^e} \left\{ \overrightarrow{\text{grad}}^T(N^e) \right\} \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(T) dv \end{aligned} \right. \quad [4.7]$$

In order to estimate the element residual, it is necessary to express the variation rate of the material enthalpy, \dot{H} , according to temperature and time. In many instances, the material enthalpy depends on temperature only. The material specific heat $C = \frac{dH}{dT}$ can then be introduced to write $\rho\dot{H} = \rho C\dot{T}$. Term \dot{T} can then be approximated by a nodal approximation (equation [4.5]) as:

$$\begin{aligned}\forall \vec{x} \in \Omega^e, \forall t \in [0, t_f[, \quad \dot{T}(\vec{x}, t) &= \sum_{i=1}^{n^e} N_i^e(\vec{x}) \dot{T}_i^e \\ &= \langle N^e(\vec{x}) \rangle \cdot \{\dot{T}^e\}\end{aligned} \quad [4.8]$$

In this equation, the vector $\{\dot{T}^e\}$ includes in a column the temperature change rates at the nodes of element e . If these expressions are inserted into the element residual (equation [4.7]), we obtain:

$$\{R^e(T, t)\} = \begin{vmatrix} \int_{\Omega^e} \{N^e\} Q \, dv + \int_{\partial\Omega^e \cap \partial\Omega_q} \{N^e\} q \, ds \\ - \left(\int_{\Omega^e} \rho C \{N^e\} \langle N^e \rangle \, dv \right) \cdot \{\dot{T}^e\} \\ - \int_{\Omega^e} \left\{ \overrightarrow{\text{grad}}^T(N^e) \right\} \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(T) \, dv \end{vmatrix}$$

Note that this expression differs from the steady state (equation [2.6] in Chapter 2) only by adding a term representing the quantity $\rho \dot{H}$ into the thermal balance (equation [4.1]). That term produces the definition of a new element matrix, called the capacitance matrix, heat capacity matrix or element mass matrix. It is a symmetric matrix which is written as:

$$[C^e] = \int_{\Omega^e} \rho C \{N^e\} \langle N^e \rangle \, dv$$

Further on in this chapter, as in Part 1, we will restrict our study to the linear case to fully express the element residual and its assembly. Non-linearities will be addressed in the next chapter.

4.1.3. Linear case

If the results of Part 1 are used, we see that the element residual will be expressed linearly according to the problem unknowns if:

- the material heat conductivity λ does not depend on the temperature;
- the term Q , the heat volumetric source, is independent of the temperature (in fact, as in the steady state, it could depend on it linearly, but as far as we know, this corresponds to no physical instance);
- the heat flux q applied to the solid boundary $\partial\Omega_q$ is a linear function of the temperature (i.e. $q = h(T_{\text{ext}} - T) + q_0$, where h and q_0 are independent of the temperature);
- the material volumetric mass ρ and specific heat C are independent of the temperature.

This element residual will then be expressed as:

$$\{R^e(T)\} = \{F^e\} - [C^e] \cdot \{\dot{T}^e\} - [K^e] \cdot \{T^e\}$$

The element terms: load vector $\{F^e\}$, capacitance matrix $[C^e]$ and conductance matrix $[K^e]$ are written as follows:

$$\{F^e\} = \int_{\Omega} \{N^e\} Q dv + \int_{\partial\Omega^e \cap \partial\Omega_q} \{N^e\} (q_0 + hT_{\text{ext}}) ds \quad [4.9]$$

$$[C^e] = \int_{\Omega^e} \rho C \{N^e\} \langle N^e \rangle dv \quad [4.10]$$

$$[K^e] = [\Lambda^e] + [H^e] \text{ with: } \begin{cases} [\Lambda^e] = \int_{\Omega^e} \left\{ \overrightarrow{\text{grad}}^T(N^e) \right\} \cdot \underline{\lambda} \cdot \left\langle \overrightarrow{\text{grad}}(N^e) \right\rangle dv \\ [H^e] = \int_{\partial\Omega^e \cap \partial\Omega_q} \{N^e\} h \langle N^e \rangle ds \end{cases} \quad [4.11]$$

Finally it is possible to assemble the element quantities as in the steady case, which allows the global problem of the following nodal residual to cancel:

$$\{R\} = \{F\} - [C] \cdot \left\{ \dot{T} \right\} - [K] \cdot \{T\} \quad [4.12]$$

In this equation, the capacitance and conductance global matrices $[C]$ and $[K]$, as well as the load vector $\{F\}$, are obtained by traditional assembly operations:

$$\left\{ \begin{array}{l} \{F\} = \sum_{e=1}^m [A^e]^T \cdot \{F^e\} \\ [C] = \sum_{e=1}^m [A^e]^T \cdot [C^e] \cdot [A^e] \\ [K] = \sum_{e=1}^m [A^e]^T \cdot [K^e] \cdot [A^e] \end{array} \right. \quad [4.13]$$

Solving a transient heat conduction problem leads us to solve a first order differential system with respect to time:

$$[C] \cdot \left\{ \dot{T} \right\} + [K] \cdot \{T\} = \{F\} \quad [4.14]$$

It is appropriate to note at this point that this differential equation does not take account of the essential boundary conditions (known temperature), which are to be inserted as the solution is in progress. Inserting these boundary conditions can be done in the same way as in the steady case. We will not go into the details here.

4.2. Time integration

The integration of differential system [4.14] gives the time evolution of the temperatures at the mesh nodes and, by means of a nodal approximation, the temperature and its gradient at any point of the solid. To do so, it is simple to use a time-stepping solution method for [4.14] by means, for example, of an implicit finite difference scheme. The temperature change rate at the mesh nodes is written as:

$$\left\{ \dot{T}(t + \Delta t) \right\} = \frac{\{T(t + \Delta t)\} - \{T(t)\}}{\Delta t}$$

The column vector $\{F\}$ and the matrices $[C]$ and $[K]$ being known at instant $t + \Delta t$, as well as the nodal temperatures at instant t (i.e. vector $\{T(t)\}$), the nodal temperatures at instant $t + \Delta t$ are obtained by solving differential system [4.14] at $t + \Delta t$. This yields:

$$([C] + \Delta t [K]) \cdot \{T(t + \Delta t)\} = \Delta t \{F\} + [C] \cdot \{T(t)\}$$

Finally, when using such a method, we start with an initial condition at $t = 0$ (i.e. temperature $\{T(0)\}$), then estimate the solution at each instant by consecutive time steps Δt . This is a space approximate solution (by the finite element method) and a time approximate solution (by the finite difference method). It should be noted that the time step Δt may change during the analysis process.

We will now describe other numerical methods to integrate this system over time.

4.2.1. Modal method

The modal method seeks a solution to the transient heat conduction problem as follows:

$$T(\vec{x}, t) = \sum_{k=1}^n \beta_k(t) \tau_k(\vec{x}) \quad [4.15]$$

where the n functions $\tau_k(\vec{x})$ form a basis of the space of admissible solutions (E_T^n).

Applying the finite element method then leads us to look for a solution to the differential system of the form:

$$\{T(t)\} = \sum_{k=1}^n \beta_k(t) \{\tau_k\} = \langle \{\tau_1\}, \dots, \{\tau_n\} \rangle \cdot \begin{Bmatrix} \beta_1(t) \\ \vdots \\ \beta_n(t) \end{Bmatrix} \quad [4.16]$$

where the n vectors $\{\tau_k\}$ are a basis of \mathcal{R}^n , called the modal basis. The modal method determines a modal basis then expresses the system to solve on that basis to obtain the solution in that basis.

4.2.1.1. Determining the modal basis

To determine the modal basis, differential system [4.14] with no load vector is considered:

$$[C] \cdot \{\dot{T}\} + [K] \cdot \{T\} = 0 \quad [4.17]$$

We are then tempted to seek solutions of the form $\{T(t)\} = \{\tau\} e^{-\alpha t}$, so that this system becomes:

$$([K] - \alpha [C]) \cdot \{\tau\} = 0 \quad [4.18]$$

Matrix $[C]$, being symmetric positive definite, can be decomposed as $[C] = [L] \cdot [L]^T$, where $[L]$ is a lower triangular matrix. It would have been possible to choose to decompose matrix $[K]$ as well. Choosing $[C]$ here is preferable since $[K]$ is not necessarily defined. This decomposition produces the system $([K] - \alpha [L] \cdot [L]^T) \cdot \{\tau\} = 0$ which, multiplied on the left by $[L]^{-1}$, yields:

$$([L]^{-1} \cdot [K] \cdot [L]^{-T} - \alpha [I]) [L]^T \cdot \{\tau\} = 0$$

In this equation, $[L]^{-T}$ represents the transpose of the inverse of matrix $[L]$. We therefore go back to the solution of an eigenvalue problem whose general form is written as:

$$([A] - \alpha [I]) \cdot \{\Psi\} = 0 \text{ with } \begin{cases} [A] = [L]^{-1} \cdot [K] \cdot [L]^{-T} \\ [I] : \text{identity matrix} \\ \{\Psi\} = [L]^T \cdot \{\tau\} \end{cases} \quad [4.19]$$

Matrix $[A]$ being symmetric, this problem admits n real eigenvalues $\alpha_1, \dots, \alpha_k, \dots, \alpha_n$. To each eigenvalue α_k corresponds an eigenvector $\{\Psi_k\}$ which can be written as $\{\Psi_k\} = [L]^T \cdot \{\tau_k\}$. If we reverse the process which led us from system [4.18] to system [4.19], it is easy to demonstrate that the vectors $\{\tau_k\}$ satisfy the following condition:

$$[K] \cdot \{\tau_k\} = \alpha_k [C] \cdot \{\tau_k\} \quad [4.20]$$

Matrix $[C]$ being symmetric positive definite and matrix $[K]$ symmetric but only positive semi-definite, we can write:

$$\underbrace{\langle \tau_k \rangle \cdot [K] \cdot \{\tau_k\}}_{\geq 0} = \alpha_k \underbrace{\langle \tau_k \rangle \cdot [C] \cdot \{\tau_k\}}_{> 0}$$

We deduce that the n real eigenvalues α_k are positive or equal to zero:

$$0 \leq \alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n$$

To define the eigenvectors in a single way, it is necessary to norm them. Several norms are able to do so; the most traditional are:

- The infinite norm: $\text{Max}_{i=1 \dots n}(\tau_{ki})$.
- The $[C]$ -norm: $\langle \tau_k \rangle \cdot [C] \cdot \{\tau_k\}$.

Now consider 2 vectors $\{\tau_j\}$ and $\{\tau_k\}$, corresponding to 2 different eigenvalues $\alpha_j \neq \alpha_k$. It is then possible to determine equation [4.20] for each value and derive:

$$\begin{aligned} \langle \tau_k \rangle \cdot [K] \cdot \{\tau_j\} &= \alpha_j \langle \tau_k \rangle \cdot [C] \cdot \{\tau_j\} \\ \langle \tau_j \rangle \cdot [K] \cdot \{\tau_k\} &= \alpha_k \langle \tau_j \rangle \cdot [C] \cdot \{\tau_k\} \end{aligned}$$

Given the symmetry of the matrices $[K]$ and $[C]$, we can write:

$$\begin{aligned} \langle \tau_k \rangle \cdot [K] \cdot \{\tau_j\} &= \langle \tau_j \rangle \cdot [K] \cdot \{\tau_k\} \\ \langle \tau_k \rangle \cdot [C] \cdot \{\tau_j\} &= \langle \tau_j \rangle \cdot [C] \cdot \{\tau_k\} \end{aligned}$$

From the above equations, we can derive:

$$(\alpha_j - \alpha_k) \langle \tau_k \rangle \cdot [C] \cdot \{\tau_j\} = 0$$

then:

$$\langle \tau_k \rangle \cdot [K] \cdot \{\tau_j\} = \langle \tau_k \rangle \cdot [C] \cdot \{\tau_j\} = 0 \quad [4.21]$$

In the multiple mode case, i.e. if the eigenspace corresponding to an eigenvalue α_k is of a dimension strictly greater than 1, we can construct an orthogonal basis of that space in the sense of the scalar product $L(\{\tau_j\}, \{\tau_k\}) = \langle \tau_j \rangle \cdot [C] \cdot \{\tau_k\}$. It is obvious that equation [4.21] remains valid.

Equation [4.21] thus shows that the $\{\tau_k\}$ for $k = 1 \dots n$ form an orthonormal basis of \mathcal{R}^n called a modal basis.

We write:

$$\begin{aligned} c_j &= \langle \tau_j \rangle \cdot [C] \cdot \{\tau_j\} \\ k_j &= \langle \tau_j \rangle \cdot [K] \cdot \{\tau_j\} \end{aligned} \quad [4.22]$$

As in mechanics, we can call c_j and k_j respectively generalized capacitance and conductivity. Note that k_j and c_j verify $\alpha_j = \frac{k_j}{c_j}$.

The description of the calculation methods of eigenvalues and vectors is not dealt with in this book. It is possible for interested readers to consult [ZIE 91] or [BAT 96].

4.2.1.2. Projection on the modal basis

We therefore seek a solution of the differential system in the form given by [4.16], where $(\{\tau_k\})_{k=1 \dots n}$ is the modal basis. By projecting the differential system onto each vector $\{\tau_j\}$, we obtain:

$$\begin{aligned} \langle \tau_j \rangle \cdot [C] \cdot \sum_{k=1}^n \dot{\beta}_k(t) \{\tau_k\} + \langle \tau_j \rangle \cdot [K] \\ \cdot \sum_{k=1}^n \beta_k(t) \{\tau_k\} = \langle \tau_j \rangle \cdot \{F(t)\} \end{aligned}$$

Given [4.21] and [4.22], we then obtain n first order differential equations, each now including only one unknown β_j :

$$c_j \dot{\beta}_j(t) + k_j \beta_j(t) = f_j(t) \quad \text{with} \quad f_j(t) = \langle \tau_j \rangle \cdot \{F(t)\} \quad [4.23]$$

To integrate equation [4.23] over time, we can use one of the direct integration algorithms presented hereafter. It is equally possible to use the analytical solution of equation [4.23] which is written:

$$\beta_j(t) = e^{-\alpha_j(t-t_0)} \left(\beta_j(t_0) + \frac{1}{c_j} \int_{t_0}^t e^{\alpha_j(s-t_0)} f_j(s) ds \right) \quad [4.24]$$

where $\beta_j(t_0)$ is the initial solution obtained from the initial temperature by:

$$\beta_j(t_0) = \langle \tau_j \rangle \{T(t_0)\}$$

Most of the time, the right-hand side of [4.24] cannot be calculated analytically. However, it is possible to calculate a time-stepping solution by assuming that, on each time step, $f_j(t)$ of [4.23] varies linearly. For each time step, we obtain:

$$\begin{aligned} \beta_j(t + \Delta t) = & \beta_j(t)e^{-\alpha_j \Delta t} + \frac{f_j(t + \Delta t)}{k_j} \left(1 + \frac{e^{-\alpha_j \Delta t} - 1}{\alpha_j \Delta t} \right) \\ & - \frac{f_j(t)}{k_j} \left(e^{-\alpha_j \Delta t} + \frac{e^{-\alpha_j \Delta t} - 1}{\alpha_j \Delta t} \right) \end{aligned}$$

In practice, it is unnecessary to consider a modal basis including all the problem eigenmodes. Only the n' first modes corresponding to the lowest eigenvalues and allowing us to give an accurate representation of the temperature are necessary. The number of differential equations [4.23] is consequently reduced.

The use of this method, also designated modal superposition method, requires on the one hand the calculation of the n' first eigenmodes and eigenvalues and on the other hand the integration of n' uncoupled differential equations [4.23].

This method is widely used in mechanics, but fairly rarely in thermal science.

4.2.2. Direct time integration

We consider the following first order differential equation which we will try to integrate over time:

$$\dot{u}(t) = g(u, t) \quad [4.25]$$

In equation [4.25], if we refer to equation [4.23], we see that u represents the projection of the nodal temperature vector onto a vector $\{\tau_j\}$ of the modal basis, and $g(u, t) = \frac{1}{c_j} (f_j(t) - k_j u(t))$. In the general case of differential system [4.12], u represents the nodal temperature vector $\{T\}$ and g the vector $[C]^{-1} (\{F\} - [K] \cdot \{T\})$.

Direct integration algorithms make it possible to integrate an equation like [4.25] step by step over time. Thus, to calculate the quantities at instant $t + \Delta t$, we assume all the quantities at the previous instants (t , $t - \Delta t$, etc.) to be known and apply a numerical scheme to replace the differential equation with an algebraic equation whose unknowns are the velocities at $t + \Delta t$.

There is a large number of integration schemes of differential equation [4.25]. Among the most widely used are the generalized central difference algorithm and the generalized trapezoidal rule.

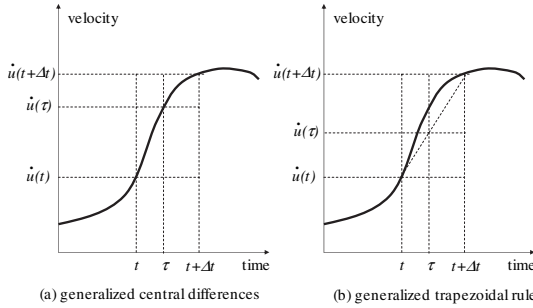


Figure 4.1. Time integration schemes: (a) generalized central difference method and (b) generalized trapezoidal rule

The generalized central difference algorithm (Figure 4.1a) assumes that the velocity \dot{u} is constant over each time interval $[t, t + \Delta t]$ and equal to the value at an intermediate instant $t + \nu\Delta t$ (with $0 \leq \nu \leq 1$):

$$\frac{u(t + \Delta t) - u(t)}{\Delta t} = \dot{u}(t + \nu\Delta t) = g(u(t + \nu\Delta t), t + \nu\Delta t) \quad [4.26]$$

In the generalized trapezoidal algorithm (Figure 4.1b), the velocity \dot{u} is also assumed to be constant, but is obtained as a weighted mean velocity of the velocities at times t and $t + \Delta t$. With $0 \leq \nu \leq 1$ we obtain:

$$\begin{aligned} \frac{u(t + \Delta t) - u(t)}{\Delta t} &= (1 - \nu)\dot{u}(t) + \nu\dot{u}(t + \Delta t) \\ &= (1 - \nu)g(u(t), t) + \nu g(u(t + \Delta t), t + \Delta t) \end{aligned} \quad [4.27]$$

Actually, these two families of numerical schemes are very close. They even coincide when the function g is a linear function of u , as is the case in linear conduction. In non-linear cases, the slight differences can be seen in Figure 4.1.

In the above numerical schemes, $u(t + \Delta t)$ is the unknown. It is obvious that if the right-hand side of [4.26] or [4.27] is independent of this unknown, then we directly (or explicitly) obtain $u(t + \Delta t) = u(t) + g(u(t), t)\Delta t$. Then the numerical scheme is called explicit.

For the numerical scheme families considered here, only the case $\nu = 0$ corresponds to an explicit scheme, designated as an Euler explicit scheme. In the case $\nu = 1$, we obtain an Euler implicit scheme. The intermediate algorithms corresponding to $0 < \nu < 1$ are said to be semi-implicit. The central difference

method (also called the Crank-Nicolson algorithm) corresponds to equation [4.26] with $\nu = 1/2$ and the trapezoidal rule to equation [4.27] with $\nu = 1/2$.

To apply the generalized central difference scheme to the transient heat conduction problem, we start again with equation [4.14] which is considered at instant $t + \nu\Delta t$:

$$[C_\nu] \cdot \{\dot{T}_\nu\} + [K_\nu] \cdot \{T_\nu\} = \{F_\nu\}$$

In the above equation, index ν is used to locate the quantities evaluated at instant $t + \nu\Delta t$. Unless otherwise indicated, this index will be used later on. In addition, index 0 will be used for quantities evaluated at t and index 1, for those evaluated at $t + \Delta t$.

The variable change rate $\{T\}$ being assumed to be constant over the time interval considered, we can write:

$$\{T_\nu\} = \{T(t + \nu\Delta t)\} = (1 - \nu) \{T_0\} + \nu \{T_1\}$$

In that expression, vector $\{T_0\}$ is known (it is the one at instant t), whereas vector $\{T_1\}$ is the one we are trying to determine. Inserting this expression into the differential equation to solve then gives the following equation linear system:

$$\left(\frac{1}{\Delta t} [C_\nu] + \nu [K_\nu] \right) \cdot \{T_1\} = \underbrace{\left(\frac{1}{\Delta t} [C_\nu] - (1 - \nu) [K_\nu] \right) \cdot \{T_0\} + \{F_\nu\}}_{G_\nu} \quad [4.28]$$

The solution of this linear system allows us to obtain $\{T_1\}$, i.e. the temperatures at the mesh nodes at instant $t + \Delta t$, from vector $\{T_0\}$, i.e. the temperatures at the mesh nodes at instant t . In the linear case being dealt with here, the other vectors and matrices involved are not time-dependent.

Now apply the generalized trapezoidal scheme to equation [4.14] considered at instant $t + \Delta t$:

$$[C_1] \cdot \{\dot{T}_1\} + [K_1] \cdot \{T_1\} = \{F_1\}$$

We obtain:

$$\left(\frac{1}{\Delta t} [C_1] + \nu [K_1] \right) \cdot \{T_1\} = \begin{vmatrix} \frac{1}{\Delta t} [C_1] \cdot \{T_0\} \\ + (1 - \nu) [C_1] \cdot \{\dot{T}_0\} \\ + \nu \{F_1\} \end{vmatrix} \quad [4.29]$$

The generalized trapezoidal scheme requires at each instant vector $\{T_0\}$ and vector $\{\dot{T}_0\}$ which is calculated gradually by means of formula [4.27]. At the initial instant, $\{\dot{T}_0\}$ is calculated by solving equation [4.14] at that instant, the initial temperature vector $\{T_0\}$ being assumed to be known.

Equations [4.28] and [4.29] do highlight the differences between both algorithm families, resulting mainly, in non-linear cases, from the fact that the quantities involved are not evaluated at the same instant ($t + \nu\Delta t$ for the generalized central difference algorithm and $t + \Delta t$ for the generalized trapezoidal algorithm).

Consider again equation [4.28] in the case of the Euler explicit scheme corresponding to $\nu = 0$. We then obtain the following system:

$$\frac{1}{\Delta t} [C_0] \cdot \{T_1\} = \{G_0\} \text{ with } \{G_0\} = \left(\frac{1}{\Delta t} [C_0] - [K_0] \right) \cdot \{T_0\} + \{F_0\} \quad [4.30]$$

The matrix to inverse is simply $[C_0]$, the capacitance matrix, resulting from the assembly of the element capacitance matrices $[C^e]$ calculated at instant t . The conductance matrix $[K_0]$ is not involved in this matrix to inverse, but only in the load vector.

The strict application of the finite element method produces a capacitance matrix $[C_0]$, symmetric and sparse, with non-diagonal terms different from zero. $[C_0]$ is called a consistent capacitance matrix (with the finite element formulation). However, in order to facilitate the inversion of the linear system obtained, matrix $[C_0]$ is made diagonal. This operation keeps the model's total capacitance while eliminating the coupling terms between two nodes. All the thermal capacitance is then transferred to the diagonal terms of the matrix. In fact, with this technique, the specific heat is no longer distributed evenly in the material volume but is lumped at the mesh nodes.

The most common lumping technique transfers on the diagonal the sum of the terms of each line:

$$[C_0] = \begin{bmatrix} C_{11} & C_{12} & \cdots \\ C_{21} & C_{22} & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}$$

$$\Rightarrow [C_0] \approx \begin{bmatrix} C_{11} + C_{12} + \cdots & 0 & 0 \\ 0 & C_{21} + C_{22} + \cdots & 0 \\ 0 & 0 & \cdots \end{bmatrix}$$

This approximation considerably simplifies the solution of linear system [4.30] as it is reduced to n scalar equations yielding the components of $\{T_1\}$ directly. The problem unknowns at instant $t + \Delta t$ are thus obtained explicitly without solving a linear system. Furthermore, storing the matrix only requires storing the n diagonal terms.

However, the explicit scheme has a major drawback. The algorithm is conditionally stable, i.e. the time step must never exceed a limit value. How to obtain this limit value is described in the next section. That is the reason why it is restricted to strongly non-linear problems, involving rapid phenomena already requiring, for physical reasons, very small time steps.

When an implicit or semi-implicit scheme ($0 < \nu \leq 1$) is used, the temperatures at the mesh nodes at instant $t + \Delta t$ (i.e. vector $\{T_1\}$) are obtained by solving equation system [4.28]. When the problem is linear, matrices $[C_\nu]$ and $[K_\nu]$ involved in the system are independent of the temperature. Therefore, they are constant over time, and the matrix to invert in equation system [4.28] is determined by the time step Δt . The assembly and triangulation of this matrix will only be necessary when the time step Δt is modified during the calculation process.

Implicit schemes are more complicated to process than explicit schemes. However, their stability conditions are less severe. This explains why they are frequently used in thermal science.

4.2.3. Accuracy and stability of a direct integration algorithm

In order to analyze the stability and accuracy of the integration schemes presented in section 4.2.2, we will cover the linear case only, and assume that the time step used, Δt , is constant. Thus, we can designate with index p the quantities estimated at instant $p\Delta t$. $\{T_0\}$ thus gives the temperatures at the mesh nodes at $t = 0$, $\{T_p\}$ the temperatures estimated at $t = p\Delta t$, and $\{T_{p+\nu}\}$ those estimated at $t = (p + \nu)\Delta t$, i.e. during the time interval, $[p\Delta t, (p + 1)\Delta t]$, with $0 \leq \nu \leq 1$. Besides, restricting the study to the linear case allows us to write matrices $[C_\nu]$ and $[K_\nu]$ as $[C]$ and $[K]$, since they are constant.

Now it is important to differentiate the accuracy and stability conditions of an algorithm:

- the algorithm accuracy is given by the truncation error made in formula [4.26] or [4.27];
- the stability is the capacity of an algorithm to reduce or amplify the errors made during the consecutive time steps.

4.2.3.1. Accuracy

The variation of a quantity u during a time step is approximated as:

$$\frac{u_{p+1} - u_p}{\Delta t} = \begin{cases} \dot{u}_{p+\nu} & \text{in the central difference method} \\ (1 - \nu)\dot{u}_p + \nu\dot{u}_{p+1} & \text{in the trapezoidal rule} \end{cases}$$

In order to evaluate the truncation error made with the generalized central difference method, the following Taylor series expansions are written:

$$\begin{cases} u_{p+1} = u_{p+\nu} + (1 - \nu)\Delta t \dot{u}_{p+\nu} + (1 - \nu)^2 \frac{\Delta t^2}{2} \ddot{u}_{p+\nu} + \mathcal{O}(\Delta t^3) \\ u_p = u_{p+\nu} - \nu \Delta t \dot{u}_{p+\nu} + \nu^2 \frac{\Delta t^2}{2} \ddot{u}_{p+\nu} + \mathcal{O}(\Delta t^3) \end{cases}$$

By subtracting these two expressions then dividing the result by Δt , we obtain:

$$\frac{u_{p+1} - u_p}{\Delta t} - \dot{u}_{p+\nu} = (1 - 2\nu) \frac{\Delta t}{2} \ddot{u}_{p+\nu} + \mathcal{O}(\Delta t^2)$$

Likewise, the truncation error made with the generalized trapezoidal rule can be obtained from the following Taylor series expansions:

$$\begin{cases} u_{p+1} = u_p + \Delta t \dot{u}_p + \frac{\Delta t^2}{2} \ddot{u}_p + \mathcal{O}(\Delta t^3) \\ \dot{u}_{p+1} = \dot{u}_p + \Delta t \ddot{u}_p + \mathcal{O}(\Delta t^2) \end{cases}$$

From these expressions we immediately find that:

$$\frac{u_{p+1} - u_p}{\Delta t} - ((1 - \nu)\dot{u}_p + \nu\dot{u}_{p+1}) = \left(\frac{1}{2} - \nu\right) \Delta t \ddot{u}_p + \mathcal{O}(\Delta t^2)$$

We can see that both families produce first order algorithms when $\nu \neq \frac{1}{2}$. The numerical schemes are said to be coherent at order 1 for differential equation [4.25]. This means that the errors in the velocity \dot{u} are in the order of Δt . When $\nu = \frac{1}{2}$, the algorithm is coherent at order 2 and the errors made are in the order of Δt^2 .

The conclusion of this might be that the Crank-Nicolson algorithm ($\nu = \frac{1}{2}$) is, in any case, the best possible choice. In fact, it is not so obvious, for, as will be seen later on, this scheme produces oscillating solutions when the time step is too great. With strong non-linearities, like those resulting from state changes for example, the Euler implicit algorithm is generally preferred.

4.2.3.2. Stability

We have just seen that at each time step, the solution obtained is erroneous. Imagine these errors aggregated in the case of Euler implicit algorithm ($\nu = 1$). We saw previously that this algorithm is coherent at order 1. Therefore, there is a constant such that:

$$\left| \frac{u_{p+1} - u_p}{\Delta t} - \dot{u}_{p+1} \right| \leq k\Delta t$$

The error would then be less than the product $Nk\Delta t$ where N represents the number of time steps performed.

Unfortunately, these are not the only errors that add up because numerical scheme [4.26] or [4.27], uses numerical values u_p which are also erroneous. The numerical scheme may therefore amplify the errors to the point of producing an unacceptable solution. We say that a numerical scheme is stable if the error amplification is bounded, i.e. if the very solution of the discretized problem is bounded. We can demonstrate that if a numerical scheme is coherent and stable, then the numerical solution converges toward the continuous solution when the time step tends to 0.

In order to analyze the algorithm stability condition, it is useful to consider the eigenvector basis. We therefore consider the linear conduction case and go back to equation [4.23], which is re-written as:

$$\dot{\beta}_j(t) + \alpha_j \beta_j(t) = \frac{1}{c_j} f_j(t) \quad [4.31]$$

Applying the generalized central difference algorithm to the equation above over the interval $[p\Delta t, (p+1)\Delta t]$; we obtain:

$$\left(\frac{1}{\Delta t} + \nu \alpha_j \right) (\beta_j)_{p+1} = \left(\frac{1}{\Delta t} - (1 - \nu) \alpha_j \right) (\beta_j)_p + \frac{1}{c_j} (f_j)_{p+\nu}$$

which, once rearranged, becomes:

$$(\beta_j)_{p+1} = \frac{1 - (1 - \nu) \alpha_j \Delta t}{1 + \nu \alpha_j \Delta t} (\beta_j)_p + \frac{\Delta t}{1 + \nu \alpha_j \Delta t} \frac{1}{c_j} (f_j)_{p+\nu} \quad [4.32]$$

By recurrence this relationship yields $(\beta_j)_{p+1}$ as a function of $(\beta_j)_0$. The algorithm will be stable if all the $(\beta_j)_{p+1}$ remain bounded, regardless of j and p . The algorithm stability condition is therefore written as:

$$\forall j, \quad \left| \frac{1 - (1 - \nu) \alpha_j \Delta t}{1 + \nu \alpha_j \Delta t} \right| \leq 1, \quad \text{i.e.} \quad \forall j, \quad (1 - 2\nu) \alpha_j \Delta t \leq 2$$

From this relationship it can be deduced that if $1/2 \leq \nu \leq 1$, then the scheme used is unconditionally stable. On the contrary, if $0 \leq \nu < 1/2$, then the scheme is conditionally stable, and the time step used, Δt , must be less than or equal to the critical stability step, Δt_{sta} :

$$\Delta t \leq \Delta t_{\text{sta}} \quad \text{with} \quad \Delta t_{\text{sta}} = \frac{2}{(1 - 2\nu)\alpha_{\text{max}}} \quad [4.33]$$

where α_{max} represents the greatest eigenvalue.

When going back to relationship [4.32], it is clear that the solution oscillates when $\frac{1 - (1 - \nu)\alpha_j \Delta t}{1 + \nu\alpha_j \Delta t} < 0$. A non-oscillation condition of the solution over time can be derived from this. This condition is written as:

$$\begin{aligned} \forall j, \quad \frac{1 - (1 - \nu)\alpha_j \Delta t}{1 + \nu\alpha_j \Delta t} \geq 0, \quad \text{i.e.} \quad \Delta t \leq \Delta t_{\text{osc}} \\ \text{with} \quad \Delta t_{\text{osc}} = \frac{1}{(1 - \nu)\alpha_{\text{max}}} \end{aligned} \quad [4.34]$$

We can see that relationship [4.34] is always satisfied for the Euler implicit algorithm. Therefore, this algorithm never produces an oscillating solution, whatever the time step value.

4.2.3.3. Simplified analysis of the stability condition

Relationship [4.33] gives the stability condition in a stringent manner. However, to be applied practically, this relationship requires a previous calculation of the problem eigenvalues and modes, which increases the global cost of the analysis. Thus, it is useful to have a stability condition generated by a simplified analysis.

To do so, we consider a 1D case with no source term and a regular mesh composed of n 2-noded elements of length e . The complete mesh thus consists of $n + 1$ nodes. Besides, we consider a capacitance thermal matrix lumped by means of the technique described in section 4.2.3.2. It is then easy to obtain:

$$[K^e] = \frac{\lambda}{e} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \Rightarrow [K] = \frac{\lambda}{e} \begin{bmatrix} 1 & -1 & 0 & \dots & \dots & \dots \\ -1 & 2 & -1 & 0 & \dots & \dots \\ 0 & -1 & 2 & -1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & 0 & -1 & 2 & -1 \\ \dots & \dots & \dots & 0 & -1 & 1 \end{bmatrix}$$

$$[C^e] = \rho C e \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} \Rightarrow [C] = \rho C e \begin{bmatrix} 1/2 & 0 & 0 & \cdots & \cdots & \cdots \\ 0 & 1 & 0 & 0 & \cdots & \cdots \\ 0 & 0 & 1 & 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & 0 & 0 & 1 & 0 \\ \cdots & \cdots & \cdots & \cdots & 0 & 1/2 \end{bmatrix}$$

The generic equation corresponding to node $1 < j < n + 1$ is written as:

$$\rho C e \dot{T}_j + \frac{\lambda}{e} (-T_{j-1} + 2T_j - T_{j+1}) = 0$$

Applying the generalized central difference algorithm then yields, after a few mathematical manipulations:

$$(T_j)_{p+1} - (T_j)_p + Fo \left(\begin{aligned} &(1 - \nu) \left(- (T_{j-1})_p + 2(T_j)_p - (T_{j+1})_p \right) \\ &+ \nu \left(- (T_{j-1})_{p+1} + 2(T_j)_{p+1} - (T_{j+1})_{p+1} \right) \end{aligned} \right) = 0 \quad [4.35]$$

where Fo is the Fourier number:

$$Fo = \frac{\lambda \Delta t}{\rho C e^2} \quad [4.36]$$

To study the stability, the Fourier method is used here. This method analyzes the amplitude variation of each term of the Fourier series development of a disturbance of the initial solution. We therefore consider a disturbance in the form $\sum_l V^l(t) e^{ilx}$, i being the complex number such that $i^2 = -1$, and we study the variation of the l^{th} harmonic V^l when equation [4.35] is applied to it.

If we term x_j the x -coordinate of node j and given that $x_{j-1} = x_j - e$ and $x_{j+1} = x_j + e$, we obtain:

$$\left((V^l)_{p+1} - (V^l)_p \right) e^{ilx_j} + Fo \left(\begin{aligned} &(1 - \nu) (V^l)_p (-e^{-ile} + 2 - e^{ile}) \\ &+ \nu (V^l)_{p+1} (-e^{-ile} + 2 - e^{ile}) \end{aligned} \right) e^{ilx_j} = 0$$

from which we derive:

$$\frac{(V^l)_{p+1}}{(V^l)_p} = \frac{1 + 4Fo(\nu - 1) \sin^2(le/2)}{1 + 4Fo\nu \sin^2(le/2)}$$

It is obvious that the stability condition, also called the von Neumann condition, is written as:

$$\left| \frac{1 + 4Fo(\nu - 1) \sin^2(le/2)}{1 + 4Fo\nu \sin^2(le/2)} \right| \leq 1$$

For this condition to be satisfied whatever the harmonic l considered, it is sufficient that:

$$(1 - 2\nu)Fo \leq \frac{1}{2} \quad [4.37]$$

Of course, for $1/2 \leq \nu \leq 1$, condition [4.37] is always satisfied. In the other cases, the stability time step derived from [4.37] is given by:

$$\Delta t_{\text{sta}} = \frac{\rho C e^2}{2\lambda(1 - 2\nu)} \quad [4.38]$$

This stability condition is only approximated. It is valid only in the conditions considered to establish it. In 2D and 3D cases, and for elements of the same size in the 2 or 3 directions, it is possible to demonstrate that the stability condition related to the Euler explicit scheme is written respectively as $Fo \leq \frac{1}{4}$ and $Fo \leq \frac{1}{6}$. When all the elements are not the same size (which is the most common case), the stability criterion must be calculated with the smallest element.

4.2.4. Practical complementary rules

We have just seen that, for reasons of stability or non-oscillation of the solution over time, the time step Δt used when simulating heat exchanges with finite elements, was always to remain less than two limit values (Δt_{sta} defined by equation [4.33] and Δt_{osc} defined by equation [4.34]). In this section, a few complementary rules to choose this time step Δt are given. These complementary rules illustrate the fact that the choice of the time step depends not only on the time integration algorithm selected (value of ν in particular) and the material properties (eigenvalues of the matrix $[C]^{-1} [K]$) but also on the mesh chosen to represent the structure on which the simulation is performed.

4.2.4.1. Space oscillations during thermal shock simulation

In the case of a thermal shock, if the terms of the capacitance matrix $[C]$ are not lumped on the diagonal, the simulation time step must remain greater than a value Δt_s , so as to avoid the appearance of space oscillations of the solution during the initial moments of the calculation process. This phenomenon will be illustrated on a 1D problem.

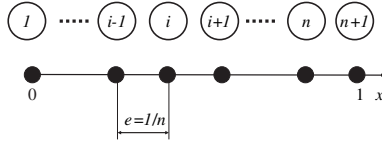


Figure 4.2. Mesh of the segment $[0, 1]$ with n elements

Consider the case of a body initially at homogenous temperature, put in sudden contact with another body at a different temperature (thermal shock problem). To represent this thermal shock without losing the general characteristic of the results obtained, the following 1D problem is used:

Find $T(x, t)$ satisfying at each instant t :

$$\begin{aligned} \forall x \in]0, 1[, \quad \rho C \dot{T} &= \lambda \frac{\partial^2 T}{\partial x^2} \\ \frac{\partial T}{\partial x}(0, t) &= 0 \\ T(1, t) &= 1 \end{aligned}$$

with the initial condition:

$$\forall x \in [0, 1[, \quad T(x, 0) = 0$$

To solve this finite element problem, the segment $[0, 1]$ is divided into n first order elements (with linear interpolation) of length $e = 1/n$ (Figure 4.2). The mesh obtained includes $n + 1$ nodes. The resolution of the problem leads us to seek the vector $\{T\}$ including the nodal values of the temperature field at each instant (T_i is the temperature value at node i at that instant), and satisfying:

$$[C] \cdot \{\dot{T}\} + [K] \cdot \{T\} = \{0\} \quad \text{with} \quad \begin{cases} T_{n+1} = 1 & \text{for any } t > 0 \\ T_i = 0 \text{ at } t = 0 & \text{for any } 1 \leq i \leq n \end{cases}$$

Conductance matrix $[K]$ and capacitance matrix $[C]$ can easily be calculated from the element matrices:

$$[K^e] = \frac{\lambda}{e} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \Rightarrow [K] = \frac{\lambda}{e} \begin{bmatrix} 1 & -1 & 0 & \dots & \dots & \dots \\ -1 & 2 & -1 & 0 & \dots & \dots \\ 0 & -1 & 2 & -1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & 0 & -1 & 2 & -1 \\ \dots & \dots & \dots & 0 & -1 & 1 \end{bmatrix}$$

$$[C^e] = \frac{\rho C e}{6} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \Rightarrow [C] = \frac{\rho C e}{6} \begin{bmatrix} 2 & 1 & 0 & \dots & \dots & \dots \\ 1 & 4 & 1 & 0 & \dots & \dots \\ 0 & 1 & 4 & 1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & 0 & 1 & 4 & 1 \\ \dots & \dots & \dots & \dots & 1 & 2 \end{bmatrix}$$

The generic equation at one node $2 \leq i \leq n$ is then written as:

$$\dot{T}_{i-1} + 4\dot{T}_i + \dot{T}_{i+1} + \frac{6Fo}{\Delta t} (-T_{i-1} + 2T_i - T_{i+1}) = 0$$

Now we will apply the generalized trapezoidal algorithm to estimate the temperatures T_i at the beginning of the thermal shock at $t = \Delta t$. By writing the previous equation at time Δt , and considering that the terms T_i and \dot{T}_i are equal to zero at $t = 0$, we get $\dot{T}_i = T_i/\nu\Delta t$ for each index i and obtain:

$$T_i = A(T_{i-1} + T_{i+1}) \quad \text{with} \quad A = \frac{6Fo\nu - 1}{4 + 12Fo\nu}$$

If we write $D_i = T_i - T_{i-1}$, we see that the above equation is satisfied in the same manner by these quantities, i.e.:

$$D_i = A(D_{i-1} + D_{i+1})$$

In order to avoid space oscillations of the solution, Figure 4.3 suggests that all the D_i must have the same sign. We derive from this a non-oscillation condition of the solution which is written $A \geq 0$, i.e.:

$$6\nu Fo \geq 1 \Rightarrow \Delta t \geq \Delta t_s \quad \text{with} \quad \Delta t_s = \frac{\rho C e^2}{6\lambda\nu} \quad [4.39]$$

If condition [4.39] is not satisfied, then the solution obtained will produce space oscillations contrary to physical laws. In fact, this condition must be generally considered as a constraint on the size e of the first element affected by the thermal shock. This constraint gives the element maximum size according to the time step used:

$$e \leq e_{\max} \quad \text{with} \quad e_{\max} = \sqrt{\frac{6\lambda\nu\Delta t}{\rho C}} \quad [4.40]$$

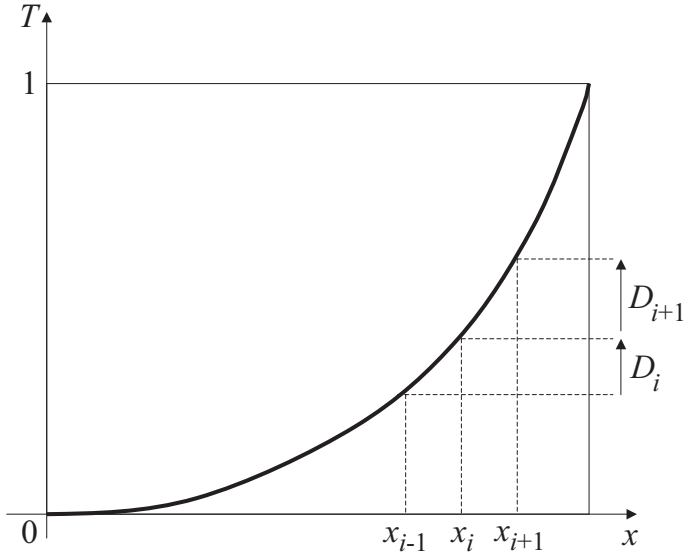


Figure 4.3. General aspect of the problem solution

Consider a capacitance matrix obtained by lumping the thermal capacitance on the nodes (see section above). In our simple example, this comes down to using the following matrix $[C]$:

$$[C] \approx \frac{\rho C e}{6} \begin{bmatrix} 3 & 0 & 0 & \cdots & \cdots & \cdots \\ 0 & 6 & 0 & 0 & \cdots & \cdots \\ 0 & 0 & 6 & 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & 0 & 0 & 6 & 0 \\ \cdots & \cdots & \cdots & \cdots & 0 & 3 \end{bmatrix}$$

If the same method as before is used, the following relationship is then obtained:

$$T_i = A'(T_{i-1} + T_{i+1}) \quad \text{with} \quad A' = \frac{6F_{ov}}{6 + 12F_{ov}}$$

The constant A' thus defined remains always positive or equal to zero, so that the solution obtained does not oscillate, whatever the time step Δt used. In fact, this lumped approximation of the capacitance matrix consists of introducing an additional

conduction term. Indeed, the constant A' can easily be expressed in the same form as the constant A previously defined, but by means of a modified thermal conductivity:

$$A' = \frac{6Fo'\nu - 1}{4 + 12Fo'\nu} \quad \text{with} \quad Fo' = \frac{\lambda'\Delta t}{\rho C e^2} \quad \text{and} \quad \lambda' = \lambda + \frac{\rho C e^2}{6\nu\Delta t}$$

4.2.4.2. Discrete maximum principle

We saw that in thermal shock problems, space oscillations of the solution could appear if the time step chosen was less than a value Δt_s . In fact for time steps $\Delta t < \Delta t_s$, the discrete maximum principle is no longer satisfied.

Generally speaking, it is reported that a function defined in a domain Ω , as the solution of a boundary value problem, satisfies the maximum principle if it reaches its maximum on the boundary $\partial\Omega$ of Ω .

In practice, the temperature satisfies the maximum principle in many thermal problems. For instance, we can list the three following situations:

- in steady state conduction problems, if temperature boundary conditions alone are applied, the temperature is maximum on Ω boundary;
- in the case of transient heat conduction, with no internal heat source, the maximum temperature is reached at the initial instant or on the boundary;
- for a solid heated in a bath by means of an exchange coefficient, the temperature of the solid is bounded by that of the bath.

The first two examples are still valid for diffusion-convection problems.

When the continuous problem is discretized it is termed the discrete maximum principle. The satisfaction of this principle ensures that the discretized solution has the same bounds as the exact solution of the continuous problem. Indeed, the discrete maximum principle is a simple principle which any discrete system must satisfy for its solution to make sense from a physical point of view.

The finite element discretization of the continuous problem does not necessarily produce systems satisfying the discrete maximum principle. We saw an example earlier with the thermal shock problem and we will see another with the diffusion-convection problem.

As was seen in the previous chapters, after the possible application of a time integration algorithm in transient problems, the finite element method leads to the solution, at each time step, of a linear system of the type:

$$[A] \cdot \{T\} = \{Q\}$$

where vector $\{T\}$ includes the unknown temperatures and $[A]$ and $\{Q\}$ depend on the problem being dealt with.

It must be emphasized that the above linear system includes the boundary conditions with prescribed temperature.

Sufficient conditions for matrix $[A]$ to satisfy the discrete maximum principle were proposed by Ciarlet [CIA 70]. They are the following:

- C1: the matrix $[A]$ is invertible and has a dominant diagonal;
- C2: all the diagonal terms of $[A]$ are positive and the non-diagonal terms are negative or equal to zero.

Consider once again the thermal shock problem dealt with in the previous section. Matrix $[A]$ in this case is the following:

$$[A] = [K] + \frac{1}{\nu \Delta t} [C]$$

$$[A] = \frac{\rho C e}{6 \nu \Delta t} \begin{bmatrix} a & b & 0 & \cdots & \cdots & 0 \\ b & 2a & b & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & b & 2a & b \\ 0 & \cdots & \cdots & 0 & b & a \end{bmatrix}$$

with

$$a = 6\nu Fo + 2 \quad b = -6\nu Fo + 1$$

Matrix $[A]$ above is invertible and has a dominant diagonal. All the diagonal terms are positive, but the non-diagonal terms are not necessarily negative or equal to zero.

For condition C2 to be satisfied, we must have:

$$b \leq 0 \iff \Delta t \geq \Delta t_s \quad \text{with} \quad \Delta t_s = \frac{\rho C e^2}{6\lambda\nu}$$

With such a condition, the discrete maximum principle is satisfied.

We can also check that conditions C1 and C2 are always satisfied when a diagonal specific heat matrix is used.

Here the analysis was carried out in a 1D theoretical case by means of first order finite elements. In 2D and 3D cases, and for isoparametric finite elements of a higher order, it is necessary to study matrix $[A]$. This is all the more difficult as the Fourier number changes from one element to the other.

The discrete maximum principle combined with the finite element method has been extensively studied, especially by Lobo and Emery [LOB 93, LOB 95], for thermal shock problems with convective or radiative boundary conditions. The authors suggest the use of certain types of preferable elements.

The main elements they recommend are the following:

- first order elements with a lumped capacitance matrix. In a transient problem, with first order elements, the capacitance matrix brings a positive contribution to the non-diagonal terms of the matrix $[A] = [K] + \frac{1}{\nu\Delta t} [C]$. Using a lumped capacitance matrix eliminates this problem and gives great efficiency to this type of element. Introducing thermal exchange conditions by means of an exchange coefficient produces the same type of problem;

- high order (quadratic or cubic) isoparametric elements make it possible to satisfy the discrete maximum principle over a wider range of the Fourier number. This is due to the fact that the positive contribution of the mass matrix to the non-diagonal terms of $[A]$ is distributed over a larger number of terms.

The authors also recommend first order elements modified by the addition of second order terms to satisfy the discrete maximum principle and of C^1 elements as well. However, the scope of application and efficiency of these elements are rather limited. Moreover, these elements are generally not present in major computer codes.

4.2.4.3. *Initial temperatures during thermal contact simulation*

The thermal contact problem occurs when casting processes are simulated. To illustrate this, a simple 1D example will be used as in the previous section.

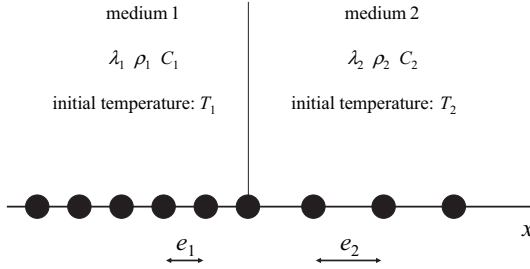


Figure 4.4. Mesh of the media for the thermal contact study

Consider two media initially at homogenous temperatures T_1 and T_2 , being put in sudden contact at instant $t = 0$. In order to calculate the transient thermal states with the finite element method, a mesh consisting of first order elements and uniform in each of the two media (Figure 4.4) is constructed. e_1 and e_2 will denote the sizes of the elements in media 1 and 2.

The contact is assumed to be perfect, so that the mesh is continuous when passing between the two media 1 and 2. The main problem in this type of simulation is to choose the initial temperature at the interface node between the two media.

If the media are large enough to behave as if they were infinite, then the problem has an analytical solution [CAR 59]. In this case, the temperature at the contact node is known to remain constant and takes the following value:

$$T_c = \frac{b_1 T_1 + b_2 T_2}{b_1 + b_2} \quad \text{with} \quad b_1 = \sqrt{\lambda_1 \rho_1 C_1} \quad \text{and} \quad b_2 = \sqrt{\lambda_2 \rho_2 C_2}$$

In this equation, b_1 and b_2 represent the thermal effusivity of media 1 and 2 respectively.

Choosing T_c as the initial value of the contact temperature is not always the best choice. In order to calculate the temperature to use, which will be written T'_c , we start with the initial distribution (at $t = 0$) of the temperatures in both media, resulting from the finite element discretization. This distribution is illustrated in Figure 4.5.

The only exchanges possible, when both media are in contact, are related to the two media only. If the temperature T'_c at the contact node is to be physically acceptable, it must reproduce the fact that the quantity of heat gained by one medium is equal to that lost by the other medium. By writing Q_i the quantity of heat gained by medium i

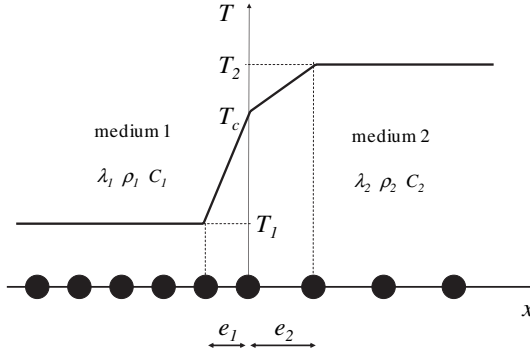


Figure 4.5. Initial temperature distribution

between instants $-\epsilon$ and 0 ($\epsilon > 0$, so that instant $-\epsilon$ represents the instant immediately before contact), we obtain:

$$Q_i = \int_{-\epsilon}^0 \int_{\Omega_i} \rho_i C_i \dot{T} dx dt = \int_{\Omega_i} \rho_i C_i (T(x, 0) - T(x, -\epsilon)) dx$$

Given that the temperatures at instant $-\epsilon$ are T_1 in medium 1 and T_2 in medium 2, that the function $T(x, 0)$ is linear in x due to the elements used, and assuming that the characteristics of the media are homogenous and are not temperature-dependent, the quantity of heat Q_i can be written as follows:

$$Q_i = \frac{1}{2} \rho_i C_i e_i (T'_c - T_i)$$

The thermal balance between the two media prescribes $Q_1 + Q_2 = 0$, which produces the following expression for the contact temperature:

$$T'_c = \frac{\rho_1 C_1 e_1 T_1 + \rho_2 C_2 e_2 T_2}{\rho_1 C_1 e_1 + \rho_2 C_2 e_2}$$

This physically coherent value of the contact node temperature is significantly different from the theoretical value T_c obtained for infinite media. In fact, it is possible to choose the sizes of the elements close to the contact, e_1 and e_2 , such that T_c and T'_c coincide. To do so, it is sufficient that the ratios of the coefficients b_i (producing T_c) to $\rho_i C_i e_i$ (producing T'_c) of each medium are equal. The following mesh criterion is obtained:

$$\frac{e_2}{e_1} = \sqrt{\frac{\frac{\lambda_2}{\rho_2 C_2}}{\frac{\lambda_1}{\rho_1 C_1}}}$$

This criterion is also written $F_{O_1} = F_{O_2}$, where F_{O_i} is a Fourier number, defined by equation [4.36,] corresponding to the medium i .

According to this criterion, the square ratio of the sizes of the first elements in media 1 and 2 must be equal to the ratio of the thermal diffusivity of these media. This criterion can also be obtained directly, and in a more general framework, by writing $\dot{T} = 0$ at $t = 0$ at the interface level, from the discretized equations. By choosing T_c as the temperature at the contact node, and satisfying that criterion when meshing these two media in contact, a physically coherent value is used, because it satisfies the thermal balance between both media.

In the case of the solidification of a cast iron piece in a sand mold, hence a sand-cast iron interface (a traditional casting problem) the criterion established produces very different element sizes on each side of the interface [BER 86]. Indeed, consider the following mean characteristics of the two media:

– $\lambda_1 = 0.4 \text{ W/m/K}$, $\rho_1 = 1,500 \text{ kg/m}^3$ and $C_1 = 1,150 \text{ J/kg/K}$ for sand (medium 1);

– $\lambda_2 = 40 \text{ W/m/K}$, $\rho_2 = 7,500 \text{ kg/m}^3$ and $C_2 = 820 \text{ J/kg/K}$ for cast iron (medium 2).

Applying the criterion produces $e_2/e_1 = 5.3$. If it is not satisfied, we can show that:

– if $e_2/e_1 < 5.3$, then the metal cooling process slows down, and the contact temperature becomes greater than the analytical value T_c ;

– if $e_2/e_1 > 5.3$, then the cast iron cooling process accelerates, and the contact temperature becomes less than the analytical value T_c .

We have established this mesh criterion in a simple 1D case. It is easily extended to 2D and 3D cases. On each side of the interface, the mesh must necessarily be in the form of regular layers whose thicknesses are in a ratio satisfying the criterion. For instance, in the case of the cast iron-sand interface, the first layer regularity on each side of the interface is a strict obligation, otherwise it will significantly affect the positions of the hot points. Indeed, the local cooling kinetics will depend on the satisfaction or the non-satisfaction of this criterion.

Some computer codes include meshing procedures, enabling them to make successive layers of elements with controlled thicknesses on each side of the interface. It is then possible to satisfy the criterion previously established. Figure 4.6 illustrates a 2D mesh example performed from the layer mesher of the software *SYSTUSTM*.

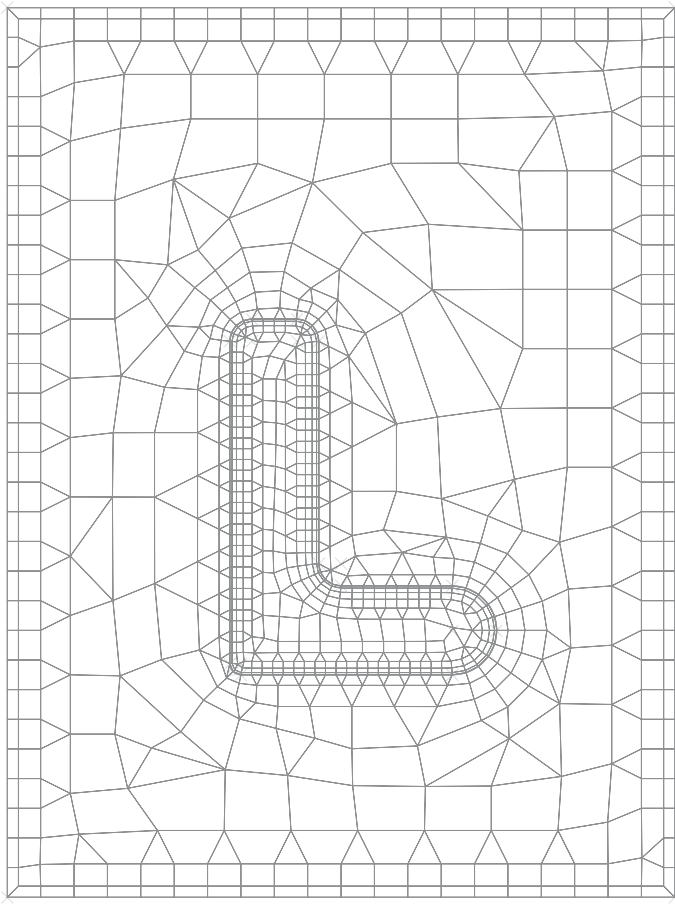


Figure 4.6. *Mesh of a casting and its mold*

One way of obtaining, with certainty, and independently of the mesh used, physically coherent initial conditions at the interface of a perfect thermal contact consists of:

- assuming first of all that both media are at the same initial temperature, that of medium T_1 , for instance;
- calculating the energy required to bring medium 2 to temperature T_2 ;
- prescribing that energy as an internal heat source distributed in the volume of the elements representing medium 2 for a very short time.

Very short time means the time allowing no significant gradient to settle down beyond the first layer of elements on each side of the interface. The interface temperature then adjusts itself to the local value T'_c which will coincide with the value T_c if the element thicknesses satisfy the criterion.

Another way of dealing with thermal interfaces is to perform two separate meshes, and introduce on the interface a thermal contact resistance. The interface heat flux will then be given locally according to the temperatures of each medium obtained by nodal approximation. If T_1 is the local temperature of medium 1, and T_2 that of medium 2, then the heat flux Q_1 gained by medium 1 is equal to $Q_1 = (T_2 - T_1)/R_c$, whereas that gained by medium 2 is equal to $Q_2 = -Q_1 = (T_1 - T_2)/R_c$. In these expressions, R_c is a characteristic introduced for the interface, often called thermal contact resistance. If a very low value is used for R_c , the situation is close to that of a perfect contact.

4.3. Working example

The aim of this section is to illustrate with a simple example the notions introduced in this chapter to simulate transient heat conduction in a solid. To do so, the induction heating example used to illustrate the formulation of the steady state conduction problem is considered again. Firstly, the physical modeling process used in this example and its finite element approximation are studied in depth, then a few numerical applications are performed to illustrate the notions seen in this chapter.

4.3.1. Physical modeling and approximation

The induction heating of a plate is illustrated in Figure 4.7. In Part 1 of this book, this example was dealt with using a heating volumetric term Q and an exchange coefficient h independent of time and the steady state corresponding to those conditions was sought. Here, a thermal capacitance of the material is introduced and

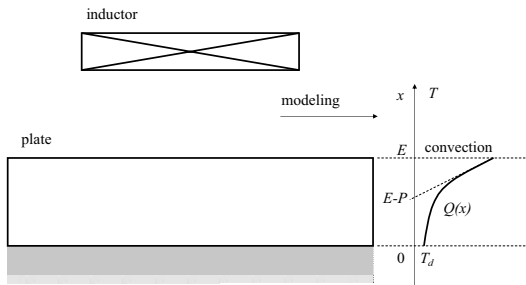


Figure 4.7. Working example: plate induction quenching

the study of transient states allows us to use the heating term Q and cooling term h as functions of time. The volumetric heat source is expressed as $Q(x, t) = Q_0(t)e^{\frac{x-E}{P}}$, where only the term $Q_0(t)$ is a function of time. The exchange coefficient h will simply be written as $h(t)$ for the moment.

In this example, for convenience, we choose physical properties λ , ρ and C independent of temperature. The solution of the thermal conduction problem in the plate consists, in the transient state, of seeking a temperature field $T(x, t)$ which verifies at any point $x \in [0, E]$ and at any instant t :

$$\begin{cases} \lambda \frac{\partial^2 T}{\partial x^2} + Q(x, t) = \rho C \dot{T}(x, t) \\ \lambda \frac{\partial T}{\partial x} = h(t)(T_d - T(x, t)) & \text{in } x = E \\ T(x, t) = T_d & \text{at } x = 0 \\ T(x, t) = T_0 & \text{at } t = 0 \end{cases}$$

We added to this equation the term T_0 which represents the initial plate temperature (at $t = 0$). This temperature is assumed to be constant over time at point $x = 0$. On the contrary, the external temperature at point $x = E$ was set at T_d .

To solve the problem, the finite element method is used. To do so, we discretize axis $0x$ in Figure 4.7 between 0 and E by means of $n+1$ nodes at uniform distances of $e = E/n$. The node coordinates are $x_i = ie$, with $i = 0, 1, \dots, n$. It is now possible to choose the type of element. We select n two-noded isoparametric elements with linear interpolation. We term $\Omega^i = [x_{i-1}, x_i]$, with $i = 1, 2, \dots, n$, the domain occupied by element i . With this discretization, the following nodal approximation is used for the temperatures:

$$\forall x \in \Omega^i, \forall t \geq 0, \quad T(x, t) = N_1^i(x)T_1^i(t) + N_2^i(x)T_2^i(t) \quad \text{with} \quad \begin{cases} N_1^i(x) = \frac{x_i - x}{e} \\ N_2^i(x) = \frac{x - x_{i-1}}{e} \end{cases}$$

In that case, the assembly of thermal conductance and capacitance matrices, and that of the element load vector, produces the following expressions for matrices $[K]$ and $[C]$ and for vector $\{F\}$:

$$[K] = \frac{\lambda}{e} \begin{bmatrix} 1 & -1 & 0 & \dots & \dots & \dots & 0 \\ -1 & 2 & -1 & 0 & \dots & \dots & \dots \\ 0 & -1 & 2 & -1 & 0 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & 0 & -1 & 2 & -1 & 0 \\ \dots & \dots & \dots & 0 & -1 & 2 & -1 \\ 0 & \dots & \dots & \dots & 0 & -1 & 1 \end{bmatrix} + h(t) \begin{bmatrix} 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \dots & \dots & 0 & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix}$$

$$[C] = \frac{\rho C e}{6} \begin{bmatrix} 2 & 1 & 0 & \dots & \dots & \dots & 0 \\ 1 & 4 & 1 & 0 & \dots & \dots & \dots \\ 0 & 1 & 4 & 1 & 0 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & 0 & 1 & 4 & 1 & 0 \\ \dots & \dots & \dots & 0 & 1 & 4 & 1 \\ 0 & \dots & \dots & \dots & 0 & 1 & 2 \end{bmatrix}$$

$$\{F\} = Q_0(t) \left\{ \begin{array}{c} -\frac{P^2}{e} \left(e^{\frac{x_0-E}{P}} - e^{\frac{x_1-E}{P}} \right) + P e^{\frac{x_1-E}{P}} \\ -\frac{P^2}{e} \left(2e^{\frac{x_1-E}{P}} - e^{\frac{x_0-E}{P}} - e^{\frac{x_2-E}{P}} \right) \\ -\frac{P^2}{e} \left(2e^{\frac{x_2-E}{P}} - e^{\frac{x_1-E}{P}} - e^{\frac{x_3-E}{P}} \right) \\ \dots \\ -\frac{P^2}{e} \left(2e^{\frac{x_{n-1}-E}{P}} - e^{\frac{x_{n-2}-E}{P}} - e^{\frac{x_n-E}{P}} \right) \\ -\frac{P^2}{e} \left(2e^{\frac{x_n-E}{P}} - e^{\frac{x_{n-1}-E}{P}} - e^{\frac{x_{n+1}-E}{P}} \right) \\ -\frac{P^2}{e} \left(e^{\frac{x_{n+1}-E}{P}} - e^{\frac{x_n-E}{P}} \right) + P e^{\frac{x_{n+1}-E}{P}} \end{array} \right\} + h(t) \left\{ \begin{array}{c} 0 \\ 0 \\ 0 \\ \dots \\ 0 \\ 0 \\ T_d \end{array} \right\}$$

The integrals involved here were calculated analytically. We have deliberately emphasized in these expressions the presence of the two time functions $Q_0(t)$ and $h(t)$, which are used to model induction heating. Only the capacitance matrix is not time-dependent in this example.

Now it is necessary to seek the temperature change at the n nodes over time. Indeed the temperature of node 0 is known: it is constant and equal to T_d . The problem to solve is written:

Find at each instant t the nodal temperature vector $\{T(t)\}$ such that:

$$\begin{cases} [K(t)] \cdot \{T(t)\} + [C] \cdot \{\dot{T}(t)\} = \{F(t)\} \\ \{T(0)\} = \{T_0\} \end{cases}$$

To obtain the variation of the temperature vector at the nodes $\{T(t)\}$ it is now necessary to choose a time integration scheme. The generalized trapezoidal rule will be used. Thus, over time, if we know the temperature vector at time t , $\{T(t)\}$, $\{\dot{T}(\tau)\}$ at any instant $\tau \in [t, t + \Delta t]$ is expressed as (see equation [4.27]):

$$\begin{aligned} \forall \tau \in [t, t + \Delta t], \quad \{\dot{T}(\tau)\} &= \frac{\{T(t + \Delta t)\} - \{T(t)\}}{\Delta t} \\ &= (1 - \nu) \{\dot{T}(t)\} + \nu \{\dot{T}(t + \Delta t)\} \text{ with } 0 \leq \nu \leq 1 \end{aligned}$$

In practice, the previous expression is used to express the quantity $\{\dot{T}(t + \Delta t)\}$ as follows:

$$\{\dot{T}(t + \Delta t)\} = -\frac{1 - \nu}{\nu} \{\dot{T}(t)\} + \frac{1}{\nu \Delta t} (\{T(t + \Delta t)\} - \{T(t)\})$$

The problem to solve at instant $t + \Delta t$ is therefore written:

Knowing $\{T(t)\}$ and $\{\dot{T}(t)\}$, find the vector $\{T(t + \Delta t)\}$ such that:

$$\begin{cases} ([C] + \nu \Delta t [K(t + \Delta t)]) \cdot \{T(t + \Delta t)\} = [C] \cdot \{T(t)\} + \nu \Delta t \{F(t + \Delta t)\} \\ \quad + (1 - \nu) \Delta t [C] \cdot \{\dot{T}(t)\} \\ T_0(t + \Delta t) = T_d \end{cases}$$

Note here that the matrix involved in this linear system can be written with the addition of Δt_s defined by equation [4.39]:

$$[C] + \nu \Delta t [K] = \frac{\rho C e}{6} \begin{bmatrix} 2 + \frac{\Delta t}{\Delta t_s} & 1 - \frac{\Delta t}{\Delta t_s} & 0 & \cdots & 0 \\ 1 - \frac{\Delta t}{\Delta t_s} & 4 + 2 \frac{\Delta t}{\Delta t_s} & 1 - \frac{\Delta t}{\Delta t_s} & 0 & \cdots \\ 0 & \cdots & \cdots & \cdots & 0 \\ \cdots & 0 & 1 - \frac{\Delta t}{\Delta t_s} & 4 + 2 \frac{\Delta t}{\Delta t_s} & 1 - \frac{\Delta t}{\Delta t_s} \\ 0 & \cdots & 0 & 1 - \frac{\Delta t}{\Delta t_s} & 2 + \frac{\Delta t}{\Delta t_s} \end{bmatrix}$$

The non-diagonal terms of this matrix will change signs according to the ratio of the time step used Δt to the minimum time step to use Δt_s . In practice, the capacitance matrix $[C]$ is approximated by a lumped matrix, obtained by transferring on the diagonal the sum of the terms on each line. This produces the following expression for the matrix involved in the linear system to solve:

$$[C] + \nu \Delta t [K] \approx \frac{\rho C e}{6} \begin{bmatrix} 3 + \frac{\Delta t}{\Delta t_s} & -\frac{\Delta t}{\Delta t_s} & 0 & \dots & 0 \\ -\frac{\Delta t}{\Delta t_s} & 6 + 2\frac{\Delta t}{\Delta t_s} & -\frac{\Delta t}{\Delta t_s} & 0 & \dots \\ 0 & \dots & \dots & \dots & 0 \\ \dots & 0 & -\frac{\Delta t}{\Delta t_s} & 6 + 2\frac{\Delta t}{\Delta t_s} & -\frac{\Delta t}{\Delta t_s} \\ 0 & \dots & 0 & -\frac{\Delta t}{\Delta t_s} & 3 + \frac{\Delta t}{\Delta t_s} \end{bmatrix}$$

In that expression, whatever the time step Δt chosen, no term of the matrix changes signs. It follows that there will be no significant change in the result (no numerical instability) when the time step Δt is modified.

4.3.2. Numerical applications

The numerical applications shown in this book are used to illustrate the notions introduced in this chapter. To do so, we have chosen some fixed values, and others which will change. The fixed values are the following:

- plate thickness: $E = 0.1$ m;
- number of finite elements: $n = 10$;
- material thermal conductivity: $\lambda = 30$ W/m/K;
- external medium temperature: $T_d = 20^\circ\text{C}$;
- induction heating depth: $P = 0.02$ m;
- volumetric mass: $\rho = 7,800$ Kg/m³;
- specific heat: $C = 500$ J/Kg/K.

The aim of the first simulation is to show that the simple introduction of the capacitance term without modifying the other data with respect to the steady state calculation, only delays the implementation of the steady state. Indeed, Figure 4.8 shows a calculation performed with:

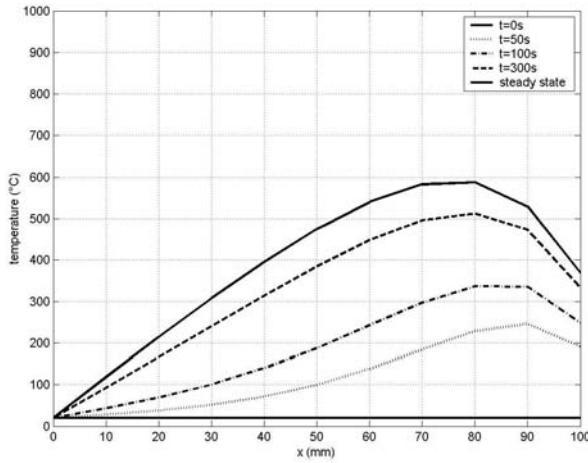


Figure 4.8. *Variation of the temperature profile over time*

- a constant exchange coefficient: $H = 2,000 \text{ W/m}^2/\text{K}$;
- a constant heating power: $Q_0 = 5.10^7 \text{ W/m}^3$;
- a plate initial temperature: $T_0 = 20^\circ\text{C}$;
- a time integration scheme (trapezoidal rule): $\nu = 0.5$;
- a constant time step: $\Delta t = 1 \text{ s}$.

In this figure, we note that the temperature profile changes gradually from a constant value T_0 at $t = 0$ toward the steady state profile (calculated in the first part of this book) for an “infinite” time. Indeed, after a fairly long time, the temperature does not change over time, so that $\{\dot{T}\} = \{0\}$ in the differential equation to solve. The solution is then obtained in the steady state.

The second simulation illustrates the limit time step Δt_s highlighted in the previous equations. To do so, we have restricted our study to the cooling phase, starting with a high initial temperature:

- a constant exchange coefficient: $H = 10,000 \text{ W/m}^2/\text{K}$;
- a heating power equal to zero: $Q_0 = 0$;
- a plate initial temperature: $T_0 = 700^\circ\text{C}$;
- a time integration scheme: $\nu = 1$.

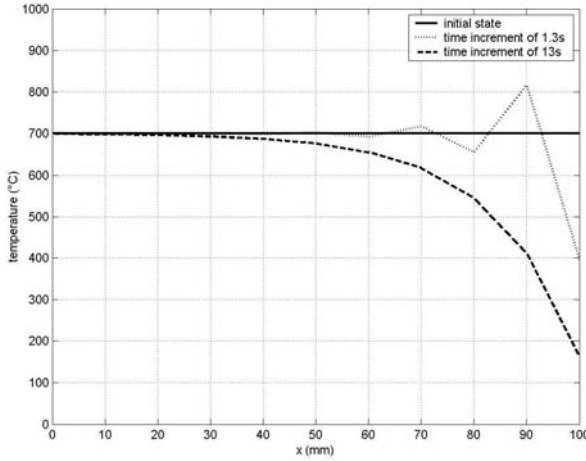


Figure 4.9. Plate quenching – temperature profile at $t = \Delta t$ for $\Delta t = 1.3 \text{ s}$ and $\Delta t = 13 \text{ s}$

We can see in Figure 4.9 that the solution obtained after the first time increment depends on the time step chosen. With $\Delta t = 13 \text{ s}$, the solution is relatively conform to physics, whereas with $\Delta t = 1.3 \text{ s}$ there are oscillations contrary to physics. In fact the threshold time step in this example being $\Delta t_s = \frac{\rho C e^2}{6\lambda} \approx 2.166 \text{ s}$, using a time step lower than this value produces space oscillations.

The last simulation shows that our simple model makes it possible to represent an induction quenching sequence. Indeed, induction quenching consists of two phases. First, the plate is heated rapidly by induction, so that the convection exchanges are negligible. Then the heating system is turned off and the plate is cooled abruptly, for instance with an atomized jet of water projected onto the face $x = E$. This can be illustrated as follows:

- between instants $t = 0$ and $t = t_0$, the heating period is given by $Q_0(t) = Q_0$ and $h(t) = 0$. An exchange coefficient equal to zero comes down to assume that there is no exchange in $x = E$, between the piece and the external medium. This is a simplification of the problem of course, in which the exchange is considered to be small enough to be neglected;

- from instant $t = t_0$, the quenching period is given by $Q_0(t) = 0$ and $h(t) = H$. The value of H will depend on the quenching medium considered (air, water, atomized water, etc.).

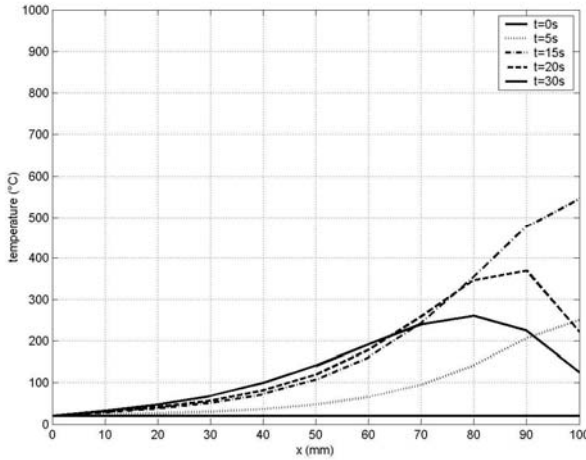


Figure 4.10. *Induction quenching – temperature profile over time*

The numerical values used are the following:

- constant exchange coefficient (when it is different from zero): $H = 6,000 \text{ W/m}^2/\text{K}$;
- heating power (when it is different from zero): $Q_0 = 2.10^8 \text{ W/m}^3$;
- plate initial temperature: $T_0 = 20^\circ\text{C}$;
- time integration scheme: $\nu = 1$;
- heating time: $t_0 = 15 \text{ s}$.

Figure 4.10 is a good illustration of the temperature profiles obtained during induction quenching. The heating phase produces a gradual increase of the temperature at the plate surface (in $x = E$) and a positive thermal gradient. This thermal gradient strongly depends on the depth P introduced to model induction heating. This depth depends on the induction current frequency. A high frequency will produce surface heating and a high thermal gradient (low value of P), whereas a lower frequency will heat the plate in its thickness (high value of P) with a lower thermal gradient as well. Then the quenching phase produces rapid cooling of the surface, and a local sign change in the thermal gradient. The cooling rapidity depends directly on the thermal exchange coefficient H .

Chapter 5

Non-linearities

5.1. Formulation and solution techniques

In the previous chapters, we introduced the finite element method applied to the solution of permanent or transient heat conduction problems. Generally speaking we assumed that the thermal and physical properties of the materials concerned did not depend on the temperature and restricted our analysis to boundary conditions expressed linearly according to the temperature by means of a constant exchange coefficient. Applying the finite element method then leads to the solution, possibly at different moments, of a linear equation system. We then speak of a linear problem solution.

Practically speaking, the various parameters (thermal and physical properties, boundary conditions, etc.) involved in mathematical modeling are rarely expressed so simply. Such is the case when the application to simulate covers an extensive temperature range or involves phase changes. The problem formulation no longer produces a linear equation system and the solution techniques have to be appropriate.

5.1.1. Formulation

The finite element method applied to a thermal conduction problem leads to the cancelation of a residual vector $\{R(T)\}$ obtained by assembling element residuals $\{R^e(T)\}$ (equation [2.6] in Part 1 of this book for the steady state, equation [4.7] in Part 2 for the transient state):

$$\{R(T)\} = \sum_{e=1}^m [A^e]^T \cdot \{R^e(T)\} = \{0\} \quad [5.1]$$

In a linear problem, we saw that this residual can be expressed as follows:

$$\begin{cases} \{R(T)\} = \{F\} - [K] \cdot \{T\} & \text{(steady state)} \\ \{R(T)\} = \{F\} - [K] \cdot \{T\} - [C] \cdot \{\dot{T}\} & \text{(transient state)} \end{cases}$$

In these expressions, the thermal conductance matrix $[K]$, the capacitance matrix $[C]$ and the load vector $\{F\}$ are independent of the temperature. It follows that in this case the finite element method leads to the solution of linear equation systems. In the transient state case, this system is obtained after applying a time integration scheme such as that of generalized central differences.

When, for instance, the thermal conductivity λ or the specific heat C depend on temperature, matrices $[K]$ and $[C]$ also depend on temperature and the residual vector becomes a non-linear function of the nodal temperature vector $\{T\}$ which can be written as follows:

$$\begin{cases} \{R(T)\} = \{F\} - [K(T)] \cdot \{T\} & \text{(steady state)} \\ \{R(T)\} = \{F\} - [K(T)] \cdot \{T\} - [C(T)] \cdot \{\dot{T}\} & \text{(transient state)} \end{cases}$$

We then obtain a non-linear equation system whose unknowns are the temperatures at the mesh nodes.

In fact, strictly speaking, the size of the non-linear system to solve [5.1] is not always exactly equal to the number of mesh nodes. Indeed, as in the linear case, the temperatures of certain nodes can be prescribed. This is even necessary in the case of a steady state problem to obtain a single solution. In the non-linear case, the application techniques for this type of condition are the same as those already introduced in Part 1 of this book. Therefore, we will not mention this aspect hereafter and assume that the non-linear system [5.1] is expressed in the form of n equations with n unknowns, where n is the number of mesh nodes, possibly reduced with the elimination method by the number of nodes at which the temperature is known, or increased by this number of nodes with the Lagrange multipliers method.

5.1.2. Non-linear equation system solution methods

In order to solve equation [5.1], most computer codes follow two steps. In the first, they try to predict the solution and in the second, they solve the non-linear equation system iteratively as indicated in Figure 5.1 using the prediction of step one as an initial estimation.

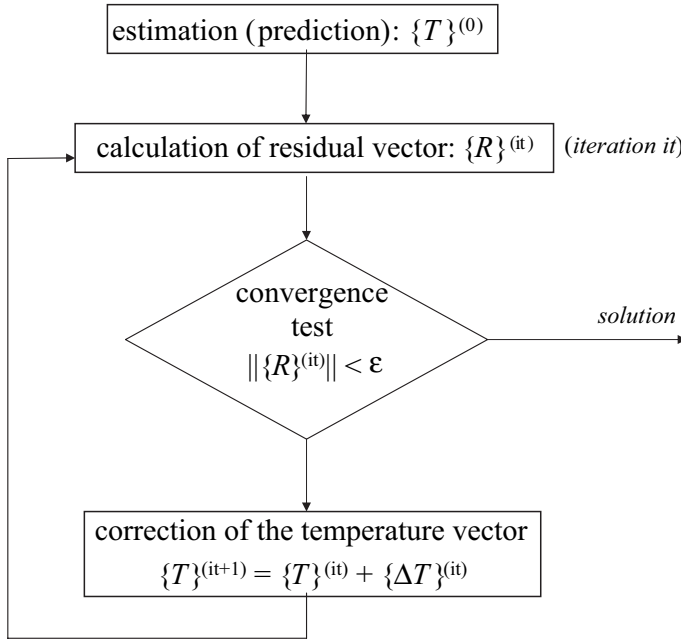


Figure 5.1. *Non-linear equation system solution principle*

The estimation $\{T\}^{(0)}$ plays an important part in the iterative process convergence. In a steady state problem, it is difficult to obtain a reliable estimation in all possible cases. Conversely, an estimation can be obtained at any moment in a transient problem from the solutions at the previous moments. For instance, it is possible to choose a linear extrapolation of the values calculated at the two previous time steps $\{T(t - \Delta t)\}$ and $\{T(t)\}$:

$$\{T\}^{(0)} = \{T(t + \Delta t)\}^{(0)} = 2\{T(t)\} - \{T(t - \Delta t)\}$$

or a quadratic extrapolation of the values calculated at the three previous steps $\{T(t - 2\Delta t)\}$, $\{T(t - \Delta t)\}$ and $\{T(t)\}$:

$$\{T\}^{(0)} = \{T(t + \Delta t)\}^{(0)} = 3(\{T(t)\} - \{T(t - \Delta t)\}) + \{T(t - 2\Delta t)\}$$

At each iteration it of the process described in Figure 5.1, we have an estimation of the solution $\{T\}^{(it)}$. Then the residual vector $\{R\}^{(it)}$ corresponding to it is calculated.

A convergence test is then carried out by comparing the norm of the residual $\{R\}^{(it)}$ with a threshold value ϵ_R given by the user. The norm considered can be

the Euclidian norm (root-sum square of each component of the vector), but some computer codes give other measurements like the vector's largest component in absolute value (called the "infinite norm"). If the residual vector norm is less than the limit value prescribed by the user, the algorithm is then considered to have converged.

If not, a correction $\{\Delta T\}^{(it)}$ of the nodal temperature vector is calculated in order to approximate the optimal solution. This correction is obtained by solving a linear system of the following type:

$$[M]^{(it)} \cdot \{\Delta T\}^{(it)} = \{R\}^{(it)} \quad [5.2]$$

In this equation, $[M]^{(it)}$ is a matrix of dimension $n \times n$ depending on the solution method used, as will be seen in the next section. The nodal temperature vector is then updated as follows before being used during the iteration $it + 1$:

$$\{T\}^{(it+1)} = \{T\}^{(it)} + \{\Delta T\}^{(it)} \quad [5.3]$$

If, during these iterations, the norm of the vector $\{\Delta T\}^{(it)}$ is less than a threshold ϵ_T defined by the user, most computer codes consider that the iterative process has converged. However, we must bear in mind that the only reliable convergence indicator is the residual norm value. As a matter of fact, from a mathematical point of view, the condition $\|\{\Delta T\}^{(it)}\| \leq \epsilon_T$ shows that the sequence made up with $\{T\}^{(it)}$ is a Cauchy sequence. Moreover, we know that in \mathcal{R}^n , any Cauchy sequence converges. However, the criterion $\|\{\Delta T\}^{(it)}\| \leq \epsilon_T$ does not indicate that the sequence of $\{T\}^{(it)}$ is very close to its limit. In all cases, it would be wise to check that the value of the residual vector norm is small enough to be acceptable.

It is important to note that the choice of the matrix $[M]^{(it)}$ in equation [5.2] does not interfere with the solution quality but only with the speed at which it can be obtained. In fact there is no method guaranteeing the convergence of the iterative process in all cases. Therefore, choosing the best method to solve [5.1] as effectively as possible is not simple. This is all the more difficult as the number of iterations is not the only criterion. The time required for assembling $[M]^{(it)}$ and solving linear system [5.2] has a part to play as well. The user will always choose a strategy for solving a given problem, according to his/her experience and the nature of this problem (non-linearities of thermal and physical properties, phase change problem, thermal radiation, etc.).

The main methods are illustrated in Figure 5.2 for a problem with one degree of freedom. The reader may go into the detail of this point with the help of specialized books [DEN 76].

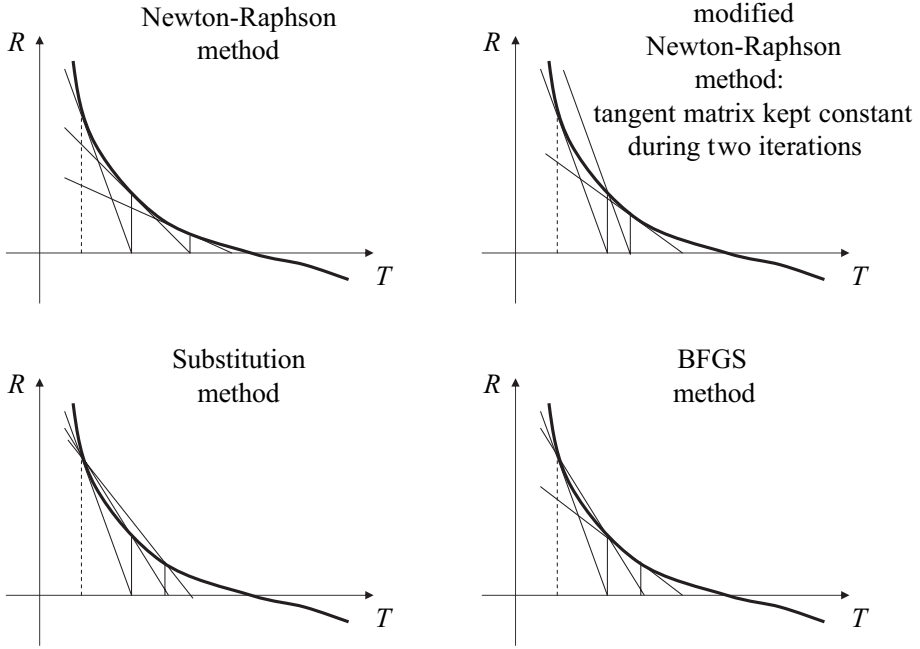


Figure 5.2. Most common methods for the solution of a non-linear equation

5.1.2.1. Newton-Raphson method

This method consists of determining the temperature correction from a first order Taylor expansion of the residual in the vicinity of the current solution, which means that the residual can be differentiated close to the solution sought:

$$R_i = R_i^{(it)} + \sum_{j=1}^n \frac{\partial R_i}{\partial T_j} (T_j - T_j^{(it)})$$

In that expansion, the components of the matrix involved are the partial derivatives of the components of the residual vector, with respect to the components of the vector's nodal temperatures calculated for $\{T\} = \{T\}^{(it)}$. A linear approximation of the residual close to the current solution is therefore obtained. The correction $\{\Delta T\}^{(it)}$ is calculated in order to cancel this approximation of the residual:

$$\{R\}^{(it)} + \left[\frac{\partial R}{\partial T} \right]^{(it)} \cdot \{\Delta T\}^{(it)} = \{0\}$$

The method therefore consists of selecting as matrix $[M]^{(it)}$ the tangent matrix of the non-linear system to solve. This tangent matrix is obtained at each iteration by assembling element tangent matrices:

$$[M] = \sum_{e=1}^m [A^e]^T \cdot [M^e] \cdot [A^e] \quad \text{with} \quad M_{ij}^e = -\frac{\partial R_i^e}{\partial T_j^e}$$

The Newton-Raphson method converges in a quadratic way. To demonstrate this, consider the non-linear scalar equation $r(u) = 0$, where r is a derivable function near the solution. Now apply Newton-Raphson's iterative process. In this case, the recurrence relationship between two consecutive estimations gives:

$$u^{(it+1)} = u^{(it)} - \frac{r^{(it)}}{r'(it)} \quad \text{with} \quad r^{(it)} = r(u^{(it)}) \text{ and } r'(it) = \frac{dr}{du}(u^{(it)})$$

Denoting by $e^{(it)} = u^{(it)} - u$ the error obtained during the iteration (it) , u being the exact solution, we obtain:

$$e^{(it+1)} = e^{(it)} - \frac{r^{(it)}}{r'(it)}$$

Now assume the function r to be derivable twice near the solution. A development of r near $u^{(it)}$ then gives:

$$\begin{aligned} r(u) &= r^{(it)} + r'(it)(u - u^{(it)}) + \frac{1}{2}r''(it)(u - u^{(it)})^2 + \dots \\ &= r^{(it)} - r'(it)e^{(it)} + \frac{1}{2}r''(it)e^{(it)^2} + \dots \\ &= 0 \end{aligned}$$

hence

$$e^{(it)} - \frac{r^{(it)}}{r'(it)} = \frac{1}{2}e^{(it)^2} \frac{r''(it)}{r'(it)} + \dots$$

i.e., if we restrict our operation to the second-order terms:

$$e^{(it+1)} = \frac{1}{2}e^{(it)^2} \frac{r''(it)}{r'(it)} \quad \text{with} \quad r''(it) = \frac{d^2r}{du^2}(u^{(it)})$$

Therefore, there is a constant $k > 0$ such that $e^{(it+1)} = ke^{(it)^2}$, hence the quadratic convergence.

The Newton-Raphson method therefore converges with, in general, relatively few iterations. However, the assembly of it requires matrix $[M]^{it}$ as well as the solution of a new linear system of type [5.2] at each iteration. It is therefore very costly from a practical point of view and must be reserved for strongly non-linear problems for which it has been impossible to reach convergence with another method.

Variants of this method, called modified Newton methods, keep the matrix $[M]^{(it)}$ constant during a certain number of iterations, before recalculating it. The assembly of $[M]^{(it)}$ is then performed only for certain iterations. Moreover, using a direct method based upon the matrix triangulation to solve linear system [5.2] can appear to be particularly efficient insofar as triangulation is performed only when the matrix is modified. The adverse effect is that the convergence domain of these methods is restricted compared with the Newton-Raphson method. In addition, these methods require more iterations to converge than the Newton-Raphson method does.

5.1.2.2. Substitution method

In a large number of situations, the residual vector can be written as follows:

$$\{R(T)\} = \{F(T)\} - [M(T)] \cdot \{T\} \quad [5.4]$$

Such is the case of steady state or transient heat conduction problems, for which we have:

$$\begin{cases} [M(T)] = [K(T)] & \text{(steady state)} \\ [M(T)] = [K(T)] - \frac{1}{\nu \Delta t} [C(T)] & \text{(transient state)} \end{cases}$$

The substitution method consists of taking matrix $[M(T^{(it)})]$ for matrix $[M]^{(it)}$.

By combining equations [5.2], [5.3] and [5.4], we demonstrate that this method comes down to choosing $\{T\}^{(it+1)}$ as follows:

$$\{T\}^{(it+1)} = [M(T^{(it)})]^{-1} \cdot \{F(T^{(it)})\} \quad [5.5]$$

The substitution method can also be considered as a fixed point method. Sometimes it is also called the Picard iterative method.

Like the Newton-Raphson method, it requires the assembly and solution of a linear system at each iteration and is therefore very costly. Its interest lies in the

fact that the matrix $[M]^{(it)}$ calculated in this way is generally symmetric whereas the tangent matrix calculated in the Newton-Raphson method is not necessarily so, which generates additional storage and solution costs.

5.1.2.3. *Quasi-Newton methods*

Quasi-Newton methods generalize to non-linear equation systems the secant iterative method well-known for solving a non-linear equation of the type $f(x) = 0$.

The principle of these methods is to replace the tangent matrix with an approximation which can be obtained more simply. The aim is to obtain convergence properties close to those of the Newton-Raphson method while avoiding the assembly and possibly the triangulation of a linear system at each iteration. Generally speaking, rather than construct an approximate tangent matrix, it is preferable to directly seek an approximation of its inverse.

Therefore, we will seek a matrix $[S]^{(it)} = ([M]^{(it)})^{-1}$ satisfying the following (secant) condition:

$$\{T\}^{(it)} - \{T\}^{(it-1)} = [S]^{(it)} \cdot (\{R\}^{(it-1)} - \{R\}^{(it)}) \quad [5.6]$$

On the other hand, we prescribe that two consecutive matrices $[M]^{(it-1)}$ and $[M]^{(it)}$ differ only in the direction $\{\Delta T\}^{(it-1)} = \{T\}^{(it)} - \{T\}^{(it-1)}$. Therefore, we have:

$$\begin{aligned} \forall \{U\} \text{ such that } \langle \Delta T \rangle^{(it-1)} \cdot \{U\} &= 0, \\ [M]^{(it)} \cdot \{U\} &= [M]^{(it-1)} \cdot \{U\} \end{aligned} \quad [5.7]$$

Allow $\{\Delta R\}^{(it-1)} = \{R\}^{(it-1)} - \{R\}^{(it)}$ and seek $[S]^{(it)}$ as follows:

$$[S]^{(it)} = ([I] + [A]^{(it)}) \cdot [S]^{(it-1)} \quad [5.8]$$

where $[I]$ is the identity matrix.

It is easy to demonstrate that condition [5.7] then becomes:

$$\forall \{U\} \text{ such that } \langle \Delta T \rangle^{(it-1)} \cdot \{U\} = 0, \quad [A]^{(it)} \cdot \{U\} = \{0\} \quad [5.9]$$

which leads us to write:

$$[A]^{(it)} = \{W\}^{(it-1)} \cdot \langle \Delta T \rangle^{(it-1)}$$

where $\{W\}^{(it-1)}$ is a vector which is still to be determined.

Considering equation [5.6] again, we obtain:

$$\{\Delta T\}^{(it-1)} = ([I] + \{W\}^{(it-1)} \cdot \langle \Delta T \rangle^{(it-1)}) \cdot [S]^{(it-1)} \cdot \{\Delta R\}^{(it-1)}$$

from which we immediately derive:

$$\{W\}^{(it-1)} = \frac{\{\Delta T\}^{(it-1)} - [S]^{(it-1)} \cdot \{\Delta R\}^{(it-1)}}{\langle \Delta T \rangle^{(it-1)} \cdot [S]^{(it-1)} \cdot \{\Delta R\}^{(it-1)}} \quad [5.10]$$

The correction $\{\Delta T\}^{(it)}$ to make to the nodal temperature vector $\{T\}^{(it)}$ is then written as follows, considering [5.2] and [5.8]:

$$\{\Delta T\}^{(it)} = ([I] + \{W\}^{(it-1)} \cdot \langle \Delta T \rangle^{(it-1)}) \cdot [S]^{(it-1)} \cdot \{R\}^{(it)} \quad [5.11]$$

Among the various existing quasi-Newton methods, the BFGS method is known to be one of the most efficient. Matrices $[S]^{(it)}$ are calculated so that:

- if $[S]^{(it-1)}$ is symmetric, then so is $[S]^{(it)}$;
- if $[S]^{(it-1)}$ is positive definite, then so is $[S]^{(it)}$.

By assuming $[S]^{(1)}$ to be positive definite, matrices $[S]^{(it)}$ can be sought as follows:

$$\begin{aligned} [S]^{(it)} &= ([I] + \{W\}^{(it-1)} \langle V \rangle^{(it-1)}) \cdot [S]^{(it-1)} \\ &\quad \cdot ([I] + \{V\}^{(it-1)} \langle W \rangle^{(it-1)}) \end{aligned} \quad [5.12]$$

If the same course of action as above is taken, it is then demonstrated that:

$$\begin{aligned} \{W\}^{(it-1)} &= \frac{\{\Delta T\}^{(it-1)}}{\langle \Delta R \rangle^{(it-1)} \cdot \{\Delta T\}^{(it-1)}} \\ \{V\}^{(it-1)} &= \{R\}^{(it)} - \frac{\langle \Delta R \rangle^{(it-1)} \cdot \{\Delta T\}^{(it-1)}}{\langle \Delta T \rangle^{(it-1)} \cdot \{R\}^{(it-1)}} \{R\}^{(it-1)} \end{aligned}$$

The temperature correction $\{\Delta T\}^{(it)}$ is written as follows:

$$\begin{aligned} \{\Delta T\}^{(it)} = & \left(\prod_{k=it-1}^1 \left([I] + \{W\}^{(k)} \langle V \rangle^{(k)} \right) \right) \cdot [S]^{(1)} \\ & \cdot \left(\prod_{k=1}^{it-1} \left([I] + \{V\}^{(k)} \langle W \rangle^{(k)} \right) \right) \cdot \{R\}^{(it)} \end{aligned} \quad [5.13]$$

In the equation above, it is obvious that the calculation of the temperature correction does not require a linear system solution or even the explicit calculation of matrices $[S]^{(it)}$. This makes the method particularly efficient in practical cases. If the computer code includes the BFGS method, it is most certainly the method to use first and foremost to solve any non-linear problem.

5.1.3. Line search method

Once $\{\Delta T\}^{(it)}$ has been obtained by means of equation [5.2], the *line search* method consists of trying to cancel the residual in that direction. By considering the 1D vectorial space generated by direction $\{\Delta T\}^{(it)}$, this comes down to canceling the projection of the residual $\{R\}$ in that space, i.e. canceling the scalar function f of the scalar variable α as follows:

$$f(\alpha) = \sum_{i=1}^n \Delta T_i^{(it)} R_i^{(it)} \left(\{T\}^{(it)} + \alpha \{\Delta T\}^{(it)} \right)$$

Function f being a non-linear function with real values, we seek α such that $f(\alpha) = 0$ by means of a secant iterative method. At each iteration (k) , a new value of α is obtained by:

$$\alpha^{(k+1)} = \alpha^{(k)} - \frac{\alpha^{(k)} - \alpha^{(k-1)}}{f(\alpha^{(k)}) - f(\alpha^{(k-1)})} f(\alpha^{(k)})$$

The process stops when $|f(\alpha^{(k)})| \leq \eta |f(0)|$, or when $|\alpha^{(k+1)} - \alpha^{(k)}| \leq \delta$, the coefficients η and δ being thresholds defined by the user.

Note that this method is substantially more costly as it requires a new estimation of the residual vector at each iteration (k) of the *line search* problem solution. This method is often combined with quasi-Newton or modified Newton methods.

A similar method, called the under-relaxation method, is sometimes used to solve non-linear systems. It consists of reducing the modification of the unknown to a fraction of the quantity $\{\Delta T\}^{(it)}$ calculated by equation [5.2], so as to satisfy an additional condition (for instance, all the components of the unknown must remain positive) or to better control strong variations in the solution.

5.2. Traditional non-linearities

As mentioned earlier on in this chapter, traditional non-linearities in a thermal conduction problem are:

- physical properties depending on the temperature;
- flux boundary conditions or a volumetric heat source depending on the temperature.

With regard to physical properties, we saw in previous chapters that the problem becomes non-linear as soon as the thermal conductivity λ , the volumetric mass ρ , or the mass heat C , of the material is a function of the temperature. In the case of flux q boundary conditions or the volumetric heat source Q , the problem remains linear as long as these very functions remain linear according to the temperature.

There are other causes for the non-linearity of the problem. For example, the case of the thermal contact processed with a contact resistance most often leads to a non-linear problem. In this section, we will restrict our study to classic non-linearities.

5.2.1. Physical properties

The case of temperature-dependent physical properties is relatively simple.

In the case of a temperature-dependent thermal conductivity, $\lambda(T)$, the term involved in the element residual expression for steady state or transient cases is:

$$\begin{aligned} R_i^{(\lambda)} &= - \int_{\Omega^e} \overrightarrow{\text{grad}}^T(N_i^e) \cdot \underline{\lambda}(T) \cdot \overrightarrow{\text{grad}}(T) dv \\ &= - \sum_j \left(\int_{\Omega^e} \overrightarrow{\text{grad}}^T(N_i^e) \cdot \underline{\lambda}(T) \cdot \overrightarrow{\text{grad}}(N_j^e) dv \right) T_j^e \end{aligned}$$

which is no longer a linear function of the vector $\{T^e\}$ including the node temperatures of element Ω^e .

If, for example, a Newton-Raphson method is used for the solution, we obtain a tangent matrix including the following term:

$$M_{ij}^{(\lambda)} = -\frac{\partial R_i^{(\lambda)}}{\partial T_j^e} = \begin{cases} \int_{\Omega^e} \overrightarrow{\text{grad}}^T(N_i^e) \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(N_j^e) dv \\ + \int_{\Omega^e} \overrightarrow{\text{grad}}^T(N_i^e) \cdot \frac{d\underline{\lambda}}{dT} \cdot \overrightarrow{\text{grad}}(T) N_j^e dv \end{cases}$$

In addition to a part similar to that already encountered in the linear case, we can note in this expression the appearance of a second term due to the fact that the thermal conductivity changes with the temperature. This term is not symmetric to indices i and j and so as to avoid the necessity of processing non-symmetric matrices that require appropriate storage and solution methods, the term is neglected. This comes down to considering the part related to thermal conductivity as a “substitution type” matrix. This approximation is all the more reasonable as the non-linearity introduced by the temperature-dependent conductivity is generally weak.

In the case of a transient calculation, the volumetric mass ρ and the specific heat capacity C of the material can also depend on the temperature. In that case, the product ρC involved in the element residual expression is a function of the temperature and the term involved is:

$$R_i^{(\rho C)} = - \int_{\Omega^e} N_i^e \rho(T) C(T) \dot{T} dv$$

If a generalized trapezoidal rule is used to integrate the differential equation system obtained between instants t and $t + \Delta t$, we obtain:

$$\dot{T}(t + \Delta t) = -\frac{1-\nu}{\nu} \dot{T}(t) + \frac{1}{\nu \Delta t} (T(t + \Delta t) - T(t)) \quad \text{with } \nu \in]0, 1]$$

It follows from this that the residual parts in which the product ρC is involved are written as follows (if we write $T = T(t + \Delta t)$):

$$R_i^{(\rho C)}(t + \Delta t) = \begin{cases} \frac{1-\nu}{\nu} \int_{\Omega^e} N_i^e \rho(T) C(T) \dot{T}(t) dv \\ - \frac{1}{\nu \Delta t} \int_{\Omega^e} N_i^e \rho(T) C(T) (T - T(t)) dv \end{cases}$$

If, for instance, a Newton-Raphson method is used to solve the non-linear equation system obtained, a tangent matrix is calculated in which the contribution of the product ρC will be involved as follows:

$$M_{ij}^{(\rho C)} = -\frac{\partial R_i^{(\rho C)}}{\partial T_j^e} = \begin{cases} \frac{1}{\nu \Delta t} \int_{\Omega^e} \rho C N_i^e N_j^e dv \\ + \int_{\Omega^e} N_i^e \frac{\partial(\rho C)}{\partial T} \left(-\frac{1-\nu}{\nu} \dot{T}(t) + \frac{1}{\nu \Delta t} (T - T(t)) \right) N_j^e dv \end{cases}$$

Note that this expression is composed of two symmetric terms at i and j . In fact, these terms can be calculated in the computer codes. Indeed, they produce no dissymmetry in the tangent matrix. Moreover, specific temperature-dependent heat changes may be relatively significant. Such is the case when phase change latent heat is represented via an equivalent specific heat capacity (see section 5.2.3).

5.2.2. Flux or volumetric heat source boundary conditions

The surface flux densities q on the boundary $\partial\Omega_q$ of the solid, as well as the volumetric heat source term Q , may be temperature-dependent in a non-linear way. In that case, the problem to solve is non-linear as well. Among the problems to be considered are:

- the interactions between a coolant and a material. In the case of steel water quenching, such interactions are often modeled by an exchange coefficient varying strongly in a non-linear way according to the temperature of the piece surface. To illustrate this variation, we can say that this coefficient is low when the surface temperature is high (calefaction state), then increases abruptly when the burnout film is pierced by liquid droplets coming to the surface (nucleate boiling state), and finally decreases again when the surface temperature goes under the boiling temperature of the water (convection state);
- the material radiation effects toward an external environment, often expressed by an equivalent exchange coefficient varying strongly according to the temperature. These effects are important when the solid temperature is typically above 300°C. They are covered in detail in Chapter 7;
- the Curie transition point during steel induction heating. Above that temperature, the induction effect on the material volume changes substantially, so that the volumetric heat source used becomes a strongly non-linear function of the temperature.

When the flux density on the boundary $\partial\Omega_q$ of the solid and/or the volumetric heat source become non-linear functions of the temperature, i.e. $q(T)$ and $Q(T)$, the corresponding terms in the element residual also become non-linear:

$$R_i^{(Q)} = \int_{\Omega^e} N_i^e Q(T) dv$$

$$R_i^{(q)} = \int_{\partial\Omega^e \cap \partial\Omega_q} N_i^e q(T) ds$$

Using a tangent matrix for the solution then leads to considering the variation of q and Q with the temperature. This is expressed as follows:

$$M_{ij}^{(Q)} = -\frac{\partial R_i^{(Q)}}{\partial T_j^e} = -\int_{\Omega^e} N_i^e \frac{\partial Q}{\partial T} N_j^e dv$$

$$M_{ij}^{(q)} = -\frac{\partial R_i^{(q)}}{\partial T_j^e} = -\int_{\partial\Omega^e \cap \partial\Omega_q} N_i^e \frac{\partial q}{\partial T} N_j^e ds$$

These expressions create no particular problem for the solution. They keep the tangent matrix symmetric.

For instance, in the case of radiation exchange in an infinite environment, we have $q(T) = \sigma\epsilon(T_\infty^4 - T^4)$, where ϵ is the wall emissivity and σ the Stefan constant. The tangent matrix is written as follows:

$$M_{ij}^{(q)} = \int_{\partial\Omega^e \cap \partial\Omega_q} 4\sigma\epsilon T^3 N_i^e N_j^e ds$$

A substitution matrix can also be defined by writing $q(T) = \sigma\epsilon(T_\infty^2 + T^2)(T_\infty + T)(T_\infty - T)$. This comes down to defining an equivalent exchange coefficient $H(T) = \sigma\epsilon(T_\infty^2 + T^2)(T_\infty + T)$. The substitution matrix is written as follows:

$$M_{ij}^{(q)} = \int_{\partial\Omega^e \cap \partial\Omega_q} \sigma\epsilon(T_\infty^2 + T^2)(T_\infty + T) N_i^e N_j^e ds$$

Using such a matrix is of no significant interest in this case, insofar as the tangent matrix, which is more efficient, is already symmetric.

5.2.3. Modeling state changes

State change problems are traditional in thermal science. To make things clear, a solidification problem will be examined hereafter.

When dealing with a pure substance, starting with the liquid state, solidification occurs at a given temperature expressed as T_f . Then a Stefan problem has to be treated; the equations to solve are written as follows:

- in the domain occupied by the liquid phase Ω_l

$$\rho_l C_l \frac{\partial T}{\partial t} - \text{div} (\lambda_l \overrightarrow{\text{grad}}(T)) = 0$$

- in the domain occupied by the solid phase Ω_s

$$\rho_s C_s \frac{\partial T}{\partial t} - \text{div} (\lambda_s \overrightarrow{\text{grad}}(T)) = 0$$

- at the liquid-solid interface

$$T_l = T_s = T_f \quad [5.14]$$

$$\lambda_s \left(\frac{\partial T}{\partial n_{s/l}} \right)_s - \lambda_l \left(\frac{\partial T}{\partial n_{s/l}} \right)_l = (\rho L)_{s/l} V_{s/l} \quad [5.15]$$

In the above equations, indices l and s correspond respectively to the liquid phase and to the solid phase, $n_{s/l}$ represents the norm at the liquid/solid interface from the solid to the liquid, $(\rho L)_{s/l}$ is the fusion/solidification latent heat per volume unit and $V_{s/l}$ the normal component of the liquid/solid interface moving speed.

The numerical solution of such a problem requires front tracking methods. These can be combined with various numerical methods like the boundary integral equation method, which is particularly interesting since only the mobile interface is meshed, or the finite element method. The finite element formulation of such a problem requires remeshing techniques. At each moment, the mesh is composed of 2 sub-domains, one for the liquid and the other for the solid, interconnected by a set of nodes located on the mobile interface. For each sub-domain, in addition to the classic capacitance and conduction terms, the problem formulation reveals a transport term (see Chapter 6) related to the movements of the nodes located on the mobile interface. The interface position is updated by solving equations [5.14] and [5.15].

However, for most metal alloys used in industry, solidification does not occur at a given temperature but on a range $[T_s, T_l]$, T_s and T_l representing respectively the *solidus* and *liquidus* temperatures. The interface between the two phases is no longer clear. The transformation takes place in an intermediate zone called a mushy zone. The phase change latent heat can then be introduced as a source term as follows:

$$Q = (\rho L)_{s/l} \frac{df_s}{dt} \quad [5.16]$$

where f_s represents the volumetric fraction of the solid. In the solidification case, f_s is generally assumed to depend on the temperature only, as follows:

$$f_s = \begin{cases} 1 & \text{if } T \leq T_s \\ \frac{T - T_l}{T_s - T_l} & \text{if } T_s < T < T_l \\ 0 & \text{if } T \geq T_l \end{cases}$$

If we assume that we operate in a constant stress environment, we can introduce the volumetric enthalpy H of the medium with:

$$H(T) = (\rho L)_{s/l} (1 - f_s) + \int_{T_{ref}}^T \rho C dT' \quad [5.17]$$

It is obvious that equation [5.17] immediately gives:

$$\dot{H} = \frac{dH}{dT} \dot{T} = -(\rho L)_{s/l} \dot{f}_s + \rho C \dot{T}$$

so that the heat equation with the source term [5.16] can now be written:

$$\dot{H} - \text{div} (\underline{\lambda} \cdot \overrightarrow{\text{grad}}(T)) = 0 \quad [5.18]$$

5.2.3.1. Equivalent specific heat method

Starting with equation [5.17], we can formally define an equivalent specific heat with:

$$C_{eq} = \frac{1}{\rho} \frac{dH}{dT} = C - \frac{(\rho L)_{s/l}}{\rho} \frac{df_s}{dT} \quad [5.19]$$

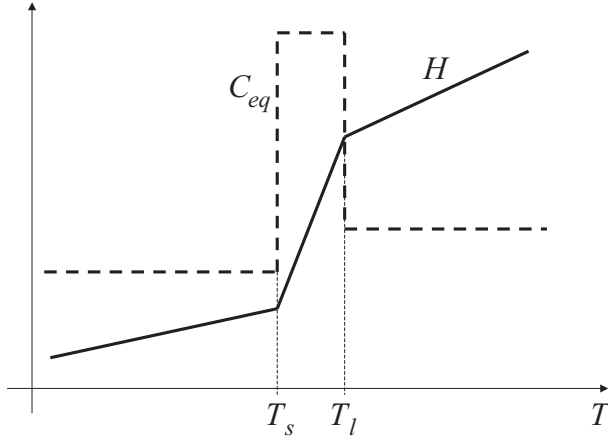


Figure 5.3. *Enthalpy and equivalent specific heat*

In fact the equivalent specific heat is the usual specific heat modified for $T_s < T < T_l$ so as to include the state change latent heat. Equation [5.18] can now be written in the usual way:

$$\rho C_{eq} \dot{T} - \text{div} (\vec{\lambda} \cdot \overrightarrow{\text{grad}}(T)) = 0 \quad [5.20]$$

The advantage of this method is its simplicity. The phase change is treated as a non-linearity of physical properties. However, it may create some numerical difficulties related to the abrupt changes of C_{eq} at temperatures T_s and T_l , as illustrated in Figure 5.3. Besides the real consideration of the latent heat requires an accurate integration of $C_{eq}(T)$ and, consequently, imposes very small time steps. One way of reducing these inaccuracies is to smooth the specific heat on a wider temperature interval $[T'_s, T'_l]$. In that case however, we have to ensure that:

$$\int_{T'_s}^{T'_l} \rho C_{eq} dT = H(T'_l) - H(T'_s)$$

The time step chosen for the analysis will have to be small enough to integrate the above equation accurately.

To avoid the problems mentioned above, some authors ([COM 74, LEW 87]) operate in a different way. Starting with the fact that enthalpy is a more regular function of the temperature than the specific heat, they suggest we estimate an equivalent specific heat from the enthalpy and temperature space variations.

A nodal approximation of enthalpy, similar in nature to that of temperature, is first made from the enthalpy values at the nodes directly deduced from the temperature values:

$$H = \sum_{i=1}^{n^e} N_i^e(\vec{x}) H_i$$

The relationship:

$$\rho C_{eq} = \frac{dH}{dT}$$

is then approximated with one of the following formulae:

$$\rho C_{eq} \approx \frac{1}{3} \sum_{i=1}^3 \frac{\partial H / \partial x_i}{\partial T / \partial x_i}$$

or otherwise, according to [LEM 81]:

$$\rho C_{eq} \approx \frac{\left(\sum_{i=1}^3 (\partial H / \partial x_i)^2 \right)^{1/2}}{\left(\sum_{i=1}^3 (\partial T / \partial x_i)^2 \right)^{1/2}}$$

However, it has been demonstrated [MOR 78] that such an approximation was unnecessary and an estimation of ρC_{eq} at instant $n\Delta t$ with:

$$\rho C_{eq} \approx \frac{H_n - H_{n-1}}{T_n - T_{n-1}}$$

yields excellent results. Enthalpy solution methods are addressed here. In [DAL 86] a comparative study of the various methods for the solution of solidification problems will be found.

5.2.3.2. Enthalpy solution method

Actually, it is not necessary to use a specific heat. Rather than apply a time integration algorithm such as [4.26] or [4.27] to the term \dot{T} of equation [5.20], it is possible to treat the term \dot{H} of equation [5.18] directly. Equation [5.18] thus becomes, after applying an implicit Euler algorithm:

$$\frac{H(T) - H_0}{\Delta t} - \text{div} \left(\vec{\lambda}(T) \cdot \overrightarrow{\text{grad}}(T) \right) = 0 \quad [5.21]$$

an equation where $H = H(T)$ is the enthalpy at $t + \Delta t$ and $H_0 = H(T_0)$, the enthalpy at instant t .

The finite element method applied to the solution of equation [5.21] then produces the following expression of the element residual:

$$R_i^e = - \int_{\Omega^e} \overrightarrow{\text{grad}}^T(N_i^e) \cdot \overrightarrow{\lambda}(T) \cdot \overrightarrow{\text{grad}}(T) dv \\ - \int_{\Omega^e} N_i^e \frac{H(T) - H_0}{\Delta t} dv$$

The enthalpy H is now directly involved in the expression of the element residual via the term:

$$R_i^{(H)} = - \int_{\Omega^e} N_i^e \frac{H(T) - H_0}{\Delta t} dv \quad [5.22]$$

Then the problem can be solved with one of the methods introduced in section 5.1.2. When a Newton-Raphson method is used, the contribution to the tangent matrix introduced by the enthalpy term is the following:

$$M_{ij}^{(H)} = - \frac{\partial R_i^{(H)}}{\partial T_j^e} = \frac{1}{\Delta t} \int_{\Omega^e} N_i^e \frac{dH(T)}{dT} N_j^e dv \quad [5.23]$$

It is obvious that the difference between the enthalpy method and the equivalent specific heat method results from the discretization of the term \dot{H} . The enthalpy change taken into consideration by the enthalpy method is accurate:

$$\int_t^{t+\Delta t} \dot{H} dt' = H - H_0$$

whereas it is only approximated with the equivalent specific heat method:

$$\int_t^{t+\Delta t} \dot{H} dt' = C_{eq}(T)(T - T_0)$$

Therefore, the enthalpy method is more accurate.

However, using an enthalpy approximation requires storing the enthalpy H , as well as \dot{H} when using an algorithm with $0 < \nu < 1$, at each integration point of

the elements. Besides, with traditional Gauss integration schemes, we may obtain oscillating solutions as the discrete maximum principle might not be satisfied. As seen in Chapter 4 concerning transient heat conduction, one way of eliminating these oscillations with first order elements, is to use a lumped capacitance matrix. In the case of the enthalpy approximation, this comes down to use a scheme in which the integration points would be located at the element nodes to integrate terms [5.22] and [5.23].

We recommend the use of first order finite elements and Euler implicit algorithm for state change problems.

5.3. A temperature-enthalpy formulation

In fact, the methods presented previously do not allow us to deal with the isothermal phase change (Figure 5.4). Indeed, in this case, T_M being the temperature of transformation, $H(T_M)$ and $\frac{dH}{dT}(T_M)$ are not defined. On the other hand, each enthalpy value has a corresponding unique temperature value. Let us denote by $T = \Theta(H)$ the function giving temperature from enthalpy. Therefore, the problem needs to be reformulated using enthalpy instead of temperature.

With that aim, Droux ([DES 87, DRO 90]) proposes an explicit finite element procedure considering enthalpy as a nodal unknown:

$$\frac{H - H_0}{\Delta t} - \text{div}(\vec{\lambda}(T_0) \cdot \vec{\text{grad}}(T_0)) = 0 \quad [5.24]$$

where H_0 and H represent enthalpy at time t and $t + \Delta t$ respectively and $T_0 = \Theta(H_0)$.

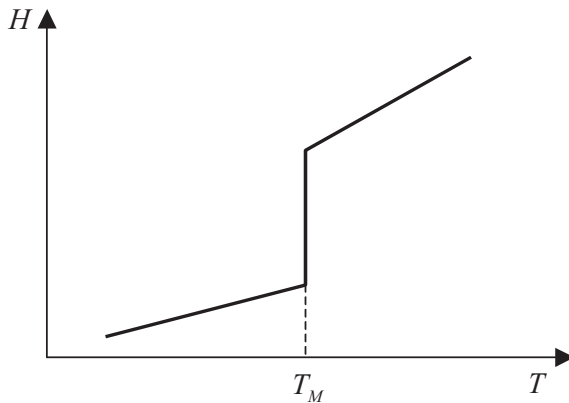


Figure 5.4. *Isothermal phase change*

The application of this explicit approach for modeling casting or welding processes leads to the use of a very large number of time steps for calculation, which can be reduced using an implicit algorithm:

$$\frac{H - H_0}{\Delta t} - \text{div} \left(\vec{\lambda} \frac{d\Theta}{dH}(H) \cdot \overrightarrow{\text{grad}}(H) \right) = 0 \quad [5.25]$$

The temperature does not appear in this formulation. Therefore, the boundary conditions need to be defined from the enthalpy H .

5.3.1. Mathematical formulation

To avoid the difficulties mentioned above, we suggest the use of formulation where both temperature and enthalpy are unknown ([FEU 07]). The problem is to find the temperature field and the enthalpy field solution of:

$$\dot{H} - \text{div} (\vec{\lambda} \cdot \overrightarrow{\text{grad}}(T)) = 0 \quad \text{in } \Omega \quad [5.26]$$

$$T - \Theta(H) = 0 \quad \text{in } \Omega \quad [5.27]$$

with classical boundary conditions:

$$T = T_d \quad \text{on } \partial\Omega_T \quad [5.28]$$

$$T - \lambda \frac{\partial T}{\partial n} = q \quad \text{on } \partial\Omega_q \quad [5.29]$$

Temperature and enthalpy are expressed from a nodal approximation:

$$T = \langle N^e \rangle \cdot \{T^e\} \quad [5.30]$$

$$H = \langle N^e \rangle \cdot \{H^e\} \quad \text{which also gives } \dot{H} = \langle N^e \rangle \cdot \{\dot{H}^e\} \quad [5.31]$$

Insofar as the same approximation scheme is used for both temperature and enthalpy and due to the strong non-linearity of the enthalpy-temperature relationship, equation [5.27] cannot be solved in a strong sense. Therefore, to obtain the finite element formulation of the problem, both equations [5.26] and [5.27] must be solved in a weak sense. Thus, both equations are classically first multiplied by the shape functions and then integrated over the whole domain Ω . After integrating the first equation by parts, including boundary conditions, and substituting the nodal

approximations into the Galerkin weighted residual form of [5.26] and [5.27], we obtain a non-linear semi-discrete system of equations giving the residual vector at each instant:

$$\begin{cases} \{R_T\} = [K] \cdot \{T\} + [M] \cdot \{\dot{H}\} - \{F\} = \{0\} \\ \{R_H\} = [M] \cdot \{T\} - \{G\} = \{0\} \end{cases} \quad [5.31a]$$

$\{T\}$ and $\{H\}$ are the vectors containing the nodal values of temperature and enthalpy. Matrices $[K]$, $[M]$ and vectors $\{F\}$, $\{G\}$ are obtained from an assembly process of element quantities defined as follows:

$$\begin{aligned} [K^e] &= \int_{\Omega^e} \left\{ \overrightarrow{\text{grad}}^T(N^e) \right\} \lambda \left\langle \overrightarrow{\text{grad}}(N^e) \right\rangle dv \\ [M^e] &= \int_{\Omega^e} \{N^e\} \langle N^e \rangle dv \\ \{F^e\} &= \int_{\partial\Omega^e \cap \partial\Omega_q} \{N^e\} q ds \\ \{G^e\} &= \int_{\Omega^e} \{N^e\} \Theta(H) dv \end{aligned} \quad [5.32]$$

Matrix $[M]$ is usually arbitrarily lumped at the nodes, resulting in a diagonal matrix $[M^*]$. This can be achieved by applying a row-sum technique:

$$M_{ij}^* = \delta_{ij} \sum_k M_{ik} \quad [5.33]$$

The non-linear semi-discrete equation system [5.31a] therefore becomes:

$$\begin{cases} \{R_T\} = [K] \cdot \{T\} + [M^*] \cdot \{\dot{H}\} - \{F\} = \{0\} \\ \{R_H\} = [M^*] \cdot \{T\} - \{G\} = \{0\} \end{cases} \quad [5.33a]$$

System [5.33a] is solved step by step in time. Let us apply for example an implicit Euler backward algorithm on the enthalpy:

$$H = H_0 + \dot{H} \Delta t \quad [5.34]$$

where H_0 and H are the enthalpy at time t and $t + \Delta t$ and \dot{H} , the enthalpy rate at time $t + \Delta t$. At each time step $t + \Delta t$, the nodal unknown values $\{T\}$ and $\{\dot{H}\}$ can be obtained using an iterative procedure like the Newton-Raphson method. Therefore, at each iteration it , we have to solve the following linear system of equations:

$$\begin{bmatrix} [K] & [M^*] \\ [M^*] & -\Delta t [Q] \end{bmatrix}^{(it)} \cdot \begin{Bmatrix} \{\Delta T\} \\ \{\Delta \dot{H}\} \end{Bmatrix}^{(it)} = - \begin{Bmatrix} \{R_T\} \\ \{R_H\} \end{Bmatrix}^{(it)} \quad [5.35]$$

where $[Q]$ results from the assembly of element matrices $[Q^e]$ defined by:

$$[Q^e] = \int_{\Omega^e} \frac{\partial \Theta}{\partial H} (\{N^e\} \langle N^e \rangle) dv \quad [5.36]$$

The second equation of [5.35] gives:

$$\{\Delta T\}^{(it)} = [M^*]^{-1} \cdot \left(\Delta t [Q]^{(it)} \cdot \{\Delta \dot{H}\}^{(it)} - \{R_H\}^{(it)} \right) \quad [5.37]$$

Because matrix $[M^*]$ is lumped, equation [5.37] gives $\{\Delta T\}^{(it)}$ explicitly as a function of $\{\Delta \dot{H}\}^{(it)}$. Taking account of equation [5.37], the first equation of [5.35] now gives $\{\Delta \dot{H}\}^{(it)}$ as the solution of the following symmetric linear system of equations:

$$\begin{aligned} & \left(\Delta t [K]^{(it)} \cdot [M^*]^{-1} \cdot [Q]^{(it)} + [M^*] \right) \cdot \{\Delta \dot{H}\}^{(it)} \\ & = [K]^{(it)} \cdot [M^*]^{-1} \cdot \{R_H\}^{(it)} - \{R_T\}^{(it)} \end{aligned} \quad [5.38]$$

Once $\{\Delta \dot{H}\}^{(it)}$ has been determined, equation [5.37] gives $\{\Delta T\}^{(it)}$ and the solution is updated by setting:

$$\begin{Bmatrix} \{T\} \\ \{\dot{H}\} \end{Bmatrix}^{(it+1)} = \begin{Bmatrix} \{T\} \\ \{\dot{H}\} \end{Bmatrix}^{(it)} + \begin{Bmatrix} \{\Delta T\} \\ \{\Delta \dot{H}\} \end{Bmatrix}^{(it)} \quad [5.39]$$

The iterative process is continued until convergence.

One drawback of the methods using enthalpy as the unknown comes from the fact that enthalpy is not continuous at the boundary between two materials. Therefore, to perform calculations on multi-material models, which is necessary for example for

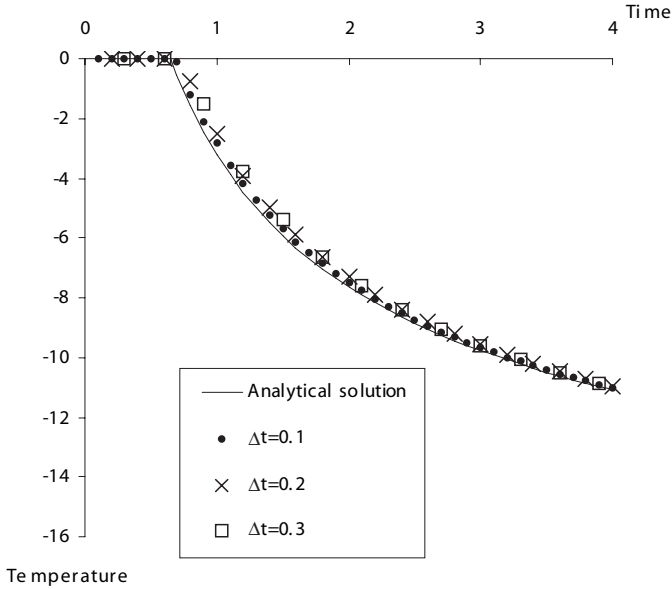


Figure 5.5. Temperature variation at $x = 1$ for different time steps (40 elements, from [FEU 07])

casting applications, it is essential to include thermal contact elements ensuring the continuity of temperature and an enthalpy jump at the interface between two different materials.

5.3.2. Example

Figures 5.5 and 5.6 give the results of a non-dimensional example similar to that proposed by Morgan *et al.* ([MOR 78]) and consisting of a unidirectional isothermal phase change in a semi infinite body. The specific heat (C), the mass density (ρ) and the thermal conductivity (λ) are set to 1. An isothermal phase change takes place at 0°C and the latent heat of transformation is assumed to be equal to 20. The length of the mesh, assumed to be equal to 4, has been chosen big enough to represent a semi-infinite body for the first seconds of simulation. Three meshes have been considered, consisting of a subdivision into 40, 20 and 8 8-noded hexahedral elements. The initial temperature and enthalpy at all the nodes are assumed to be equal to the temperature of transformation $T = 0$ and $H = 20$ respectively. For $t > 0$, the temperature at $x = 0$ is set equal to -20 . To ensure a correct convergence, $\frac{\partial \theta}{\partial H}$ should not be less than or equal to zero. It is therefore assumed to be equal to $\epsilon = 10^{-4}$ for $0 \leq H \leq 20$.

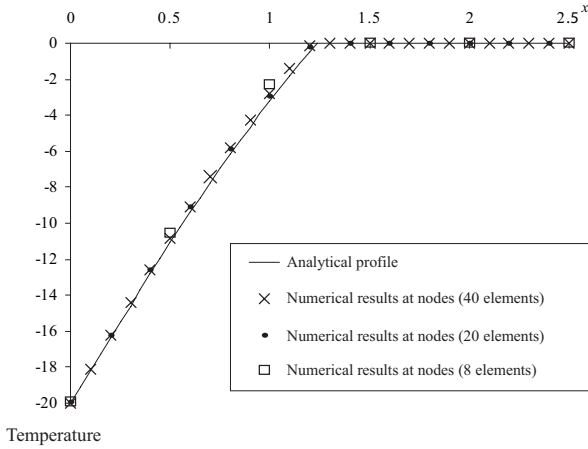


Figure 5.6. Temperature profile at $t = 1$ s for different meshes ($\Delta t = 0.1$ s, from [FEU 07])

The results obtained with the method presented above are compared with analytical results given in [CAR 59]. Figure 5.5 shows the influence of the time step on the computed temperature at $x = 1$ and Figure 5.6 demonstrates the influence of the mesh on the temperature profile at time $t = 1$. The very good agreement between the analytical solution and the numerical results obtained with different time steps and mesh sizes shows the efficiency of the method.

This page intentionally left blank

Chapter 6

Transport Phenomena

6.1. Highlighting instabilities

In this chapter, either a liquid medium or a mobile solid medium will be considered. The velocity field in the material is assumed to be known. The case of a mobile solid, which is a new extension of Part 1 of this book, is frequent enough in practice. For instance, welding, cutting or surface thermal treatment processes may involve a heat source moving at a given velocity (laser beam, electron beam, etc.), or otherwise the whole solid is moving across heating and cooling areas (line thermal treatment, continuous casting, etc.). In all cases, the thermal problem will be addressed in a framework of reference related to the stresses (heat source, etc.). Therefore, the solid is moving.

First, we will highlight the term to add to the thermal balance in order to take account for this movement. Then we will demonstrate by means of a simple 1D case that the traditional solution techniques produce spatial instabilities in the solution. In the last sections, a method commonly used in computer codes to solve these so-called convection-diffusion problems will be described.

6.1.1. Thermal balance

When the medium is moving, the temperature field sought, $T(\vec{x}, t)$, is written $T(\vec{x}(t), t)$. Indeed, the position \vec{x} of each material point is here a function of time. It follows that the temperature change over time is written as:

$$\dot{T} = \frac{\partial T}{\partial t} + \overrightarrow{\text{grad}}(T) \cdot \frac{d\vec{x}}{dt} = \frac{\partial T}{\partial t} + \vec{v} \cdot \overrightarrow{\text{grad}}(T)$$

In this equation, \vec{v} represents the velocity of the points of the solid. Of course, this velocity is *a priori* a function of the position of point \vec{x} and time t . From this result, it is possible to gather the thermal balance which must be satisfied at any point of the solid Ω :

$$\operatorname{div}(\underline{\lambda} \cdot \overrightarrow{\operatorname{grad}}(T)) + Q - \rho C \vec{v} \cdot \overrightarrow{\operatorname{grad}}(T) = \rho C \frac{\partial T}{\partial t} \quad \text{in } \Omega$$

In this chapter, our study will be restricted to the calculation of the steady state. In the stress-related frame of reference (heat or cooling source), the steady state corresponds to the condition $\frac{\partial T}{\partial t} = 0$. This finally produces the following equation:

$$\operatorname{div}(\underline{\lambda} \cdot \overrightarrow{\operatorname{grad}}(T)) + Q - \rho C \vec{v} \cdot \overrightarrow{\operatorname{grad}}(T) = 0 \quad \text{in } \Omega \quad [6.1]$$

In this type of problem, the boundary conditions are similar to those of a steady state, where flux q is prescribed on a boundary $\partial\Omega_q$ of the solid, and temperature T_d is known on the other portion $\partial\Omega_T$.

The finite element method applied to this equation and these boundary conditions obviously leads to the calculation of an element residual $\{R^e\}$. As in the transient state case, this residual will be linear according to temperature if:

- the thermal conductivity $\underline{\lambda}$ and material product ρC are independent of temperature;
- the volumetric heat source Q varies linearly according to the temperature: $Q = a_Q T + b_Q$;
- the surface flux density q varies linearly according to the temperature: $q = a_q T + b_q$.

In this case, the element residual is written $\{R^e\} = [K^e] \cdot \{T^e\} - \{F^e\}$, where the load vector $\{F^e\}$ and conductance matrix $[K^e]$ are written as:

$$K_{ij}^e = \begin{cases} \int_{\Omega^e} \overrightarrow{\operatorname{grad}}^T(N_i^e) \cdot \underline{\lambda} \cdot \overrightarrow{\operatorname{grad}}(N_j^e) dv - \int_{\Omega^e} N_i^e a_Q N_j^e dv \\ - \int_{\partial\Omega_q \cap \Omega^e} N_i^e a_q N_j^e ds - \int_{\Omega^e} N_i^e \rho C \vec{v} \cdot \overrightarrow{\operatorname{grad}}(N_j^e) dv \end{cases} \quad [6.2]$$

$$F_i^e = \int_{\Omega^e} N_i^e b_Q dv + \int_{\partial\Omega_q \cap \Omega^e} N_i^e b_q ds \quad [6.3]$$

In these expressions, we can see that the element matrix $[K^e]$ is not symmetric. The non-symmetric term is the transport term:

$$T_{ij}^e = \int_{\Omega^e} N_i^e \rho C \vec{v} \cdot \overrightarrow{\text{grad}}(N_j^e) dv$$

In a non-linear case, the tangent matrix involved would be non-symmetric too, because of the presence of the transport term. Moreover, this term cannot be neglected, as it is the only term taking into account the movement of the solid or fluid.

6.1.2. Treating a simple case

Applying the classic finite element method to the diffusion-convection problem produces instabilities. In order to highlight them, consider a fluid with physical properties, λ , ρ and C , running at constant speed, v , in a pipe. The problem is assumed to be 1D, the flowing direction is written x , and the temperature field $T(x)$ corresponding to the steady state is calculated. A temperature $T = 0$ is prescribed at the pipe entry ($x = 0$), and $T = 1$ at the exit ($x = 1$). The temperature field $T(x)$ is therefore the solution of the following boundary value problem:

$$\begin{cases} \rho C v \frac{dT}{dx} - \lambda \frac{d^2 T}{dx^2} = 0 & \text{for } 0 \leq x \leq 1 \\ T(0) = 0 \\ T(1) = 1 \end{cases}$$

This problem has a simple analytical solution:

$$T(x) = \frac{1 - e^{\alpha x}}{1 - e^{\alpha}} \quad \text{with} \quad \alpha = \frac{\rho C v}{\lambda} \quad [6.4]$$

In order to obtain a finite element solution, the weighted residual method is applied, at first by introducing weighting functions ψ , then a partial integration involving a temperature second derivative so as to end up with the weak formulation of the continuous problem:

$$\begin{cases} \text{Find } T(x) \text{ such that } T(0) = 0, T(1) = 1, \text{ and for any } \psi: \\ \lambda \int_0^1 \frac{d\psi}{dx} \frac{dT}{dx} dx + \rho C v \int_0^1 \psi(x) \frac{dT}{dx} dx = 0 \end{cases}$$

Consider a mesh obtained by the regular division of the segment $[0, 1]$ into n linear interpolation elements, of length $h = 1/n$. On each element $\Omega^e = [x_e, x_{e+1}]$ (with $e = 1, \dots, n$) two linear interpolation functions, $N_1^e = (x_{e+1} - x)/h$ (corresponding to the node with x-axis x_e) and $N_2^e = (x - x_e)/h$ (corresponding to the node with x-axis x_{e+1}) are defined. The element quantity calculation is then relatively simple. We obtain:

$$\begin{cases} K_{ij}^e = \lambda \int_{\Omega^e} \frac{dN_i^e}{dx} \frac{dN_j^e}{dx} dx + \rho C v \int_{\Omega^e} N_i^e(x) \frac{dN_j^e}{dx} dx \\ F_i^e = 0 \end{cases}$$

By analytically calculating the integrals involved, the following element quantities are obtained:

$$\begin{cases} [K^e] = \frac{\lambda}{h} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + \frac{\rho C v}{2} \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \\ \{F^e\} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \end{cases}$$

Likewise, the assembly is relatively simple in this case and produces the following matrices:

$$[K] = \frac{\lambda}{h} \begin{bmatrix} 1 & -1 & 0 & & \\ -1 & 2 & -1 & & \\ & & \ddots & & \\ & & & -1 & 2 & -1 \\ & & & 0 & -1 & 1 \end{bmatrix} + \frac{\rho C v}{2} \begin{bmatrix} -1 & 1 & 0 & & \\ -1 & 0 & 1 & & \\ & & \ddots & & \\ & & & -1 & 0 & 1 \\ & & & 0 & -1 & 1 \end{bmatrix}$$

and:

$$\{F\} = \begin{Bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{Bmatrix}$$

The generic equation at a node i such that $2 \leq i \leq n$ is written as:

$$\forall i = 2, \dots, n, \quad \frac{\lambda}{h} (-T_{i-1} + 2T_i - T_{i+1}) + \frac{\rho C v}{2} (-T_{i-1} + T_{i+1}) = 0 \quad [6.5]$$

This equation clearly reveals a Peclet number P related to an element:

$$P = \frac{\rho C v h}{\lambda} \quad [6.6]$$

Indeed, the solution of this generic equation can be written as:

$$T_{i+1} - T_i = A(T_i - T_{i-1}) = \dots = A^{i-1}(T_2 - T_1) \quad \text{with} \quad A = \frac{1 + \frac{P}{2}}{1 - \frac{P}{2}}$$

Coefficient A defined by this equation tends to infinity for a Peclet number $P = 2$, and changes signs on either side of this value. Therefore, physically realistic solutions will be obtained only for values of A positive or equal to zero, i.e. for Peclet numbers $P \leq 2$. For Peclet numbers greater than 2, the solution will produce spatial oscillations.

By consecutively applying the conditions $T_1 = 0$ and $T_{n+1} = 1$ to the generic solution, we obtain:

$$T_i = (A^{i-2} + A^{i-3} + \dots + A + 1)T_2 = \frac{1 - A^{i-1}}{1 - A}T_2$$

$$T_{n+1} = \frac{1 - A^n}{1 - A}T_2 = 1 \implies T_2 = \frac{1 - A}{1 - A^n}$$

The finite element solution can finally be written in a general way according to Peclet number P :

$$\forall i = 1, \dots, n+1, \quad T_i = \frac{1 - A^{i-1}}{1 - A^n} \quad \text{with} \quad A = \frac{1 + \frac{P}{2}}{1 - \frac{P}{2}} \quad [6.7]$$

Figure 6.1 illustrates the influence of Peclet number P on the solution of this example. To do so, a value $\alpha = \frac{\rho C v}{\lambda} = 30$ was used and the variation of P was obtained by modifying the discretization parameter h by the consecutive use of 30, 15, then 10 elements.

We can see that for values $P < 2$, the finite element approximate solution is relatively close to the analytical solution. On the contrary, when Peclet number P becomes greater than 2, the approximate solution produces spatial oscillations, due to the sign change of the coefficient A . If $P = 2$, coefficient A tends to infinity. In this case, a complete solution produces temperatures equal to zero at all the mesh nodes, except for node number $n+1$, where the condition $T_{n+1} = 1$ is satisfied.

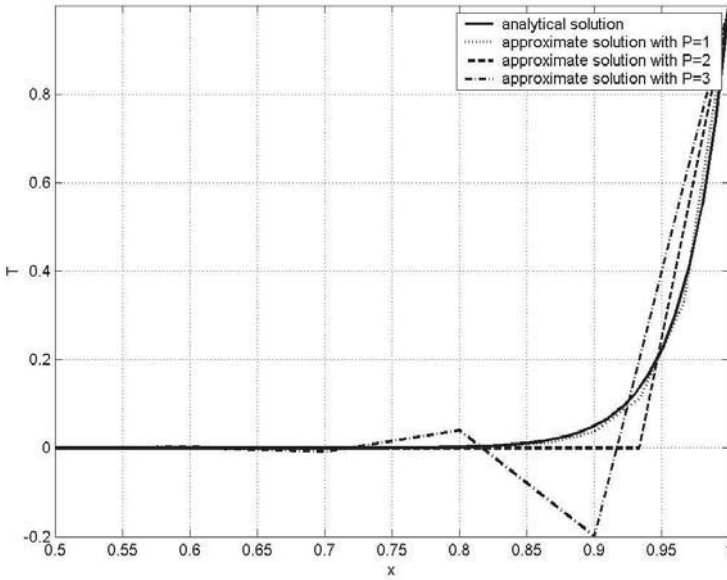


Figure 6.1. Temperature profiles obtained with $P = 1$, $P = 2$ and $P = 3$

6.2. Resolution techniques

With the classic finite element resolution, the solution obtained is only valid for low Peclet numbers P . Now, this number depends on the material physical properties (λ , ρ and C), the solid or fluid velocity v , and the length of the element h in its movement direction (equation [6.6]). For a given material and mesh, the classic finite element method is applicable only if the solid or fluid velocity v is low enough. If not, finer elements have to be used to reduce h , hence Peclet number P .

Specific resolution techniques for this type of problem have been implemented. In the case of finite differences, it is the *upwind* technique, which is described in section 6.2.1. This technique numerically solves a diffusion-convection problem with a high Peclet number. To adapt it to the finite element method, it is possible to use a Petrov-Galerkin variational formulation. This method is termed the Streamline-Upwind-Petrov-Galerkin (SUPG) method. This formulation is described in section 6.2.2.

6.2.1. Upwind technique

Numerical mathematicians using the finite difference method are very familiar with the spatial oscillation phenomenon in diffusion-convection. In fact, in the example above, the equations obtained by means of this method with centered difference schemes and a similar discretization are:

$$\begin{cases} T_1 = 0 \\ \forall i = 2, \dots, n, \quad -\frac{\lambda}{h^2}(T_{i+1} - 2T_i + T_{i-1}) + \frac{\rho C v}{2h}(T_{i+1} - T_{i-1}) = 0 \\ T_{n+1} = 1 \end{cases}$$

These equations are strictly identical to those obtained by the finite element method and therefore produce the same spatial oscillation problems for high Peclet numbers P . In order to rectify this, it is possible to write an off-centered difference scheme that estimates the first and second derivatives involved in the physical model:

$$\begin{aligned} \left. \frac{dT}{dx} \right|_{x=x_i} &= \frac{T_{i+1} - T_i}{h} \\ \left. \frac{d^2T}{dx^2} \right|_{x=x_i} &= \frac{1}{h} \left(\left. \frac{dT}{dx} \right|_{x=x_{i+1}} - \left. \frac{dT}{dx} \right|_{x=x_i} \right) = \frac{T_{i+2} - 2T_{i+1} + T_i}{h^2} \end{aligned}$$

Using this off-centered scheme on nodes $i = 1, \dots, n - 1$ produces the following equations:

$$\begin{cases} T_1 = 0 \\ \forall i = 1, \dots, n - 1, \quad -\frac{\lambda}{h^2}(T_{i+2} - 2T_{i+1} + T_i) + \frac{\rho C v}{h}(T_{i+1} - T_i) = 0 \\ T_{n+1} = 1 \end{cases}$$

The solution is written as:

$$\forall i = 1, \dots, n + 1, \quad T_i = \frac{1 - A^{i-1}}{1 - A^n} \quad \text{with} \quad A = 1 + P = 1 + \frac{\rho C v h}{\lambda} \quad [6.8]$$

Figure 6.2 shows the solutions obtained with this method for Peclet numbers P equal to 1, 2 or 3 obtained as before (Figure 6.1). It is possible to verify that the solution no longer oscillates for the high values of P .

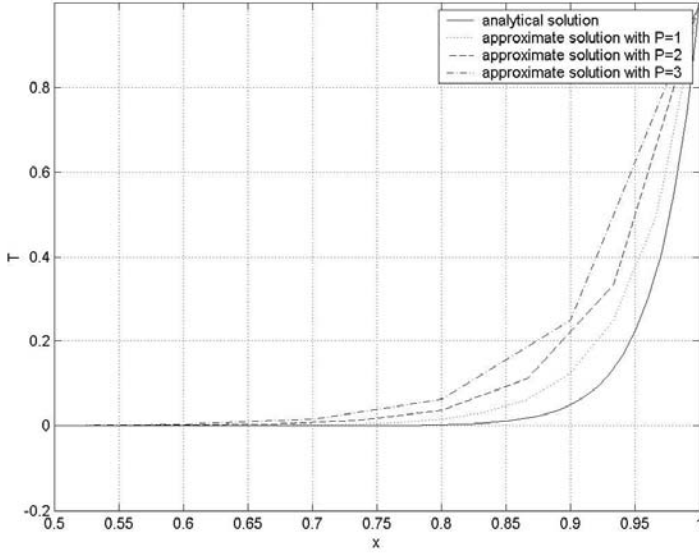


Figure 6.2. Temperature profiles obtained with an off-centered scheme

In order to compare the centered scheme (equation [6.7]) with the off-centered scheme (equation [6.8]), we analyze the generic equations E_i^c and E_i^d derived in both cases. We obtain:

$$E_i^c = \frac{\lambda}{h}(-T_{i-1} + 2T_i - T_{i+1}) + \frac{1}{2}\rho C v(-T_{i-1} + T_{i+1}) = 0$$

$$E_i^d = \frac{\lambda}{h}(-T_{i-1} + 2T_i - T_{i+1}) + \rho C v(-T_{i-1} + T_i) = 0$$

The first terms of E_i^c and E_i^d correspond to the discretizations of the conduction term, and the second terms to those of the convection term. If these expressions are analyzed in-depth, we can note that:

$$E_i^d = E_i^c + \frac{1}{2}\rho C v(-T_{i-1} + 2T_i - T_{i+1})$$

The off-centered scheme therefore corresponds to a centered scheme in which more weight would have been given to conduction. Indeed, we obtain E_i^d again by

replacing inside E_i^c the thermal conduction λ with an equivalent conduction λ' such that:

$$\lambda' = \lambda \left(1 + \frac{\rho C v h}{2\lambda} \right) = \lambda \left(1 + \frac{P}{2} \right)$$

This result is used to establish the finite element formulations corresponding to the off-centered scheme.

6.2.2. SUPG method

In order to apply the upwind technique to finite elements, it is possible to use a Petrov-Galerkin method, which uses discontinuous weighting functions on the element boundaries. In this section, the Petrov-Galerkin method will be illustrated on the simple 1D example of the previous sections. To do so, the weighted residual method is applied to the problem to solve to obtain:

$$\begin{cases} \rho C v \int_0^1 \psi(x) \frac{dT}{dx} dx - \lambda \int_0^1 \psi(x) \frac{d^2 T}{dx^2} dx = 0 \\ T(0) = 0 \\ T(1) = 1 \end{cases}$$

Make each weighting function ψ as the sum of a continuous function β on the segment $[0, 1]$ and a discontinuous function γ on the element boundaries:

$$\forall x \in [0, 1], \quad \psi(x) = \beta(x) + \gamma(x)$$

We obtain:

$$\rho C v \int_0^1 (\beta(x) + \gamma(x)) \frac{dT}{dx} dx - \lambda \int_0^1 \beta(x) \frac{d^2 T}{dx^2} dx - \lambda \int_0^1 \gamma(x) \frac{d^2 T}{dx^2} dx = 0$$

Apply an integration by parts on the second term of this equation. We obtain:

$$\int_0^1 \beta(x) \frac{d^2 T}{dx^2} dx = \left[\beta(x) \frac{dT}{dx} \right]_0^1 - \int_0^1 \frac{d\beta}{dx} \frac{dT}{dx} dx$$

Finally, by choosing functions β equal to zero in $x = 0$ and $x = 1$ (i.e. at the places where essential boundary conditions are applied), the global problem to solve, derived

from the weighted residual method and an integration by parts of the continuous terms of the weighting functions, becomes:

$$\begin{cases} \text{Find } T(x) \text{ such that } T(0) = 0, T(1) = 1, \text{ and for any } \beta(x) \text{ and } \gamma(x): \\ \lambda \int_0^1 \left(\frac{d\beta}{dx} \frac{dT}{dx} - \gamma(x) \frac{d^2T}{dx^2} \right) dx + \rho C v \int_0^1 (\beta(x) + \gamma(x)) \frac{dT}{dx} dx = 0 \end{cases} \quad [6.9]$$

The finite element discretization is based upon the regular division of the segment $[0, 1]$ into n linear interpolation elements, of length $h = 1/n$. On each element $\Omega^e = [x_e, x_{e+1}]$, with $e = 1, \dots, n$, two interpolation functions $N_1^e = (x_{e+1} - x)/h$ (corresponding to the node with x-axis x_e) and $N_2^e = (x - x_e)/h$ (corresponding to the node with x-axis x_{e+1}) are defined in a traditional way. However, these interpolation functions will only be used for the continuous fields, i.e. functions T and β . The functions γ inside each element Ω^e will be defined from two functions $H_1^e(x)$ and $H_2^e(x)$ which will be specified later on. Therefore, we obtain:

$$\forall x \in \Omega^e, \quad \begin{cases} T(x) = N_1^e(x)T_1^e + N_2^e(x)T_2^e \\ \psi(x) = (N_1^e(x) + H_1^e(x))\psi_1^e + (N_2^e(x) + H_2^e(x))\psi_2^e \end{cases}$$

Discretized equation [6.9] then leads us to solve the system $[K] \cdot \{T\} = \{F\}$, where the matrix $[K]$ and load vector $\{F\}$ are obtained as usual by assembling element quantities $[K^e]$ and $\{F^e\}$. By noting that, in this example, the interpolation functions $N_i^e(x)$ are linear and therefore their second derivative to x is equal to zero, we obtain:

$$\begin{cases} K_{ij}^e = \lambda \int_{\Omega^e} \frac{dN_i^e}{dx} \frac{dN_j^e}{dx} dx + \rho C v \int_{\Omega^e} \left(N_i^e(x) \frac{dN_j^e}{dx} + H_i^e(x) \frac{dN_j^e}{dx} \right) dx \\ F_i^e = 0 \end{cases}$$

In the expression of the terms K_{ij}^e the second integral involves the functions H_i^e . Hughes and Brooks [HUG 82] suggest we take as interpolation functions:

$$H_i^e(x) = \frac{1}{2} \xi h \frac{dN_i^e}{dx}$$

where ξ is an adjustable parameter and h the element length. This option uses the result obtained at the end of the previous section and gives more weight to thermal conduction. In fact, the components of the matrix $[K^e]$ are now written as:

$$K_{ij}^e = \lambda \left(1 + \xi \frac{P}{2} \right) \int_{\Omega^e} \frac{dN_i^e}{dx} \frac{dN_j^e}{dx} dx + \rho C v \int_{\Omega^e} N_i^e(x) \frac{dN_j^e}{dx} dx$$

In this equation, Peclet number P defined by equation [6.6] appears. In addition, note that this expression is similar to that derived from the classic approach, providing that the thermal conductivity λ is replaced with the following equivalent conductivity λ' :

$$\lambda' = \lambda \left(1 + \xi \frac{P}{2} \right)$$

If this equivalent thermal conductivity λ' is used, it is again possible to write generic equation [6.5] of the problem to solve using the finite element method. This generic equation becomes:

$$\forall i = 2, \dots, n, \quad \frac{\lambda'}{h} (-T_{i-1} + 2T_i - T_{i+1}) + \frac{\rho C v}{2} (-T_{i-1} + T_{i+1}) = 0$$

If the same method is used to solve the global system (with the boundary conditions $T_0 = 0$ and $T_{n+1} = 1$), the general solution of equation [6.7] is expressed as follows, according to Peclet number P and the coefficient ξ :

$$\forall i = 1, \dots, n+1, \quad T_i = \frac{1 - A^{i-1}}{1 - A^n} \quad \text{with} \quad A = \frac{1 + \frac{P}{2}(\xi + 1)}{1 + \frac{P}{2}(\xi - 1)} \quad [6.10]$$

Next, it is necessary to adjust this parameter ξ so as to avoid oscillations (coefficient A negative), and obtain the most accurate possible solution.

In the simple case addressed here, equation [6.4] yields the analytical solution of the problem. In particular, at the mesh nodes with x-axes $x_i = (i-1)h$, this analytical solution can be expressed as follows (P is the problem Peclet number):

$$\forall i = 1, \dots, n+1, \quad T_i = \frac{1 - e^{(i-1)P}}{1 - e^{nP}} \quad [6.11]$$

The comparison of equations [6.11] and [6.10] leads us to choose a coefficient $A = e^P$. Indeed, in this case, the finite element solution will be accurate at the mesh nodes and the coefficient A will always be positive. Then the following optimal value will be obtained for ξ :

$$\xi_{\text{opt}} = \frac{e^P + 1}{e^P - 1} - \frac{2}{P}$$

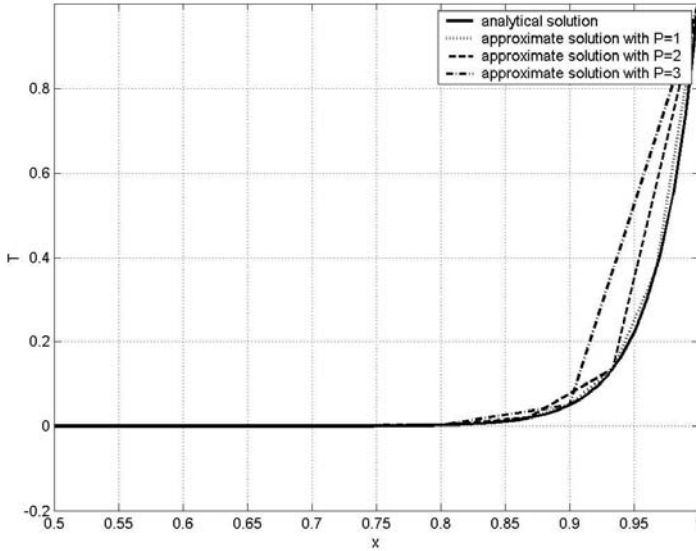


Figure 6.3. Temperature profile obtained with a Petrov-Galerkin formulation and an optimal coefficient ξ

Figure 6.3 gives the solution obtained analytically with the finite element method, for $P = 1$, $P = 2$ and $P = 3$. The solution does not oscillate anymore and coincides with the analytical solution at the mesh nodes. The discrepancies between both approaches are only due to the linear nodal approximation used.

The optimal value of the coefficient ξ obtained here is only valid in the simple case under consideration. Now it is necessary to address the general case, to check whether a similar approach can be appropriate.

6.2.3. 2D and 3D Petrov-Galerkin formulation

The Petrov-Galerkin method illustrated in the example above can be used in the general case of a 2D or 3D diffusion-convection problem, as well as with any interpolation functions N_i^e . Here again, the functions used to define the discontinuous part γ of the weighting functions ψ will be termed H_i^e and N_i^e those used for their continuous part β and for the temperature field T . In the linear case, the system to

solve is of the type $[K] \cdot \{T\} = \{F\}$ where the matrix $[K]$ and vector $\{F\}$ result from the assembly of the element quantities $[K^e]$ and $\{F^e\}$ whose general terms are:

$$\begin{aligned} K_{ij}^e &= \int_{\Omega^e} \overrightarrow{\text{grad}}^T(N_i^e) \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(N_j^e) dv + \int_{\Omega^e} \rho C(N_i^e + H_i^e) \overrightarrow{v} \cdot \overrightarrow{\text{grad}}(N_j^e) dv \\ &\quad - \int_{\Omega^e} H_i^e \text{div}(\underline{\lambda} \cdot \overrightarrow{\text{grad}}(N_j^e)) dv \\ F_i^e &= \int_{\Omega^e} (N_i^e + H_i^e) Q dv + \int_{\partial\Omega_q \cap \partial\Omega^e} N_i^e q ds \end{aligned}$$

Obtaining these element quantities results from the application of the weighted residual method and the divergence theorem to get a weak formulation of the continuous problem. However, the divergence theorem can only be applied here on the terms involving the continuous part β of the weighting functions. In matrix $[K^e]$ there are terms including second derivatives of the shape functions N_i^e . However, these terms are often neglected. They are equal to zero for linear interpolation elements and a constant thermal conductivity.

Thus, the Petrov-Galerkin method defines a working context for the finite element solution of a diffusion-convection problem. The quality of approximate solutions will be directly related to the choice of functions H_i^e . As in the 1D case above, Hughes and Brooks [HUG 82, BRO 82] suggest the following functions:

$$H_i^e = \tau \text{div}(N_i^e \overrightarrow{v})$$

In this equation, τ is an adjustable parameter. Now in the 1D case, we had $H_i^e(x) = \frac{1}{2} \xi h \frac{dN_i^e}{dx}$, i.e. $\tau = \frac{1}{2} \xi \frac{h}{v}$, where v is the velocity in the direction x . It is possible to extend this result to the 2D and 3D cases as follows:

$$\tau = \frac{\xi h_\xi \overrightarrow{v} \cdot \overrightarrow{u}_\xi + \eta h_\eta \overrightarrow{v} \cdot \overrightarrow{u}_\eta}{2 \overrightarrow{v} \cdot \overrightarrow{v}} \quad (2D \text{ case})$$

$$\tau = \frac{\xi h_\xi \overrightarrow{v} \cdot \overrightarrow{u}_\xi + \eta h_\eta \overrightarrow{v} \cdot \overrightarrow{u}_\eta + \zeta h_\zeta \overrightarrow{v} \cdot \overrightarrow{u}_\zeta}{2 \overrightarrow{v} \cdot \overrightarrow{v}} \quad (3D \text{ case})$$

In these equations, ξ , η and ζ are adjustable parameters, and h_ξ , h_η and h_ζ are lengths characteristic of the element in the three reference directions \overrightarrow{u}_ξ , \overrightarrow{u}_η and \overrightarrow{u}_ζ (Figure 6.4). Considering the results obtained in the previous example, it is possible

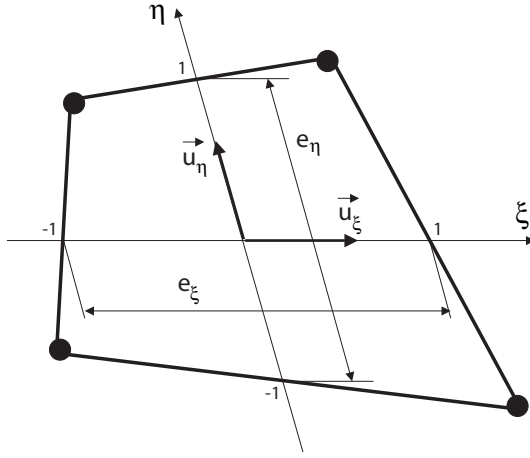


Figure 6.4. *Diagram of the 2D upwind technique principle*

for example to choose as adjustable parameters, in the case of an isotropic thermal conductivity λ :

$$\left\{ \begin{array}{l} \xi = \frac{e^{P_\xi} + 1}{e^{P_\xi} - 1} - \frac{2}{P_\xi} \\ \eta = \frac{e^{P_\eta} + 1}{e^{P_\eta} - 1} - \frac{2}{P_\eta} \\ \zeta = \frac{e^{P_\zeta} + 1}{e^{P_\zeta} - 1} - \frac{2}{P_\zeta} \end{array} \right. \quad \text{with} \quad \left\{ \begin{array}{l} P_\xi = \frac{\rho C h_\xi}{\lambda} \vec{v} \cdot \vec{u}_\xi \\ P_\eta = \frac{\rho C h_\eta}{\lambda} \vec{v} \cdot \vec{u}_\eta \\ P_\zeta = \frac{\rho C h_\zeta}{\lambda} \vec{v} \cdot \vec{u}_\zeta \end{array} \right.$$

In these expressions, the Peclet number related to each reference direction of the element under consideration appears. The method presented therefore makes it possible to extend to 2D and 3D the off-centered scheme introduced in the 1D example. The off-centering operation is carried out in the flow direction. The values of ξ , η and ζ thus defined correspond to optimal 1D values, for which the numerical and analytical solutions coincide at the mesh nodes. However, this is rarely the case when the flow is more complex, as in 2D or 3D.

PART 3

Coupled Phenomena

This page intentionally left blank

Introduction

Consider again the boundary value problem governing heat diffusion in a solid occupying a bounded domain Ω , with the boundary $\partial\Omega$, with initial conditions on T :

$$\begin{aligned} \rho \frac{dH}{dt} - \text{div} (\underline{\lambda} \cdot \overrightarrow{\text{grad}}(T)) - Q &= 0 \quad \text{in } \Omega \\ \left. \begin{aligned} (\underline{\lambda} \cdot \overrightarrow{\text{grad}}(T)) \cdot \overrightarrow{n} &= q \quad \text{on } \partial\Omega_q \\ T &= T_d \quad \text{on } \partial\Omega_T \end{aligned} \right\} \quad \text{with } \partial\Omega = \partial\Omega_q \cup \partial\Omega_T \end{aligned}$$

The finite element method applied to this problem produces an element residual $\{R^e(T)\}$ (equation [2.6] in Part 1 of this book for the steady state, equation [4.7] in Part 2 for the transient state), then performs the assembly to obtain a global residual $\{R(T)\}$ to cancel:

$$\{R(T)\} = \sum_{e=1}^m [A^e]^T \cdot \{R^e(T)\} = 0$$

As already seen in Chapter 5, this equation is solved by an iterative method. For instance, the Newton-Raphson method solves at each iteration it the following linear system, so as to obtain the nodal vector $\{T\}^{(it+1)}$ of the next iteration temperatures:

$$\{R\}^{(it)} + \left[\frac{\partial R}{\partial T} \right]^{(it)} \cdot (\{T\}^{(it+1)} - \{T\}^{(it)}) = 0$$

For instance, iterations are performed in order to obtain a norm of the residual vector $\{R\}^{(it)}$ less than a precision threshold set by the user. In this equation, the tangent matrix is obtained by assembling element tangent matrices.

The boundary value problem contains a large amount of coupling relationships with other physical phenomena. The material physical properties ρ , H and λ may thus depend on the temperature, but also the chemical composition or other state variables such as those describing the microstructure (phase proportions, grain size, etc.) or the pressure. The internal heat source Q involves other physical phenomena. They can be the Joule effect in conduction heating or induction heating applications, or plastic dissipation in forming applications. Here again, the problem involves new state variables, but this term can also express radiation exchanges in a participative medium like glass at certain temperatures. It then depends only on the temperature field in the solid, but will produce coupling phenomena on distances longer than those involved in the diffusion phenomena alone. The boundary conditions such as the prescribed thermal flux density q or the exchange coefficient h and the external temperature are also due to complementary physical phenomena. These phenomena may also involve other state variables (for example, those governing a fluid bath flow in the vicinity of the wall) or not, as in the case of thermal radiation in a chamber, but this will produce additional coupling phenomena on the temperature field at the surface of the chamber. Generally speaking, coupled problems can be arranged into three categories:

1) The first category includes problems involving similar state variables. Such is the case of coupled diffusion and radiation heat exchanges. The diffusion-related tangent matrix then couples all the nodes involved in these phenomena (for example, all the nodes at the surface of the chamber in the case of thermal radiation in a chamber). The result is a very significant increase of the non-zero terms on each line of the tangent matrix, the destruction of its band structure, and therefore a very high increase of the solution time. However, as already seen in the solution of non-linear problems, it is not necessary to use a real tangent matrix to solve this equation. It would be sufficient to introduce a partial radiation contribution into the tangent matrix, which does not destroy the good structure of the matrix related to diffusion exchanges. However, this method will inevitably require a larger number of iterations and, in some cases, will not allow any convergence.

2) The second category includes problems involving different state variables. However, additional state variables are governed by time ordinary differential equations so that they can be calculated locally. For instance, structural transformations are governed by equations depending on the (local) temperature, each phase (local) proportion and time. These equations are then solved at each finite element integration point and the dependent magnitudes are updated for the calculation of element quantities.

3) The third category includes problems involving different state variables whose changes are governed by boundary value problems. Boundary value problems may

be related to the same domain of study Ω as is generally the case for electrothermal coupling applied to conduction heating, or not, as in induction heating where modeling magnetic dynamic phenomena requires us to consider air as the domain of study.

For the problems belonging to the third category, two (or several) non-linear systems must be solved at each time step. If for instance the vector of nodal unknowns related to the coupling is termed $\{\alpha\}$, and $\{S\}$ is the residual obtained by applying the finite element method to the physical model governing these unknown changes, at each time step, the following equation has to be solved:

$$\begin{aligned} \{R(\{T\}, \{\alpha\})\} &= \{0\} \\ \{S(\{T\}, \{\alpha\})\} &= \{0\} \end{aligned}$$

Note that in this equation, vectors $\{T\}$ and $\{\alpha\}$ do not necessarily have the same number of components, as the physical phenomena involved are not necessarily related to the same domains of study and therefore their solution may be based upon different meshes. If these physical phenomena are weakly coupled, it is possible to find an approximate solution of the problem by replacing the first equation $\{\alpha\}$ with its value obtained at the previous time step, then solving this equation alone to determine $\{T\}$, and the second equation alone to determine $\{\alpha\}$. If the two problems are strongly coupled, then both equations have to be solved at the same time. To do so, an iterative procedure can be used.

If $\{T\}^{(it)}$ and $\{\alpha\}^{(it)}$ are the solutions obtained after it iterations, a better solution can be obtained by solving consecutively:

$$\{R(\{T\}^{(it+1)}, \{\alpha\}^{(it)})\} = \{0\}$$

to obtain $\{T\}^{(it+1)}$, then:

$$\{S(\{T\}^{(it+1)}, \{\alpha\}^{(it+1)})\} = \{0\}$$

to obtain $\{\alpha\}^{(it+1)}$, and so on and so forth until the norms of vectors $\{R\}$ and $\{S\}$ reach the accuracy thresholds required.

The major advantage of this method is that both physical problems can use different meshes. It is then necessary to have an algorithm for the transport of physical quantities from one mesh to the other. Moreover, this method does not require a specific formulation of the coupling as the problems are solved consecutively. It

makes it possible to use different numerical methods to solve each system. It is finally well-adapted to boundary condition coupling.

Of course it is possible to solve the problem globally and directly by means of a Newton-Raphson method. Starting with $\{T\}^{(it)}$ and $\{\alpha\}^{(it)}$, $\{T\}^{(it+1)}$ and $\{\alpha\}^{(it+1)}$ are obtained by the resolution of:

$$\begin{Bmatrix} \{R\}^{(it)} \\ \{S\}^{(it)} \end{Bmatrix} + \begin{bmatrix} \left[\frac{\partial R}{\partial T} \right]^{(it)} & \left[\frac{\partial R}{\partial \alpha} \right]^{(it)} \\ \left[\frac{\partial S}{\partial T} \right]^{(it)} & \left[\frac{\partial S}{\partial \alpha} \right]^{(it)} \end{bmatrix} \cdot \begin{Bmatrix} \{T\}^{(it+1)} - \{T\}^{(it)} \\ \{\alpha\}^{(it+1)} - \{\alpha\}^{(it)} \end{Bmatrix} = \begin{Bmatrix} \{0\} \\ \{0\} \end{Bmatrix}$$

This iterative process *a priori* converges faster than the previous one. However, this method is more complicated because it requires the development of specific elements and generally produces non-symmetric matrices. It will be restricted to particular cases for which the previous convergence method is not possible.

Chapter 7

Radiation Exchanges in a Chamber

7.1. Modeling radiative heat exchanges in a cavity

Heat radiation becomes an important mode of heat transfer as soon as high temperatures (higher than 300°C) are involved, as is the case for heat exchanges in a heat treatment furnace. Generally speaking, such exchanges are strongly non-linear and difficult to take into consideration in finite element modeling, because they couple the radiative surfaces. This results in an increase of the memory space necessary for calculation, and long calculation times.

The most commonly used methods to solve heat radiation problems assume isothermal zones, for the equations governing radiation exchanged heat fluxes are well-known and this assumption is particularly well-adapted to finite element modeling [SIE 81]. These methods can be considered to be specific cases of variational approximations in which special shape functions are used [KEA 88, LOB 95, BRE 90].

However, the coupled resolution of diffusion and radiation equations creates some difficulties [KEA 88, ENG 91, SOR 90]. One method is to solve consecutively, at each iteration, the systems describing radiation and diffusion exchanges. Radiation fluxes are calculated from the previous iteration temperatures and transferred into the finite element system to calculate new temperatures. The explicit character of this coupling requires the use of very tiny time increments for the simulation of transient phenomena and bans this method for steady solutions. However, the advantage of this method lies in the fact that it satisfies the discrete maximum principle [LOB 95], and therefore

avoids spatial oscillations of the solution which may occur in problems where the radiation phenomenon is paramount.

For more general cases, the problems related to diffusion and radiation exchanges are treated simultaneously [KEA 88, ENG 91, SOR 90, BOR 90, BER 01]. At each iteration, the solution of the non-linear problem involves a matrix including a contribution of the radiative phenomena. This contribution destroys the band structure of the matrix resulting from the diffusion phenomena and makes it non-symmetric.

In this section, the equations governing radiation exchanges are described first, then a coupling method often used in finite element computer codes is introduced.

7.1.1. *Posing the problem*

Consider a cavity whose surface consists of f plane surfaces (Figure 7.1) and seek an expression of the heat flux locally received by each surface under the action of the radiation inside the cavity. It can be noted that the approximation of the cavity surface by a set of plane surfaces results from the finite element mesh. Besides, the assumption of gray bodies is made. It is therefore assumed that the radiance of each plane surface is independent of the direction and the monochromatic emission factor is independent of the wavelength.

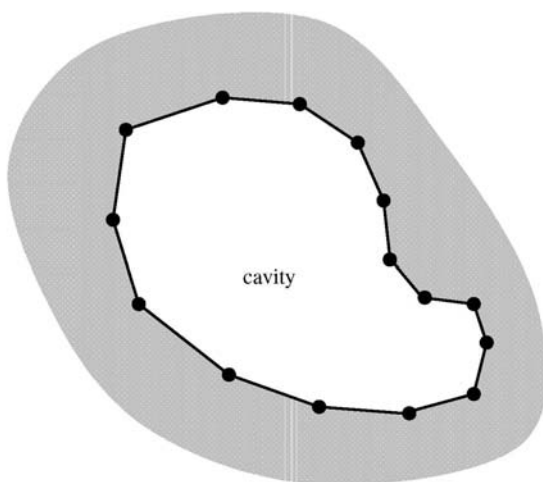


Figure 7.1. *Discretized cavity*

For each plane surface a , its surface is termed S_a , its emission coefficient (or absorption coefficient) ϵ_a , and its reflection coefficient ρ_a , ($\rho_a = 1 - \epsilon_a$). The assumption of so-called isothermal zones is made. It is therefore assumed that each plane surface has a uniform temperature written \bar{T}_a , expressed in Kelvin, and it is recalled that:

- the proper emittance of a gray plane surface a is given by:

$$M_a = \sigma \epsilon_a \bar{T}_a^4$$

where σ is a Stefan constant which is equal to $5.67 \cdot 10^{-8} \text{ W/m}^2/\text{K}^4$;

- the view factor F_{ab} from a plane surface a to another plane surface b is the energy fraction leaving the plane surface a reaching the plane surface b . The view factors F_{ab} satisfy the conditions:

$$S_a F_{ab} = S_b F_{ba} \quad \text{and for any } a, \quad \sum_{b=1}^f F_{ab} = 1 \quad [7.1]$$

For each plane surface a it is now necessary to define an apparent emittance or radiosity R_a such as the energy density escaping the plane surface:

$$R_a = \underbrace{\sigma \epsilon_a \bar{T}_a^4}_{\text{proper emission}} + \underbrace{(1 - \epsilon_a) H_a}_{\text{reflected portion of incident energy}} \quad [7.2]$$

The incident flux density on a , H_a , is obtained by figuring the following relationship:

$$\underbrace{S_a H_a}_{\substack{\text{energy} \\ \text{incident on } a}} = \sum_{b=1}^f \underbrace{F_{ab}}_{\substack{\text{fraction} \\ \text{arriving on } a}} \underbrace{S_b R_b}_{\substack{\text{energy} \\ \text{leaving } b}} = S_a \sum_{b=1}^f F_{ab} R_b \quad [7.3]$$

The net flux density received by the plane surface a , q_a is then expressed as:

$$q_a = \underbrace{\epsilon_a H_a}_{\substack{\text{absorbed part of} \\ \text{incident energy}}} - \underbrace{\sigma \epsilon_a \bar{T}_a^4}_{\text{proper emission}}$$

This expression, combined with equation [7.2], gives:

$$q_a = \frac{\epsilon_a}{1 - \epsilon_a} (R_a - \sigma \bar{T}_a^4) \quad [7.4]$$

The combination of equations [7.2] and [7.3] produces a linear system of f equations with f unknowns, making it possible to determine the plane surface radiosities. We obtain:

$$\sum_{b=1}^f (\delta_{ab} - (1 - \epsilon_a)F_{ab})R_b = \sigma\epsilon_a\bar{T}_a^4$$

In this equation, δ_{ab} represents the Kronecker symbol ($\delta_{ab} = 1$ if $a = b$ and 0 otherwise). The linear system reveals a non-symmetric matrix with $\delta_{ab} - (1 - \epsilon_a)F_{ab}$ as general term. In order to facilitate the numerical resolution of this system, it is made symmetric. To do so, it is assumed that the cavity surface consists of very gray plane surfaces ($\epsilon_a \neq 1$) numbered from 1 to g , and very black surfaces ($\epsilon_a = 1$) numbered from $g + 1$ to f . The radiosity of a black plane surface being equal to its proper emittance, while for any gray plane surface a the previous system is written:

$$\sum_{b=1}^g (\delta_{ab} - (1 - \epsilon_a)F_{ab})R_b = \sigma\epsilon_a\bar{T}_a^4 + (1 - \epsilon_a) \sum_{b=g+1}^f F_{ab}\sigma\bar{T}_b^4$$

If the two members of the equation are multiplied by $S_a/(1 - \epsilon_a)$, the following symmetric linear system of g equations for g unknowns is finally obtained:

$$\sum_{b=1}^g \underbrace{\left(\frac{S_a}{1 - \epsilon_a} \delta_{ab} - S_a F_{ab} \right)}_{G_{ab}} R_b = \sigma \frac{\epsilon_a S_a}{1 - \epsilon_a} \bar{T}_a^4 + \sum_{b=g+1}^f S_a F_{ab} \sigma \bar{T}_b^4 \quad [7.5]$$

The gray surface radiosities are determined by the solution of this symmetric linear system ($G_{ba} = G_{ab}$). If L_{ab} is the general term of the inverse matrix of the G_{ab} , the radiosity of a gray surface a can be expressed as:

$$R_a = \sum_{b=1}^g L_{ab} \left(\sigma \frac{\epsilon_b S_b}{1 - \epsilon_b} \bar{T}_b^4 + \sum_{c=g+1}^f S_b F_{bc} \sigma \bar{T}_c^4 \right) \quad [7.6]$$

It is now possible to note that the sum of the terms on the same line of the matrix formed by the G_{ab} is equal to:

$$\sum_{b=1}^g G_{ab} = \frac{S_a}{1 - \epsilon_a} - \sum_{b=1}^g S_a F_{ab} = \frac{\epsilon_a S_a}{1 - \epsilon_a} + \sum_{b=g+1}^f S_a F_{ab}$$

By considering this sum as the product of a matrix whose general term is G_{ab} by a vector including only 1, the equation above can be written as an equation system which, after inversion, gives the following relationship between the terms L_{ab} :

$$\sum_{b=1}^g L_{ab} \left(\frac{\epsilon_b S_b}{1 - \epsilon_b} + \sum_{c=g+1}^f S_b F_{bc} \right) = 1 \quad \text{for any gray plane surface } a.$$

Finally, by transferring this relationship and equation [7.6] into the expression of the net flux received q_a (equation [7.4]), we obtain for any gray plane surface a :

$$q_a = \frac{1}{S_a} \sum_{b=1, b \neq a}^f \sigma D_{ab} (\bar{T}_b^4 - \bar{T}_a^4) \quad [7.7]$$

In this equation, the terms D_{ab} are expressed as (a is a gray plane surface):

$$D_{ab} = \begin{cases} \frac{\epsilon_a S_a}{1 - \epsilon_a} L_{ab} \frac{\epsilon_b S_b}{1 - \epsilon_b} & \text{if } b \text{ is gray} \\ \frac{\epsilon_a S_a}{1 - \epsilon_a} \sum_{c=1}^g L_{ac} S_c F_{cb} & \text{if } b \text{ is black} \end{cases}$$

In the case of a black plane surface a , the expression obtained for the net flux received is the same. The terms D_{ab} are then expressed as:

$$D_{ab} = \begin{cases} S_a F_{ab} + \sum_{c=1}^g \sum_{d=1}^g S_a F_{ac} L_{cd} S_d F_{db} & \text{if } b \text{ is black} \\ \frac{\epsilon_b S_b}{1 - \epsilon_b} \sum_{c=1}^g L_{cb} S_a F_{ac} & \text{if } b \text{ is gray} \end{cases}$$

It can be noted that the matrix whose general term is D_{ab} is symmetric. The method above therefore allows for an analytical expression of the net flux received by each gray or black plane surface making up the cavity surface. There are several advantages to this method. First, it is applied regardless of the surface emissivities. Then, when the surface emissivities are independent of temperature, it only requires the inversion of a symmetric matrix whose order is equal to the number of gray plane surfaces. The calculation of this matrix and of the view factors is made in a pre-treatment phase. Finally, the expression of the net flux received by each plane surface makes it possible

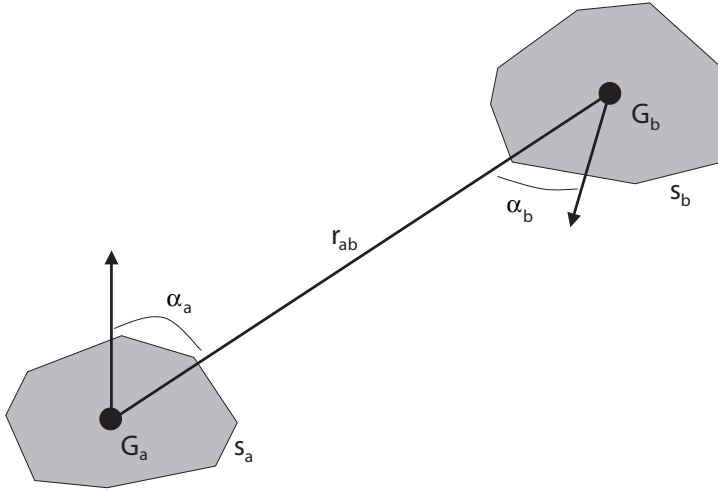


Figure 7.2. View factor between two plane surfaces

to include a contribution due to the heat radiation into the first member matrix of the problem coupling diffusion and radiation.

7.1.2. Calculation of view factors

The view factor between two plane surfaces is given by the following expression (Figure 7.2):

$$F_{ab} = \frac{1}{\pi S_a} \int_{S_a} \int_{S_b} \frac{\cos \alpha_a \cos \alpha_b}{r_{ab}^2} ds_a ds_b \quad [7.8]$$

Several calculation techniques may be used. In a 2D case, Hottel's crossed string method [SIE 81] uses equation [7.1] in a quadrangle (Figure 7.3). It is easy to obtain:

- for two adjacent plane surfaces (for example, 1 and 2 in Figure 7.3):

$$S_1 F_{12} = \frac{1}{2} (S_1 + S_2 - S_5)$$

- for two non-adjacent plane surfaces (for example, 1 and 3 in Figure 7.3):

$$S_1 F_{13} = \frac{1}{2} (S_5 + S_6 - S_2 - S_4)$$

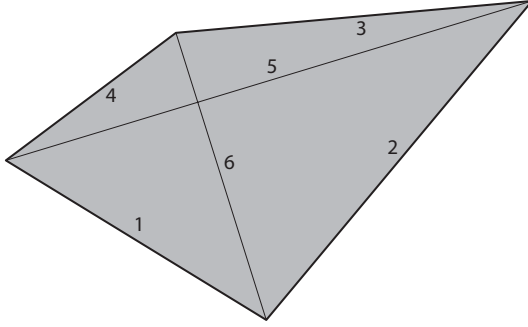


Figure 7.3. *The crossed string method*

In a 3D case¹, one possibility is to use Gauss numerical integration, which consists of replacing the integrals involved in expression [7.8] with a discrete sum on all the integration points of plane surfaces a and b . This method is described in detail in Part 1 of this book. We obtain:

$$s_a F_{ab} = \frac{1}{\pi} \sum_{i=1}^{n_a^g} \sum_{j=1}^{n_b^g} \frac{\overrightarrow{G_a^i G_b^j} \cdot (\vec{n}_a^i - \vec{n}_b^j)}{\|\overrightarrow{G_a^i G_b^j}\|^4} w_a^i w_b^j \nu_{ab}^{ij}$$

In this equation, $G_a^i, w_a^i, \vec{n}_a^i$ ($G_b^j, w_b^j, \vec{n}_b^j$) represent respectively an integration point, the related weight and the unit normal to the plane surface a (plane surface b). The coefficient ν_{ab}^{ij} takes the value 1 if the integration points are visible, and 0 should this be otherwise.

This technique does not enable us to accurately calculate (except if a large number of integration points are used) the view factors; however, it is naturally included in a finite element computer code and makes it possible to address any type of convex or non-convex cavity. Figures 7.4 and 7.5 give the accuracy expected in classic configurations.

These element cases are extremely interesting for the user as they give very useful information for determining the cavity surface discretization. It can be noted that for a reasonable discretization, the resulting accuracy is satisfactory. For 2 parallel planes,

1. Axisymmetric cases are particularly interesting and complicated to address. The reader may refer to [CHU 82].

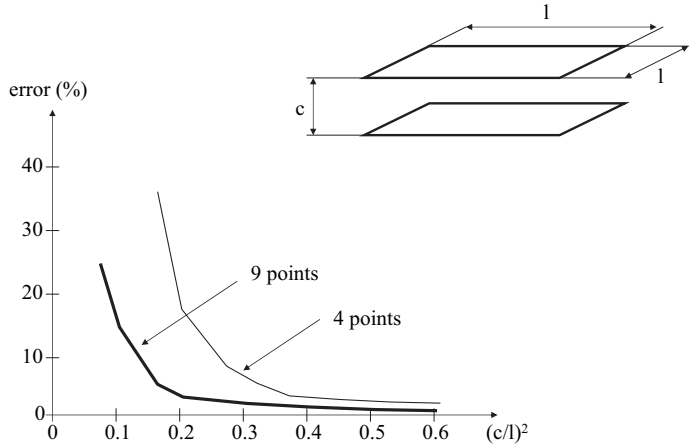


Figure 7.4. Calculation error of the view factor between two opposite plane surfaces with 4 and 9 Gauss points per plane surface

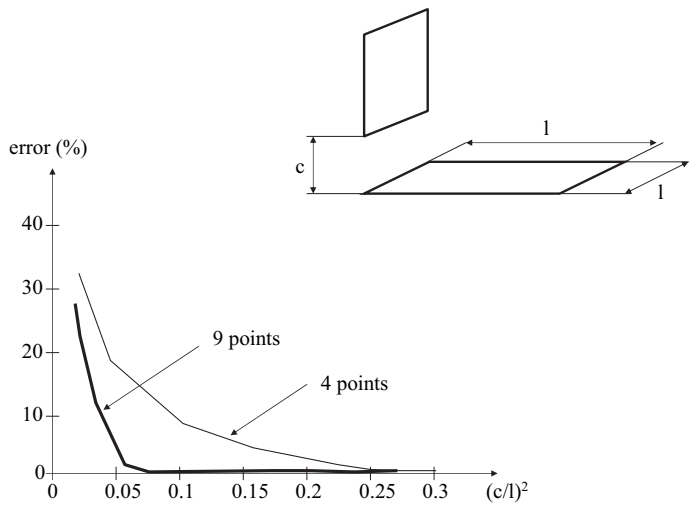


Figure 7.5. Calculation error of the view factor between two orthogonal plane surfaces with 4 and 9 Gauss points per plane surface

the error is less than 5% as soon as $\frac{c}{l} > 0.387$ for elements with 9 integration points and $\frac{c}{l} > 0.583$ with elements with 4 integration points. However, this method does not enable us to accurately calculate of the view factor between two plane surfaces with a common edge ($c = 0$ in 7.5).

The calculation of the view factor between two spheres makes it possible to validate this technique in the case of a non-convex cavity. Consider an inner sphere of radius 1 and an outer sphere of radius 2, each sphere being meshed with 96 quadrangles with 4 integration points and 32 triangles with 1 integration point. The sum of the view factors being calculated, related to each element, remains in the range of $0.9978 \leq \sum_{b=1}^f F_{ab} \leq 1.014$, which produces an error of less than 1.4%. Moreover, the view factor sought, from the inner sphere to the outer sphere, is equal to 0.25024, compared with a theoretical factor equal to 0.25, which is very satisfactory. In practice, above all in the case of a non-convex cavity, it is generally preferable to rectify the view factors after calculation, so as to satisfy equation [7.1] for each of the plane surfaces. It is therefore advisable to take the following iterative option. If $F_{ab}^{(k)}$ are the view factors obtained after k iterations, the following sequence is defined:

$$F_{ab}^{(k+1)} = \frac{1}{2} \left(\frac{1}{\phi_a^{(k)}} + \frac{1}{\phi_b^{(k)}} \right) F_{ab}^{(k)} \quad \text{with} \quad \phi_a^{(k)} = \sum_{b=1}^f F_{ab}^{(k)}$$

It is demonstrated that this sequence converges, and at the convergence, for any plane surface a , we find that $\phi_a = 1$.

Other view factor calculation techniques may be used (contour integrals, Monte Carlo method, etc.). These calculations are often parts of specialized programs [CHU 82, SHA 85]. The reader will find in [EME 91] a comparative study of the various calculation methods available.

7.1.3. Diffusion-radiation coupling

Assume a cavity surface meshed with f skin finite elements, each element a representing a plane surface S_a . In order to take the action of heat radiation inside the cavity into consideration, it is sufficient to add to the residual expression the contribution of the corresponding heat flux density q_a . This contribution is written as:

$$\{R_r\} = \sum_{a=1}^f [A^a]^T \cdot \{R_r^a\} \quad \text{with} \quad \{R_r^a\} = \int_{S_a} \{N^a\} q_a ds \quad [7.9]$$

where $[A^a]$ represents the location matrix and N^a the shape functions corresponding to element a .

We could stop at this point. However, heat radiation problems produce significant non-linearities and it is wise to insert a contribution due to these effects into the

iterative matrix $[M]^{(it)}$ corresponding to the non-linear system to solve. The additional term $\{R_r\}$ influences only the nodes attached to the cavity surface and only depends on these node temperatures. The contribution $[M_r]$ to the matrix $[M]^{(it)}$ is therefore a square matrix of dimension $n_r \times n_r$, where n_r represents the cavity's number of nodes.

Consider $\{T_r\}$ the temperature vector of the nodes attached to the cavity surface. In the case of radiative exchanges, the term $\{R_r\}$ depends on all the temperatures of the cavity surface nodes, and hence on all the components of $\{T_r\}$. It follows that matrix $[M_r]$ is full and its assembly will significantly increase the bandwidth of global matrix $[M]^{(it)}$. Another method often used is to assemble only the terms of $[M_r]$ included in the band structure of matrix $[M]^{(it)}$ corresponding to diffusion phenomena.

Two types of radiation matrix $[M_r]$ to assemble in $[M]^{(it)}$ at each resolution iteration can be considered:

– a real tangent matrix:

$$[M_r] = - \left[\frac{\partial R_r}{\partial T_r} \right]$$

– a substitution matrix:

$$[M_r] = - \left[\frac{\delta R_r}{\delta T_r} \right]$$

7.1.3.1. Tangent matrix

Reconsidering the expression of the heat flux density q_a given by equation [7.7], we obtain:

$$[M_r] = \frac{\partial}{\partial \{T_r\}} \left(\sum_{a=1}^f [A^a]^T \cdot \int_{S_a} \{N^a\} \frac{1}{S_a} \sum_{b=1, b \neq a}^f D_{ab} \sigma (\bar{T}_a^4 - \bar{T}_b^4) ds \right)$$

Then the mean temperature of a plane surface is defined with the finite element method nodal approximation:

$$\bar{T}_a = \frac{1}{S_a} \int_{S_a} T ds = \frac{1}{S_a} \left\langle \int_{S_a} N^a ds \right\rangle \cdot \{T^a\} \quad [7.10]$$

It is then easily demonstrated that the matrix² $[M_r]$ consists of the element blocks $[M_r^{ab}]$ influencing only the nodes of the plane surfaces a and b :

$$[M_r] = \sum_{a,b=1}^f [A^a]^T \cdot [M_r^{ab}] \cdot [A^b] \quad \text{with} \quad [7.11]$$

$$\begin{cases} [M_r^{aa}] = \frac{4\sigma\bar{T}_a^3}{S_a^2} \left(\sum_{b=1, b \neq a}^f D_{ab} \right) \left\{ \int_{S_a} N^a ds \right\} \left\langle \int_{S_a} N^a ds \right\rangle \\ [M_r^{ab}] = -\frac{4\sigma\bar{T}_b^3 D_{ab}}{S_a S_b} \left\{ \int_{S_a} N^a ds \right\} \left\langle \int_{S_b} N^b ds \right\rangle \quad \text{for } b \neq a \end{cases}$$

The tangent radiation matrix obtained by matrix assembly has the drawback of being non-symmetric.

7.1.3.2. Substitution matrix

Expression [7.9] of $\{R_r\}$ can be written taking into consideration expression [7.7]:

$$\{R_r\} = - \sum_{a=1}^f \frac{1}{S_a} \int_{S_a} \{N^a\} \sum_{b=1, b \neq a}^f \underbrace{D_{ab} \sigma (\bar{T}_a^2 + \bar{T}_b^2) (\bar{T}_a + \bar{T}_b) (\bar{T}_a - \bar{T}_b)}_{h_{ab}} ds$$

The quantity h_{ab} in the previous equation is similar to an exchange coefficient between two plane surfaces a and b . Given equation [7.10] defining a plane surface mean temperature, $\{R_r\}$ can be written as $\{R_r\} = -[M_r] \cdot \{T_r\}$, where the matrix $[M_r]$ is a substitution matrix which is written as:

$$[M_r] = \sum_{a,b=1}^f [A^a]^T \cdot [M_r^{ab}] \cdot [A^b] \quad \text{with}$$

$$\begin{cases} [M_r^{aa}] = \frac{1}{S_a^2} \left(\sum_{b \neq a}^f h_{ab} \right) \left\{ \int_{S_a} N^a ds \right\} \left\langle \int_{S_a} N^a ds \right\rangle \\ [M_r^{ab}] = \frac{h_{ab}}{S_a S_b} \left\{ \int_{S_a} N^a ds \right\} \left\langle \int_{S_b} N^b ds \right\rangle \quad \text{for } b \neq a \end{cases}$$

2. The possible temperature-dependence of heat emissivities is neglected for the calculation of the tangent operator.

As above, the matrices $[M_r^{ab}]$ are element blocks influencing only the nodes attached to the plane surfaces a and b . It is easy to demonstrate that the symmetry of the D_{ab} leads to that of the h_{ab} , and consequently the symmetry of $[M_r]$ thus defined. This is a major advantage of this method.

7.2. Examples

Two application examples are addressed in this section:

- radiation between two walls, which will validate the method presented in the section above;
- cylinder quenching, which uses symmetry relationships to simplify the global problem [BER 01].

7.2.1. Radiation between two walls

Consider the abstract 1D example in Figure 7.6. The walls are assumed to have a thermal conductivity $\lambda = 1.5 \text{ W/m/K}$. The analytical solution is obtained, in the steady state, by equalizing the set of flux densities involved, i.e. the conduction between the surfaces A and B , the conduction between the surfaces C and D , and the radiation between the surfaces B and C . If ϕ_r is the radiated flux density, we obtain:

$$-\lambda \frac{T_B - T_A}{x_B - x_A} = \phi_r = -\lambda \frac{T_D - T_C}{x_D - x_C}$$

It is easily demonstrated that if F represents the view factor between the surfaces B and C , then the radiation flux density between these two surfaces is written:

$$\phi_r = K\sigma(T_B^4 - T_A^4) \quad \text{with} \quad K = \frac{F\epsilon_B\epsilon_C}{1 - F^2(1 - \epsilon_B)(1 - \epsilon_C)}$$

The mesh used for the calculations is regularly divided, every 0.05 m, in the direction x . Five solution strategies are examined:

- NC (no radiation contribution) method: the conduction matrix alone is used to make the iterations;
- PS (partial substitution) method: the radiation contribution is introduced as a substitution matrix, with consideration of only those terms belonging to the band structure corresponding to the diffusion finite elements;
- TS (total substitution) method: the radiation contribution is taken into consideration by means of the symmetric substitution matrix;

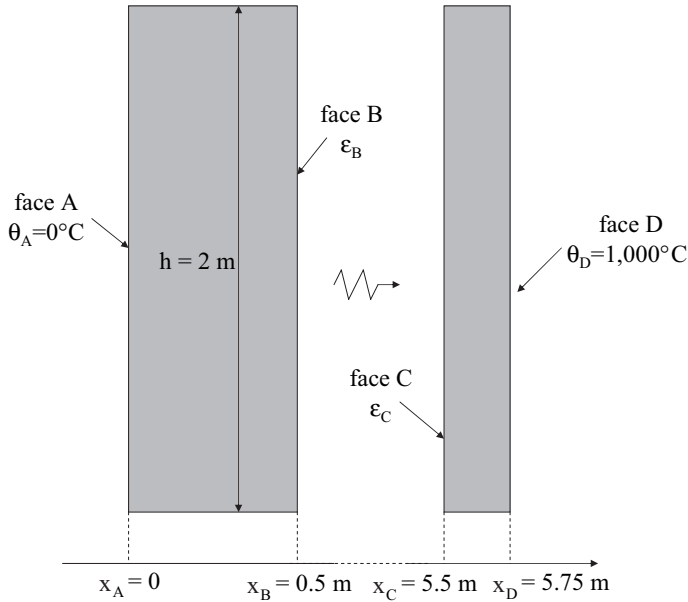


Figure 7.6. Radiation exchange between two walls

- TM (tangent matrix) method: the radiation contribution is taken into consideration by means of the tangent matrix (non-symmetric);
- BFGS method: the BFGS method is used.

In all cases, the iterative calculation is interrupted as soon as the largest residual component (in absolute value) becomes less than 0.01 W , or as soon as that of the temperature difference is less than 0.01°C . Different couples (ϵ_A , ϵ_B) are tested; all the results are in Table 7.1.

As could be expected, it is with the real tangent matrix that convergence is fastest (in number of iterations). The substitution matrix produces convergence too, without using a non-symmetric solver. In practice, these methods very significantly increase the bandwidth of the first member matrix used to make the iterations and, as a consequence, the analysis cost.

The conduction matrix alone does not produce convergence generally. Introducing a radiation partial contribution, without destroying the matrix band structure, has produced convergence in all cases (often with a large number of iterations, if truth be told). The BFGS method seems to be the most efficient.

ϵ_B	ϵ_C	T_B (°C)	T_C (°C)	method	iterations
0.1	0.3	204.4	897.8	PC	14
				SP	9
				ST	4
				MT	3
0.4	1	581.2	709.4	PC	∞
				SP	41
				ST	6
				MT	4
0.01	0.2	19.06	990.5	PC	4
				SP	4
				ST	3
				MT	3
1	1	630.9	684.6	PC	∞
				SP	93
				ST	5
				MT	4
				BFGS	12
0.2 +0.0008T	0.2 +0.0008T	599.5	700.3	PC	∞
				SP	45
				ST	7
				MT	6
				BFGS	13

Table 7.1. Work example – number of iterations at the convergence

The example considered here is much more difficult than it first appears. On the one hand, the calculation of a steady state is much more difficult to make than that of a transient state where the phenomena are regularized over time. On the other hand, the example deals with two conductive regions which are only coupled by the radiative exchanges. This coupling is only inserted into the first member matrix when a radiation contribution by means of a substitution matrix or a tangent matrix is considered. That is the reason why the methods using them converge in fewer iterations.

Generally speaking, it is recommended to treat diffusion-radiation problems with the BFGS method. Using a substitution matrix will be restricted to cases in which the latter does not converge. Using the tangent matrix, in addition to the hopeless cases in which none of the methods above would have produced a solution, will be recommended in cases where the first member matrix is already non-symmetric in nature, for example, as is the case of diffusion-convection problems.

7.2.2. Cylinder quenching

Consider here the case of five cylinders at the initial temperature of 900°C, which are cooled in a chamber whose walls and top surface are assumed to be at the constant temperature of 200°C, with an emissivity of 1 (black body), and whose bottom surface is at 60°C, with an emissivity of 0.6. The cylinder emissivity is assumed to be equal to 0.9.

We added to the radiative exchanges convection exchanges on the cylinder surfaces, with a convection coefficient $h = 10 \text{ Wm}^{-2}\text{K}^{-1}$, and a temperature equal to that of the chamber walls, 200°C. The thermal properties of the cylinders are $\lambda = 40 \text{ Wm}^{-1}\text{K}^{-1}$ and $\rho C = 4.7 \cdot 10^6 \text{ Jm}^{-3}\text{K}^{-1}$.

The numerical simulation of this problem takes the symmetries into consideration by means of the formulation described in detail in [BER 01]. In order to illustrate the influence of symmetries on calculation times, we have made two calculations with two meshes presented in Figure 7.7. The first image represents the complete structure, without the chamber walls. It altogether consists of 2,761 nodes, 2,400 3D isoparametric elements. Furthermore, 1,720 2D elements were added to illustrate the radiative exchanges in the chamber. In the second image, the problem symmetries lead us to model only one-eighth of the whole structure. The mesh used for the calculation then includes only 437 nodes, 300 3D elements and 215 2D elements.

The view factors F_{ab} between each surface and the chamber walls are calculated only once at the beginning of the analysis. The transient calculation is made with a time step of 20 s between the initial moment and 60 s, then a step of 60 s between moments 60 s and 600 s. The solution was calculated with a tangent matrix restricted to the assembly of the terms $[M_r^{aa}]$ in equation [7.11]. The convergence was set at 10^{-5} W on the residual maximum component (in absolute value), or at 10^{-2}°C on the maximum temperature variation (in absolute value) between two iterations.

Figure 7.8 gives the temperature distribution on the cylinder surfaces at $t = 240 \text{ s}$. We see that the reduced model gives the same distribution as the full model, whereas the calculation times were as follows:

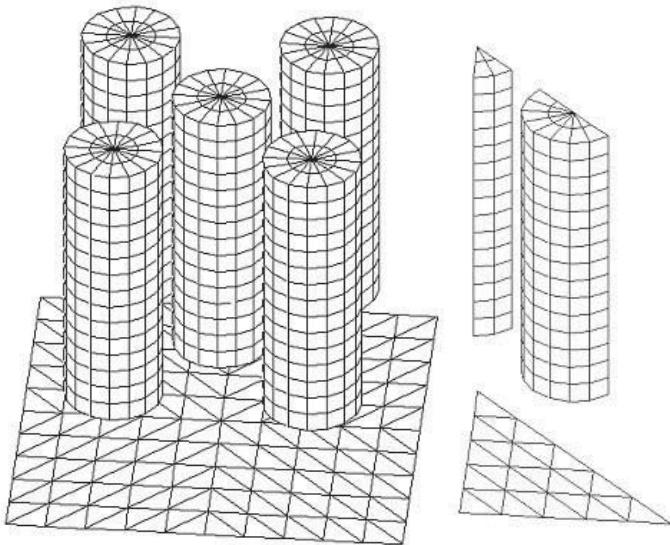


Figure 7.7. Mesh used for the calculation, from [BER 01]

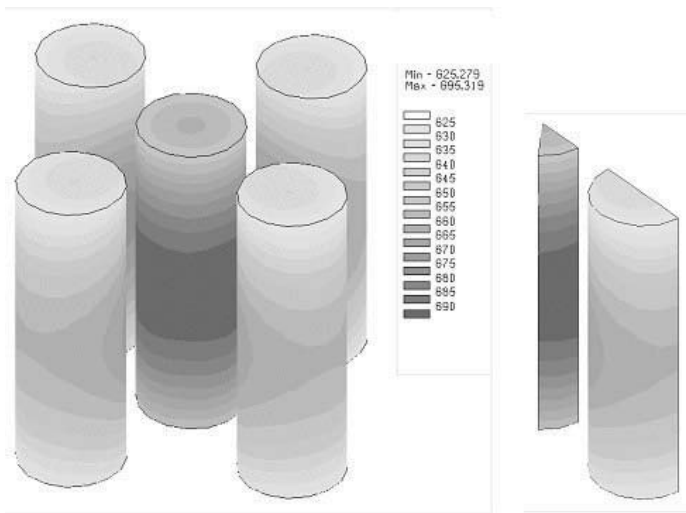


Figure 7.8. Temperature distribution at $t = 240\text{ s}$, excerpt from [BER 01]

- on the full model, 669 s for the view factors, 25 s for the calculation of the radiation contribution to the tangent matrix, and 54 s for the transient thermal calculation,
- on the reduced model (one eighth on account of symmetries), 79 s for the view factors, 0.07 s for the calculation of the radiation contribution to the tangent matrix, and 5 s for the transient thermal calculation.

These results show how important it is to use the symmetries of a problem when it is modeled, with a view to its numerical simulation. Additionally, using a partial tangent matrix (radiation contribution restricted to the terms $[M_r^{aa}]$ in equation [7.11]) was not very prejudicial, for the convergence of the solution process (at each time step) was obtained with a maximum of five iterations. Finally, the following points can be noted:

- as expected, the calculation of the view factors takes nine times as long as when the problem symmetries are not considered. Indeed, it is demonstrated that the method used here to take the symmetries into consideration reduces the time for the calculation of the view factors by a ratio $N_{sym} + 1$ where N_{sym} is the number of times when the simulation model has to be duplicated to obtain the real model (here, $N_{sym} = 8$);
- in both cases, the calculation time of the radiation contribution to the tangent matrix is negligible compared with that of the view factor calculation;
- of course, the transient thermal calculation time is significantly reduced when the problem symmetries are used.

This page intentionally left blank

Chapter 8

Fluid-Structure Coupling in a Pipe

8.1. Modeling the fluid

Many industrial applications are related to heat exchanges. Modeling such devices requires the correct consideration of the heat exchanges between the heat-carrying fluid and the wall. This problem is all the more difficult since the temperature of the fluid, because of its exchanges with the wall, changes as it moves along. Modeling must therefore make it possible to calculate these temperature changes and take into consideration the fluid local temperature in the complex heat exchanges occurring with the wall. This is a presentation of the physical interaction model and its mathematical formulation, leading axiomatically to a use of the finite element method. Finally, we will deal with the elements creating the fluid-structure coupling [BER 98].

8.1.1. *Physical model and mathematical formulation*

Modeling consists of representing the heat-carrying fluid with a set of 1D elements and solving a diffusion-convection problem in these elements by means of the technique introduced in Chapter 6. The heat exchanges between the fluid and the wall are supported by 2D elements and involve the fluid temperature and wall temperature.

The pipe is therefore subdivided into macro-elements. Each macro-element combines:

- a 2-noded 1D fluid element;
- a set of 2D elements representing the structure wall and supporting the heat exchange conditions between the fluid and the wall (exchange elements).

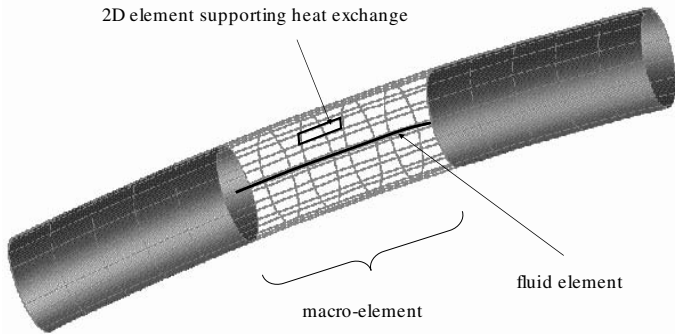


Figure 8.1. *Description of the pipe*

The macro-element mathematical formulation is based upon the following assumptions:

- A1: the fluid flow is assumed to be steady and 1D. The flow direction is given by the direction of the elements representing the fluid;
- A2: the fluid mass flow \dot{m} is assumed to be constant along the pipe, but may possibly vary over time;
- A3: inside each macro-element, heat exchanges can only occur between the fluid element and the set of exchange elements corresponding to the macro-element. The heat flux density received locally by the wall can thus be written as:

$$\phi(T_p, \bar{T}_f) = h(\bar{T}_f - T_p) \quad [8.1]$$

where T_p , \bar{T}_f and h represent respectively the wall temperature, the fluid mean temperature (in the fluid element) and the exchange coefficient between fluid and wall.

In practice, the heat flux density received locally by the wall will be calculated at each integration point of the exchange elements. The wall temperature under consideration will therefore be the one calculated at each integration point. Moreover, the expression of h depends on the heat exchange conditions (natural convection, forced convection or mixed convection) and may depend on T_p and \bar{T}_f . The expression of h results from a Nusselt number by means of various empirical relationships between numbers without dimension, which are functions of the convection type (natural, forced, mixed, etc.), and possible phase changes in the heat-carrying fluid (vaporization, etc.), among others.

The heat equation, considering assumptions A1 and A2, is written as:

$$\frac{\dot{m}}{S_f} C_f \frac{\partial T_f}{\partial l} - \frac{\partial}{\partial l} \left(\lambda_f \frac{\partial T_f}{\partial l} \right) - q = 0 \quad [8.2]$$

In this equation, l represents the curvilinear x-axis, S_f the pipe section, C_f the specific heat, λ_f the fluid thermal conductivity, and q the heat flux received per fluid volume unit. The quantity q is assumed to be constant in each fluid element. As seen in Chapter 6, using a classic Galerkin method to formulate the finite elements from equation [8.2] produces spatial oscillations of the solution when the Peclet number is greater than a critical value (2 for first order elements). It is obvious that such a limitation is not acceptable for modeling heat exchangers. To avoid these oscillations, a SUPG method can be used. This approach produces the following equation system for the fluid:

$$([K_f] + [U_f] + [D_f]) \cdot \{T_f\} = \{F_f\} \quad [8.3]$$

In this equation, $[K_f]$ represents the classic conductance matrix (obtained as assembly of the element conductance matrices in the fluid), $\{T_f\}$ represents the fluid nodal temperature vector, whereas $[U_f]$, $[D_f]$ and $\{F_f\}$ are obtained by assembling on the fluid elements Ω_f^e the following element quantities:

$$U_{fij}^e = \int_{\Omega_f^e} (N_i^e + H_i^e) \dot{m} C_f \frac{\partial N_j^e}{\partial l} dl \quad [8.4]$$

$$D_{fij}^e = -S_f \int_{\Omega_f^e} H_i^e \frac{\partial}{\partial l} \left(\lambda_f \frac{\partial N_j^e}{\partial l} \right) dl \quad [8.5]$$

$$F_{fi}^e = S_f \int_{\Omega_f^e} (N_i^e + H_i^e) q dl \quad [8.6]$$

The functions N_i^e and H_i^e appearing in these equations represent respectively the shape function and the disturbance function corresponding to node i of element e . It can be noted that, for first order elements and a constant thermal conductivity, $[D_f] = 0$.

Considering assumption A3 and equation [8.1], the thermal balance inside a macro-element is written as follows:

$$qV_f + \sum_{\text{exchange elements}} \int_{\text{element}} h(\bar{T}_f - T_p) ds = 0$$

The first term of this equation represents the heat flux received by the fluid (V_f is the fluid element volume). The second term represents the heat flux received by the structure. From this equation the expression of the quantity q is directly derived:

$$q = -\frac{1}{V_f} \sum_{\text{exchange elements}} \int_{\text{element}} h(\bar{T}_f - T_p) ds \quad [8.7]$$

8.1.2. Modeling the coupling

The equation system to solve for the solid structure can, in the steady state, be written:

$$[K_s] \cdot \{T_s\} = \{F_s\} \quad [8.8]$$

In this equation, $[K_s]$ is the traditional thermal conductance matrix of the solid, $\{T_s\}$ is the unknown (temperature) vector corresponding to the solid structure and $\{F_s\}$ is the load vector. If, for convenience, it is assumed that the solid structure exchanges heat only with the fluid flowing in the pipe, it is possible to write:

$$\{F_s\} = \sum_{\text{macro-elements } m-e} [A^{m-e}]^T \cdot \{F_s^{m-e}\} \quad [8.9]$$

In this equation, matrix $[A^{m-e}]$ is the assembly matrix described in the first chapter of this book, applied to the macro-element $m-e$. Load vector $\{F_s\}$ therefore results from the contributions of all the macro-elements $m-e$, contributions which can be written as follows:

$$\{F_s^{m-e}\} = \sum_{\text{exchange elements } e} [A^e]^T \{Q_s^e\} \quad [8.10]$$

Again, in this equation, assembly matrix $[A^e]$ is that described in Chapter 1. The components Q_{si}^e of element term $\{Q_s^e\}$ are finally written:

$$Q_{si}^e = \int_{\Omega^e} N_i^e h(\bar{T}_f - T_p) ds \quad [8.11]$$

Inside each exchange element, T_p is given by the classic finite element approximation, where T_j^e is the temperature of node number j of element e (local numbering):

$$T_p(\vec{x}) = \sum_{j=1}^{n^e} N_j^e(\vec{x}) T_j^e$$

\bar{T}_f is the fluid's mean temperature. It is obtained, like the average of temperatures T_{f1} and T_{f2} , at the two nodes of the fluid element:

$$\bar{T}_f = \frac{T_{f1} + T_{f2}}{2}$$

The non-linear problem composed of equations [8.3] and [8.8] coupled by equations [8.6], [8.7] and [8.11] can be solved by means of a Newton-Raphson iterative method. Equations [8.3] and [8.8] are written as:

$$\{R\} = \begin{Bmatrix} R_s \\ R_f \end{Bmatrix} = \begin{Bmatrix} F_s \\ F_f \end{Bmatrix} - \begin{bmatrix} K_s & 0 \\ 0 & K_f + U_f + D_f \end{bmatrix} \cdot \begin{Bmatrix} T_s \\ T_f \end{Bmatrix} \quad [8.12]$$

At each resolution iteration (k), a better solution is obtained by solving a linear system (see Chapter 5). The process is repeated until the norm (Euclidian, for instance) of the residual $\{R\}$ becomes less than the desired accuracy. Equation [8.12] can also be written:

$$\{R\} = \sum_{\text{macro-elements}} [A^{m-e}]^T \{R^{m-e}\} - \begin{bmatrix} K_s & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{Bmatrix} T_s \\ T_f \end{Bmatrix}$$

In this equation, $\{R^{m-e}\}$ represents the contribution of the macro-element $m - e$ to the residual:

$$\{R^{m-e}\} = \begin{Bmatrix} F_s^{m-e} \\ F_f^{m-e} \end{Bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & K_f^{m-e} + U_f^{m-e} + D_f^{m-e} \end{bmatrix} \cdot \begin{Bmatrix} T_s^{m-e} \\ T_f^{m-e} \end{Bmatrix} \quad [8.13]$$

In this expression, $\{F_s^{m-e}\}$ is defined by equation [8.10], whereas $\{F_f^{m-e}\}$, $[K_f^{m-e}]$, $[U_f^{m-e}]$ and $[D_f^{m-e}]$ are the load vector and element matrices of the only fluid element corresponding to the macro-element $m - e$. The terms $[U_f^{m-e}]$, $[D_f^{m-e}]$ and $\{F_f^{m-e}\}$ are given by equations [8.4], [8.5] and [8.6]. Finally, the terms $\{T_s^{m-e}\}$ and $\{T_f^{m-e}\}$ are the restrictions of the vectors $\{T_s\}$ and $\{T_f\}$ at the only nodes connected to the macro-element $m - e$.

The contribution $[K^{m-e}]$ of each macro-element to the tangent matrix $[K_T]$ of the global system is directly derived from equation [8.13]:

$$[K_T^{m-e}] = \begin{bmatrix} -\frac{\partial F_s^{m-e}}{\partial T_s^{m-e}} & -\frac{\partial F_s^{m-e}}{\partial T_f^{m-e}} \\ -\frac{\partial F_f^{m-e}}{\partial T_s^{m-e}} & K_f^{m-e} + U_f^{m-e} + D_f^{m-e} - \frac{\partial F_f^{m-e}}{\partial T_f^{m-e}} \end{bmatrix}$$

In this equation, it can be noted that the matrix $[K_T^{m-e}]$ is not symmetric.

8.2. Example

The example given here is that of the auxiliary sprinkling line of a vapor generator in a 1,300 MW PWR nuclear plant. It is a security system designed to remedy a failure in the primary pumps, by injecting water to keep the pressure constant and equal to 17.13 bars in the circuit. We first describe the physical and geometric modeling used, then analyze a few results.

8.2.1. Physical and geometric modeling

The geometry is very simple here. It is a 51.1 m long pipe modeled with 1,200 volumetric elements. The exchange with the fluid is represented by 100 fluid-structure coupling macro-elements (Figure 8.2). The geometric and physical specifications used for the simulation are the following, considering that the pipe is assumed to be adiabatic on the outer wall:

- pipe dimensions: 51.1 m long, 60.3 mm in internal diameter, 8.7 mm thick,
- initial temperature: 20°C,



Figure 8.2. Sprinkling pipe mesh

- pipe specifications: conductivity $15 \text{ W/m/}^\circ\text{C}$, volumetric mass $7,800 \text{ Kg/m}^3$, specific heat capacity $500 \text{ J/Kg/}^\circ\text{C}$,
- water specifications: conductivity $0.603 \text{ W/m/}^\circ\text{C}$, volumetric mass $1,000 \text{ Kg/m}^3$, specific heat capacity $4,180 \text{ J/Kg/}^\circ\text{C}$,
- mass flow rate: 12 tonnes/hour ,
- exchange coefficient: $10,400 \text{ W/m}^2/^\circ\text{C}$.

8.2.2. Results

Figure 8.3 gives the water temperature variation in the pipe, over time, at the entry and exit. It can be noted that the water injected under pressure at the entry reaches at the exit a temperature of 280°C after about 150 s .

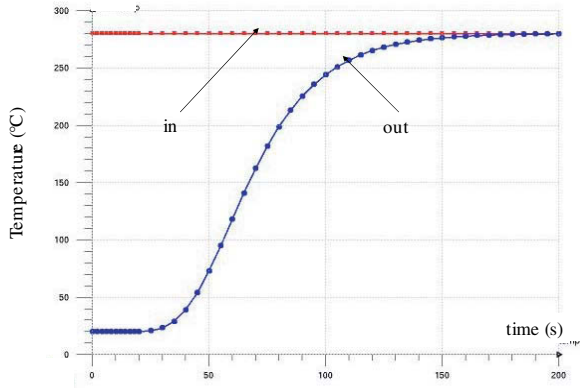


Figure 8.3. Temperature variation over time, at the sprinkler entry and exit

This page intentionally left blank

Chapter 9

Thermometallurgical Coupling

9.1. Modeling phase changes

This chapter deals with the interactions between heat exchanges and phase changes in a solid material. The first section is a description of the most common models used to describe the variations of phase proportions in a material undergoing thermal cycles [LEB 84, FER 85, INO 85]. To do so, differential equations governing each phase change are figured first, followed by the method to integrate them numerically along any thermal path. Eventually, the case of multiple transformations which may occur simultaneously will be dealt with.

9.1.1. *Rate of phase changes*

Solid state phase change models usually distinguish two types of transformation:

- those governed by diffusion, described by Avrami kinetics [AVR 40a, AVR 40b, AVR 41],
- those of the martensitic type, represented by the Koistinen-Marburger law [KOI 59].

9.1.1.1. *Avrami kinetics*

Avrami kinetics used to represent diffusion phase changes result from the extended volume concept [AVR 40a, AVR 40b, AVR 41]. In Figure 9.1, the extended volume V_e is obtained by considering that the new phase germination may occur anywhere in the material. Its variation velocity is therefore proportional to the new phase germ creation rate, \dot{N} , and the mean volume of these germs, v_g . The variation velocity of the phase

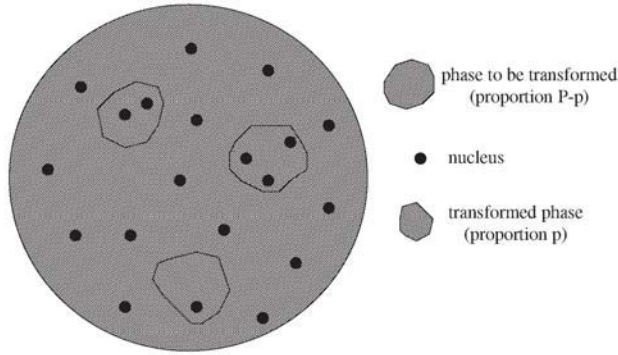


Figure 9.1. *Extended volume concept*

proportion formed is in turn proportional to the extended volume variation velocity \dot{V}_e , as well as to the phase proportion left for transformation. Thus, we obtain:

$$\dot{p} = (P - p)\dot{V}_e$$

In this equation, P designates the maximum phase proportion that can be formed, i.e. that which will be produced at the end of the transformation process.

The classic form of Avrami kinetics is obtained with the previous equation. For this purpose, it has to be integrated at a constant temperature (temperature contour path), and use a power law for the extended volume variation velocity:

$$\dot{V}_e = nfV_e^{\frac{n-1}{n}} \implies p = P(1 - e^{-(ft)^n})$$

Here, f and n are parameters of the phase change kinetics, at the temperature under consideration. If the same expression is kept for \dot{V}_e , it is also possible to define with the previous equations the variation velocity \dot{p} of the phase proportion p . This variation speed is written as follows, considering that the parameters P , f and n may be temperature-dependent:

$$\dot{p} = nf(P - p) \left[\ln \left(\frac{P}{P - p} \right) \right]^{\frac{n-1}{n}} \quad [9.1]$$

9.1.1.2. Martensitic kinetics

The martensitic transformation occurs instantly. Martensitic kinetics therefore involve only the temperature, and are represented by the Koistinen-Marburger law [KOI 59]. According to this law, the (martensitic) phase proportion p formed at temperature T is given by:

$$p(T) = \begin{cases} 0 & \text{for } T > M_s \\ 1 - e^{b(T-M_s)} & \text{for } T \leq M_s \end{cases}$$

Here, b is a homogenous kinetic factor, unlike temperature, and M_s (martensite start) is the temperature at the beginning of the transformation. These coefficients may depend on the material. For instance, in the case of steel, M_s is a function of the transformed austenite carbon content.

As with Avrami kinetics, martensitic transformation kinetics can be figured in the differential form. This is done by deriving the previous equation with respect to time. We obtain:

$$\dot{p} = \begin{cases} 0 & \text{for } T > M_s \text{ or } \dot{T} \geq 0 \\ -b(1-p)\dot{T} & \text{otherwise} \end{cases} \quad [9.2]$$

It can be noted that this equation can be written like equation [9.1]. To do so, it is sufficient to choose in this equation $n = 1$, $P = 1$, and a frequency f depending on the thermal path:

$$f = \begin{cases} 0 & \text{when } T > M_s \text{ or } \dot{T} \geq 0 \\ -b\dot{T} & \text{otherwise} \end{cases}$$

9.1.2. Numerical integration

Different methods can be used to integrate equations [9.1] and [9.2]. For instance, in the case of Avrami kinetics, the thermal path can be divided into a series of temperature contour plateaus, of length Δt , along which it is possible to apply an integrated form of the kinetics (Figure 9.2 [PUM 48, FER 85]). However, it is also possible to perform a classic numerical integration, with time steps Δt , with a second order Runge-Kutta method (Figure 9.3 [LEB 84]).

In the case of Figure 9.2, a dummy time t^* is introduced to take into consideration the phase transformation already formed at the beginning of the time increment. The

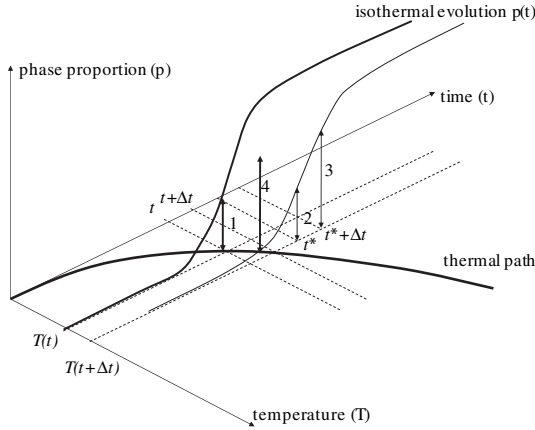


Figure 9.2. Numerical integration of a phase proportion along a thermal path with the dummy time concept

value of t^* is the time that would have been required to form a phase proportion $p(t)$ from zero, at temperature $T(t + \Delta t)$ of the temperature contour plateau (point 2 of Figure 9.2). This dummy time is obtained in the following form, all the parameters being estimated at temperature $T(t + \Delta t)$:

$$p(t) = P(1 - e^{(ft^*)^n}) \quad \text{or} \quad t^* = \frac{1}{f} \left[\ln \left(\frac{P}{P - p(t)} \right) \right]^{\frac{1}{n}}$$

The dummy time t^* thus obtained is then used as the starting point of a temperature contour transformation occurring at a temperature $T(t + \Delta t)$. The final phase proportion is therefore obtained analytically:

$$p(t + \Delta t) = P(1 - e^{(f(t^* + \Delta t))^n})$$

In the case of Figure 9.3, intermediate points are introduced by the classic second order Runge-Kutta scheme to integrate equations [9.1] or [9.2] along the thermal path $T(t)$. Point 2 is obtained by choosing a constant variation velocity of the phase proportion, on a temperature contour variation curve at temperature $T(t)$. This point is used to construct point 2', which represents a phase proportion similar to that of point 2, but on the temperature contour variation curve of p corresponding to temperature $T(t + \Delta t)$. Point 3 is then obtained like point 2, but with the variation velocity of p obtained at point 2'. Finally, the last point, 4, is obtained by calculating the mean value of the phase proportions obtained at points 2 and 3. Note that in Figure 9.3 this point is not located on the phase proportion or temperature contour variation curves. For

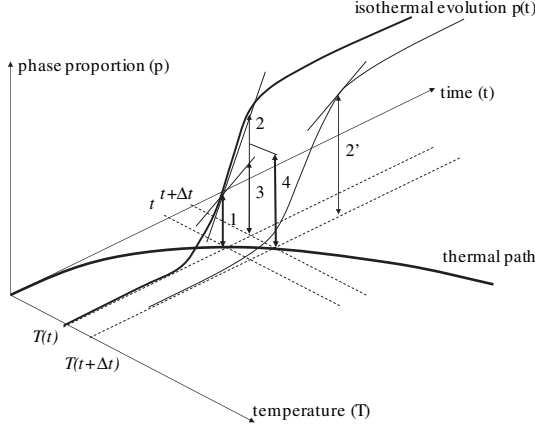


Figure 9.3. Numerical integration of a phase proportion along a thermal path with a second order Runge-Kutta scheme

instance, let $\dot{p}(p, t)$ be the phase proportion variation velocity obtained by equation [9.1], for a common proportion p , and with parameters P , f and n estimated at instant t of the thermal path (at temperature $T(t)$). The second order Runge-Kutta method then gives the phase proportion at instant $t + \Delta t$ by estimating consecutively the following quantities:

$$\begin{cases} \dot{p}_1 = \dot{p}(p(t), t) \\ p_2 = p(t) + \dot{p}_1 \Delta t \\ \dot{p}_{2'} = \dot{p}(p_2, t + \Delta t) \\ p_3 = p(t) + \dot{p}_{2'} \Delta t \end{cases} \quad \text{then } p(t + \Delta t) = \frac{1}{2}(p_2 + p_3)$$

This can be written globally as:

$$p(t + \Delta t) = p(t) + \frac{1}{2}(\dot{p}(p(t), t) + \dot{p}(p(t) + \dot{p}(p(t), t)\Delta t, t + \Delta t))\Delta t)$$

The Runge-Kutta method makes it possible to integrate, with the same method, the Avrami and martensitic kinetics, providing that the latter are formulated like Avrami kinetics with $n = 1$, $P = 1$ and a frequency f depending on the thermal path. Hereafter, therefore, general formulation [9.1] will be used for all possible phase changes. If it is of the Avrami kinetics type, the parameters n , P and f of this formulation will be temperature-dependent. If it is the martensitic type kinetics, we will choose $n = 1$, $P = 1$, and a frequency f depending on the temperature variation velocity (heating or cooling) and the value of M_s (see section 9.1.1).

9.1.3. The case of several phase changes

Actually, several different phases may be present in the material. For instance, in low-carbon steel, we can find austenite, ferrite, perlite, bainite and martensite, i.e. five phases, and transformations between these phases may occur in a similar volume element. In that case, if p_i designates the phase's common proportion i , then it is possible to define a driving force E_{ij} of the transformation $i \rightarrow j$ of the form $E_{ij} = P_{ij}(p_i + p_j) - p_j$, where P_{ij} is the maximal proportion of phase j which can be obtained from phase i . It should be noted at this point that, if the driving force E_{ij} is positive, then the driving force E_{ji} (corresponding to the transformation $j \rightarrow i$) is negative. These driving forces therefore indicate the transformation direction between phases i and j .

The problem of transformations between several phases can now be solved [LEB 84, FOR 00b]:

– for each transformation $i \rightarrow j$, by substituting p into $\frac{p_j}{p_i + p_j}$ in differential equation [9.1],

– by considering that the variation of the phase proportion j , \dot{p}_j , results from all the possible transformations $i \rightarrow j$ and $j \rightarrow i$, which are assumed to be independent of one another.

With this method, it is possible to write:

$$\dot{p}_j = \sum_{i \neq j} A_{ij} \quad [9.3]$$

where A_{ij} can be considered to be the reaction activity between phases i and j . This activity is obtained from equation [9.1] as:

$$A_{ij} = \begin{cases} n_{ij} f_{ij} E_{ij} \left[\ln \left(1 + \frac{p_j}{E_{ij}} \right) \right]^{\frac{n_{ij}-1}{n_{ij}}} & \text{if } E_{ij} > 0 \text{ (reaction } i \rightarrow j) \\ -n_{ji} f_{ji} E_{ji} \left[\ln \left(1 + \frac{p_i}{E_{ji}} \right) \right]^{\frac{n_{ji}-1}{n_{ji}}} & \text{if } E_{ji} > 0 \text{ (reaction } j \rightarrow i) \\ 0 & \text{otherwise} \end{cases} \quad [9.4]$$

The parameters P_{ij} , n_{ij} and f_{ij} of this equation depend on the transformation under consideration. Note finally that the activities A_{ij} given by equation [9.4] make up an antisymmetric matrix. Indeed, as A_{ij} represents the phase proportion i which is transformed into phase j per time unit, we obtain $A_{ij} + A_{ji} = 0$. One of the

consequences is that the sum of all the matrix terms is zero, which expresses the conservation of the phase proportions since the sum of their variation velocities is equal to zero.

Another model was suggested for steel and inserted into various calculation software [WAE 94]. By making p_1 , p_2 and p_3 the common proportions of ferrite, perlite and bainite, and M_s the temperature of initial martensite transformation, for these phases the model equations are written as:

$$\begin{cases} \dot{p}_1 = f_1(T, \dot{T}, p_1, p_2, p_3, M_s, d) \\ \dot{p}_2 = f_2(T, \dot{T}, p_1, p_2, p_3, M_s, d) \\ \dot{p}_3 = f_3(T, \dot{T}, p_1, p_2, p_3, M_s, d) \\ M_s = M_s(p_1, p_2, p_3, d) \end{cases}$$

Martensitic transformation is described by equation [9.2]. Functions f_1 , f_2 , f_3 and M_s are not made explicit, but calculated point after point from the states $(T, \dot{T}, p_1, p_2, p_3, M_s, d)$ corresponding to the continuous cooling transformation diagram (or *CCT* diagram) of the shade treated. As a consequence, the model is very close to this diagram and is no longer very accurate when the cooling conditions depart from those of the diagram. Its major advantage, however, is that it is easy to identify.

9.1.4. Modeling the coupling

Strong coupling between thermal and metallurgical phenomena is relatively well-known. Phase changes are time- and temperature-dependent, and produce in turn temperature changes (exothermic and endothermic reactions). Generally speaking, the heat equation must be solved to obtain the temperature field in a domain Ω subjected to temperature or flux density thermal boundary conditions. This problem can be formulated as follows:

$$\begin{cases} \operatorname{div}(\underline{\lambda} \cdot \overrightarrow{\operatorname{grad}}(T)) + Q - \rho \dot{H} = 0 & \text{in } \Omega \\ T = T_d & \text{on } \partial\Omega_T \\ (\underline{\lambda} \cdot \overrightarrow{\operatorname{grad}}(T)) \cdot \vec{n} = q & \text{on } \partial\Omega_q \end{cases} \quad \text{with } \partial\Omega = \partial\Omega_T \cup \partial\Omega_q \quad [9.5]$$

In this equation, Q is an internal heat source, q a flux surface density entering Ω , and λ , ρ and H represent respectively the thermal conductivity, the volumetric

mass and the specific enthalpy of the material. The material properties λ , ρ and H are temperature- and phase proportion-dependent. They are generally obtained from linear mixture laws on the corresponding properties of each phase i , λ_i , ρ_i and H_i which are temperature-dependent only. If N is the total number of phases, we therefore write:

$$\begin{cases} \lambda(p_1, \dots, p_N, T) = \sum_{i=1}^N p_i \lambda_i(T) \\ \rho(p_1, \dots, p_N, T) = \sum_{i=1}^N p_i \rho_i(T) \\ H(p_1, \dots, p_N, T) = \sum_{i=1}^N p_i H_i(T) \end{cases} \quad [9.6]$$

It follows that the enthalpy variation velocity involved in heat equation [9.5] is written:

$$\dot{H} = \dot{T} \sum_{i=1}^N p_i C_i + \sum_{i=1}^N \dot{p}_i H_i \quad \text{with} \quad C_i = \frac{dH_i}{dT} \quad [9.7]$$

In this equation, C_i represents the heat capacity of phase i , and the terms \dot{p}_i can be obtained from equations [9.3] and [9.4]. The first term of equation [9.7] involves the phase mixture heat capacity, $\sum_{i=1}^N p_i C_i$. The second term represents the latent heat effects. Indeed, the curves $H_i(T)$ giving the enthalpy of the phases i according to the temperature, are different, not only in their slopes, representing the phase heat capacities, but also in their y -axes at the origin, which includes these latent heat effects. This is illustrated by Figure 9.4.

Inserting equation [9.7] into [9.5] makes it possible to treat heat exchange problems, including phase changes, on complex geometries. For this purpose, the finite element method can be used. Before showing a few examples, we will deal with the case of phase transformation diagrams, which can be simulated to validate the parameters included in the expression of the activities of equation [9.4].

9.2. Examples

In this section, we give a few examples of numerical process simulations in which thermal and metallurgical aspects are strongly coupled. The initial focus will be on phase transformation diagrams, which are used to obtain the parameters of the phase change laws used. Later, we give the example of the simulation of steel quenching.

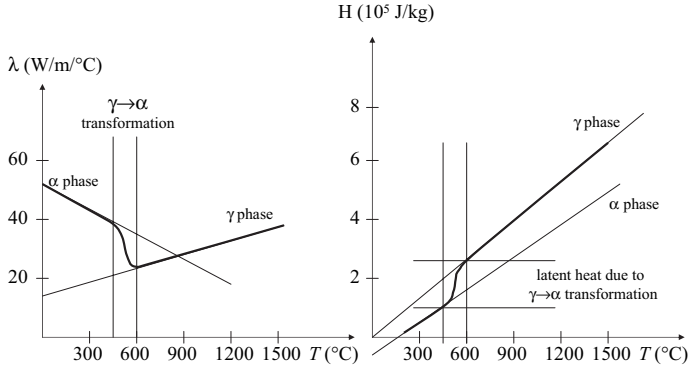


Figure 9.4. Typical variation of the thermal conductivity λ and the enthalpy H of phases α (ferrite) and phase γ (austenite) in low carbon-steel, according to temperature

9.2.1. Phase transformation diagrams

Metal and alloy transformation diagrams give the boundaries of the stability and phase change domains. Time is on the x-axis and temperature on the y-axis. There are two major types of metal and alloy phase transformation diagrams. IT diagrams (isothermal transformation diagram) are obtained by keeping a test tube at high temperature, called austenitizing temperature, for a certain time, called austenitizing time, before cooling it suddenly and keeping it at a constant temperature. The variation of austenite transformation is measured from the constant temperature. CCT diagrams are obtained with controlled thermal cycles. The test tube is kept at a high temperature, also called austenitizing temperature, for a certain time, also called austenitizing time, then continuously cooled down to ambient temperature. The variation of austenite transformation is measured during the cooling process.

Figures 9.5 and 9.6 give examples of IT and CCT diagrams obtained on steel with 0.35% carbon, 1% chromium, and 0.2% molybdenum. The austenitizing temperature is 850°C. The austenitizing time is 30 mn. In the IT diagram, there is a dotted line. It represents the temperature under which a martensitic transformation independent of time occurs. In the CCT diagram, note that, on each cooling curve, the percentage of transformed austenite is mentioned after each passage into a phase change domain. At the end of the cooling process, the test tube hardness is given.

The IT and CCT diagrams must be used to obtain the coefficients of phase change laws used during numerical simulation. For this purpose, setting methods were developed [PON 94], and some simple transformations can be set manually:

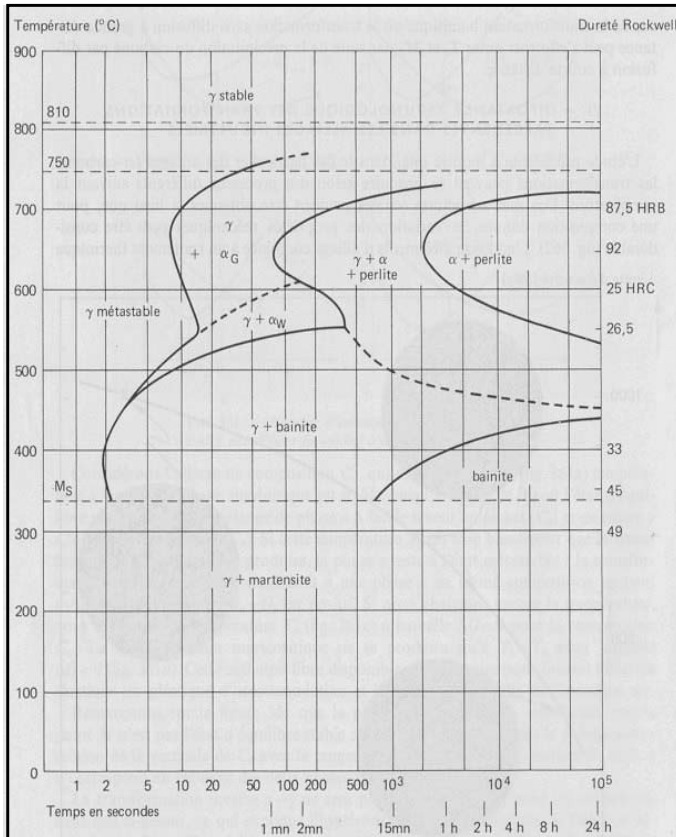


Figure 9.5. Example of IT diagram obtained on steel with 0.35% carbon, 1% chromium and 0.2% molybdenum, from [BÉN 84]

– martensitic transformation only requires the initial transformation temperature M_s and the kinetic parameter b (see equation [9.2]). The temperature M_s is directly read in the IT or CTT diagrams. For instance we obtain 335°C in Figure 9.5. The kinetic parameter b does not depend greatly on the steel chemical composition. A value of 0.011°C^{-1} is often used;

– with an IT diagram, the transformation kinetics given by equation [9.1] can be integrated at a constant temperature to obtain the classic Avrami relationship [AVR 40a, AVR 40b, AVR 41]:

$$p = P(1 - e^{-(ft)^n})$$

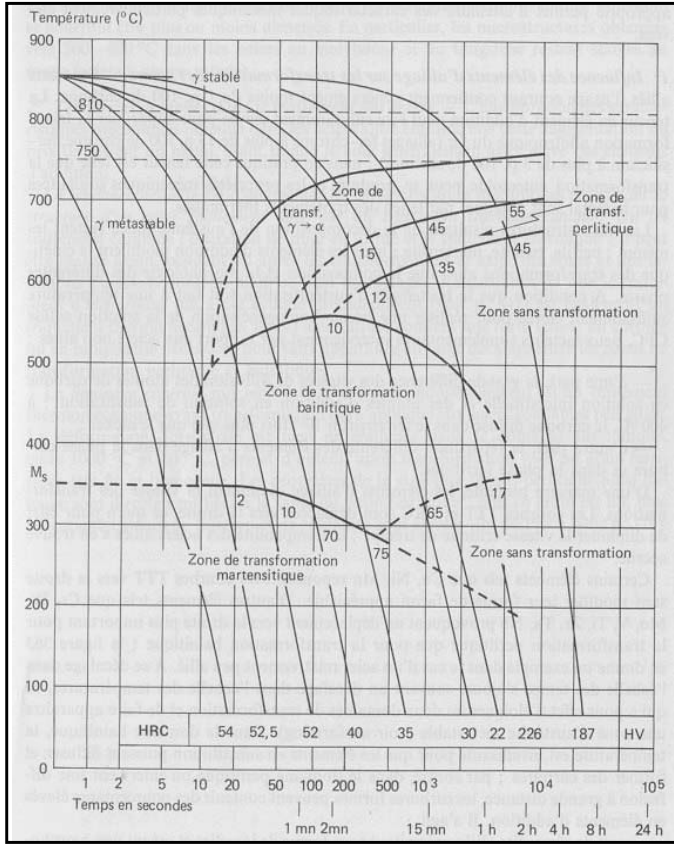


Figure 9.6. Example of CCT diagram obtained on steel with 0.35% carbon, 1% chromium and 0.2% molybdenum, from [BÉN 84]

In this equation, P , n and f are three parameters which can be determined at any temperature with the IT diagram. P is the phase proportion obtained after an infinite time, and this relationship gives:

$$\ln \left(\ln \frac{P}{P-p} \right) = n \ln f + n \ln t$$

It follows that by plotting at each temperature the quantity $\ln(\ln(P/(P-p)))$ according to $\ln(t)$, we obtain a straight line whose slope corresponds to the parameter n and the y-axis at the origin gives the parameter f .

Figure 9.5 illustrates how to obtain the perlite transformation parameters at different temperatures with an IT diagram. From equation [9.1] we find $P = 1$ and $n = 3$ at any temperature and a frequency f varying significantly according to the temperature, from 0 to 0.45 s^{-1} . The variation of this frequency is given by the right-hand diagram in Figure 9.5. The left-hand diagram in this figure gives the IT diagram used for the setting, and particularly the initial and final transformation times at temperatures 400°C , 500°C , 600°C and 700°C .

It is also possible to obtain the parameters of the transformation described in Figure 9.6 from the CCT diagram of steel. For this purpose, it is possible for instance to assume a particular form of the frequency f of the transformation according to the temperature. Figure 9.6 suggests the choice of a Gaussian function:

$$f = \begin{cases} f_0 e^{-((T-T_0)/\Delta T_i)^2} & \text{if } T \leq T_0 \\ f_0 e^{-((T-T_0)/\Delta T_s)^2} & \text{if } T \geq T_0 \end{cases}$$

In this function, the parameters f_0 , T_0 , ΔT_i and ΔT_s must be determined. Then, assuming P and n to be constant, a setting algorithm based on the Levenberg-Marquardt method [LEV 44] is applied.

Figures 9.7 and 9.8 give the results obtained by perlitic transformation in steel with 0.8% carbon. The left-hand CCT diagram is the experimental diagram, with initial and final transformation times, and the perlite quantity formed, for four cooling velocities, 80°C/s , 50°C/s , 8°C/s and 0.8°C/s . The transformation parameters obtained are $P=1$, $f_0 = 0.347 \text{ s}^{-1}$, $T_0 = 593^\circ\text{C}$, $\Delta T_i = 122^\circ\text{C}$, $\Delta T_s = 67^\circ\text{C}$ and $n = 2.78$. The right-hand CTT diagram was recalculated with these parameters. Note that it is fairly close to the experimental diagram, represented by the dotted lines.

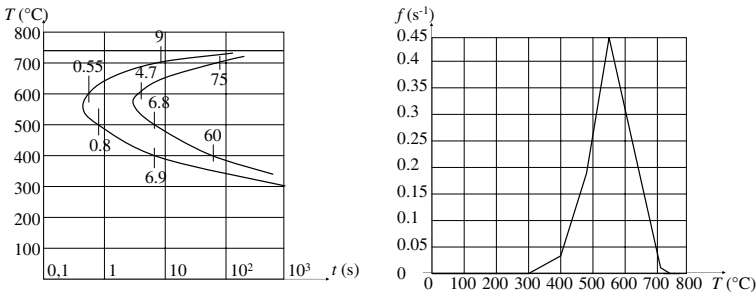


Figure 9.7. Obtaining perlite transformation parameters in steel with 0.8% carbon from an IT diagram, from [PON 94]

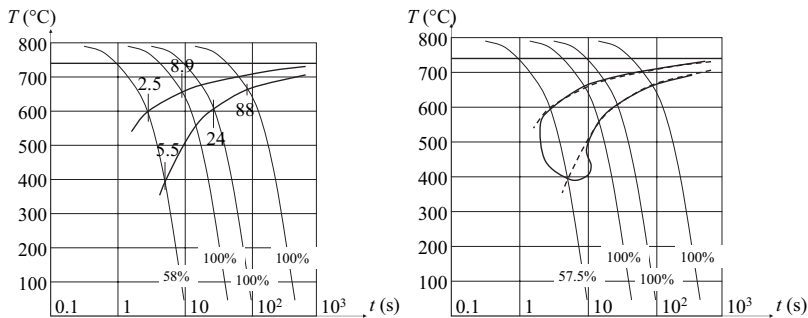


Figure 9.8. Obtaining perlitic transformation parameters in steel with 0.8% carbon from a CTT diagram, from [PON 94]

9.2.2. Steel quenching

This section examines an example of steel distortion tracking during cooling [CLA 02]. Figure 9.9 gives the geometry of the piece used. The cylindrical symmetry makes it possible to perform calculations on a meridian section of the piece in the

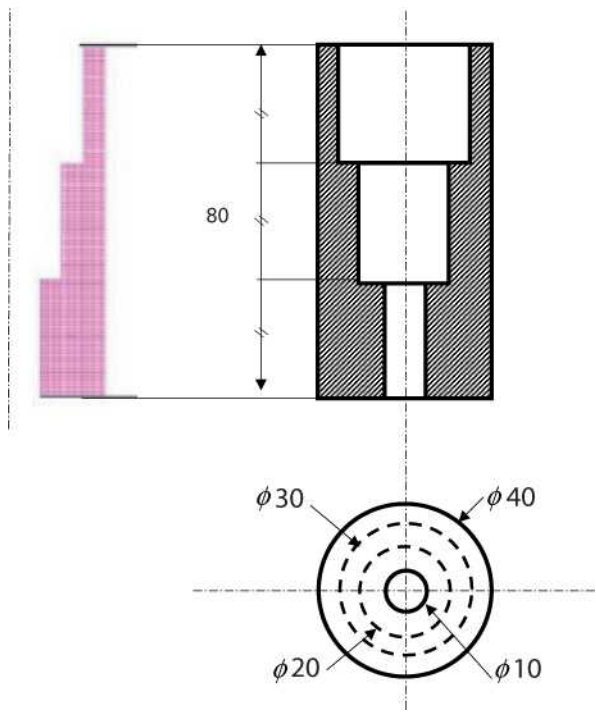


Figure 9.9. Geometry and mesh used for tracking distortions in steel quenching (sizes in mm)

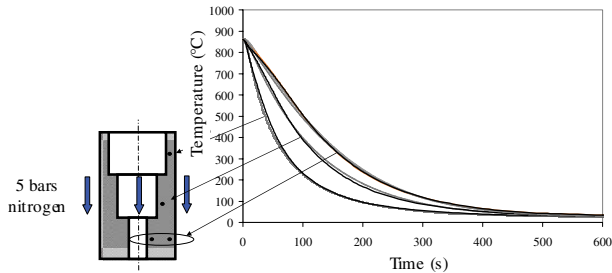


Figure 9.10. *Thermocouple position and theoretical and experimental variation of the local temperatures in the cylinder during the cooling process*

axisymmetric calculation configuration. Three bores with increasing diameters were regularly distributed on the cylinder height. They enable us to create different local thermal kinetics, and hence generate local distortion variations.

The cylinder of Figure 9.9 undergoes gas quenching with nitrogen injected vertically at a pressure of 5 bars. Figure 9.10 illustrates the quenching and the different thermal kinetics in agreement with the bores in the cylinder. Four temperatures are measured by thermocouple ($T1$ to $T4$). They were used to set the thermal exchange coefficient representing the heat extracted by the gas moving along the sample. The results of numerical simulations are also given in this figure. They were carried out on austenitic stainless steel, with no phase change during cooling. Note the good calculation-experiment agreement with an error always less than 10°C .

When the piece is cooling down, a camera is used to film its shape variation. It is thus possible to plot the shape variation of the austenitic stainless steel cylinder during the gas quenching operation (Figure 9.11). Note once again that experiment and calculations are in good agreement. This is mainly due to the good representation of the local thermal paths and the absence of phase changes.

In order to test the influence of phase changes, experiments and simulations were performed on steel 30CrNiMo8 with phase changes. The same heat exchange coefficient as for stainless steel was used. The thermal paths followed at the top and bottom of the piece were plotted on the CCT diagram of that steel. Figure 9.12 gives these thermal paths and the microstructure obtained. Note that the only phase transformation taking place is the austenite-martensite transformation.

Figure 9.13 represents the shape variation of a 30CrNiMo8 steel cylinder during cooling. It is easy to note the influence of the phase change on the diameter. Indeed,

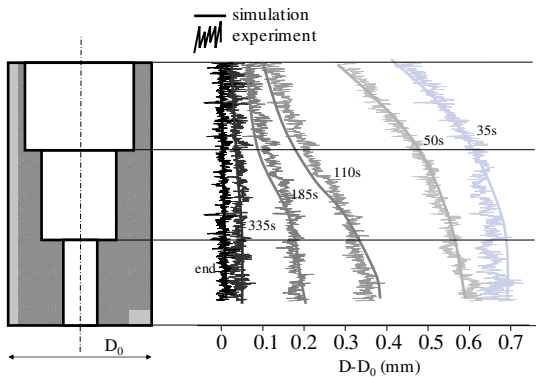


Figure 9.11. Shape variation of an austenitic stainless steel cylinder during cooling

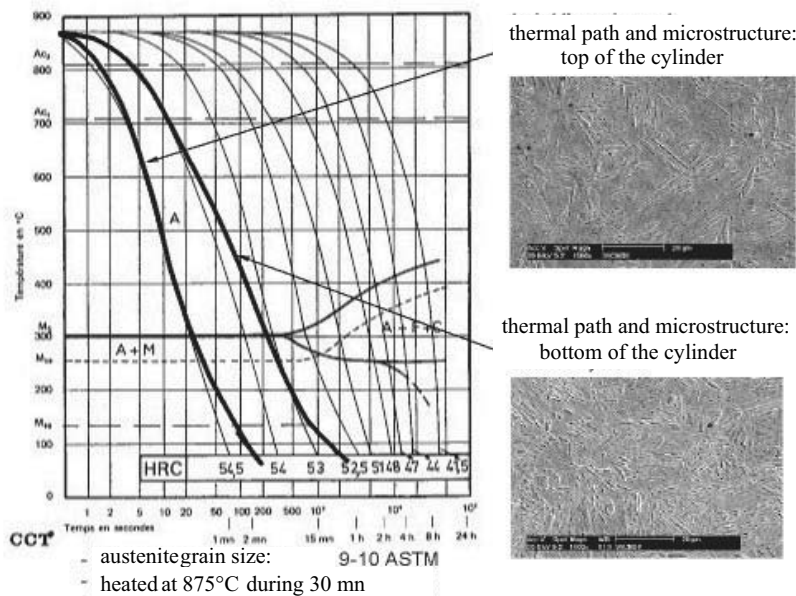


Figure 9.12. CCT diagram of steel 30CrNiMo8 – extremity thermal paths in the cylinder and corresponding final micro-structures

the structure of martensite being less compact than that of austenite, a local expansion occurs during the phase change, together with plastification called transformation plasticity [LEB 89b].

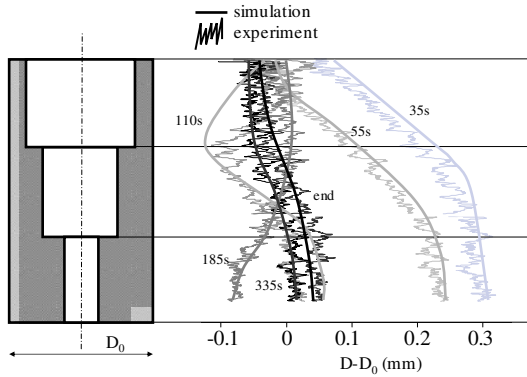


Figure 9.13. *Shape variation of a 30CrNiMo8 steel cylinder during cooling*

In Figure 9.13, experiment and simulation are in fairly good agreement. In particular, it is possible to make a good prediction of the final shape of the cylinder after quenching. If, without a phase change, it comes back to its initial shape when the thermal gradients are not sufficient to plasticize the material, when phase changes occur, the cylinder's initial shape is no longer rectilinear on the outside, even without any traditional plastification. This illustrates the presence of transformation plasticity. The plastic deformation of materials (traditional during phase changes) is not addressed in this book. The reader may refer to specialized books or articles [HIL 78, LEB 86, LEB 89b, LEB 89a, DEN 97, BES 01].

Chapter 10

Thermochemical Coupling

10.1. Finite element simulation of simultaneous diffusion and precipitation

Thermochemical treatments are widely used to improve the mechanical properties of metallic components. For example, in the case of steel, high temperature diffusion of carbon and/or nitrogen leads to an increased hardness near the surface, as well as to local compressive residual stresses. Both contribute to an increased component lifetime. Modeling diffusion can be considered relatively straightforward, since the governing equations are well-established, and are very similar to those described in the previous chapters of this book. However, we must account for the fact that during diffusion, precipitation may occur and modify the kinetics of diffusion. For example, in the case of steel, carbides and/or nitrides can form during the diffusion of carbon and/or nitrogen. This precipitation contributes to the improvement in the mechanical properties of the treated component.

The aim of this chapter is to describe a model aimed at predicting the local amount of chemical elements in a metallic component, during a thermochemical treatment. This model takes into account the diffusion and precipitation of these elements during the treatment, and has been incorporated into a commercial finite element computer code [FOR 95, FOR 00a, BER 03]. The application of the finite element method to simulate simultaneous diffusion and precipitation in metals is first recalled. Then we focus on the method used to incorporate precipitation into diffusion equations, and on the numerical scheme used to solve the global problem. Finally, numerical applications which illustrate the approach proposed are expounded.

10.1.1. Governing equations

We consider N chemical elements that can diffuse in a single matrix phase. For steels, the matrix consists of one major chemical element, iron, while alloying elements that can diffuse are mainly carbon, nitrogen, titanium or aluminum. These elements can combine together or with the matrix so as to form M precipitates. The chemical composition of the precipitates is characterized by stoichiometric coefficients $s_{i\alpha}$, each giving the number of atoms of element i contained in precipitate α . Throughout this chapter, Latin indices refer to chemical elements ($i = 1, \dots, N$), whereas Greek subscripts denote precipitates ($\alpha = 1, \dots, M$). The total fraction of element i , in all its forms (dissolved in the matrix or enclosed in a precipitate), is denoted by f_i^T , while the part dissolved in the matrix is f_i^D . If p_α denotes the fraction of precipitate α , the conservation of any element's atoms i can be expressed as:

$$f_i^T = f_i^D + \sum_{\alpha=1}^M s_{i\alpha} p_\alpha \quad [10.1]$$

It should be noted that we could choose to deal with mass fractions instead of molar fractions. In this case, we must introduce the atomic mass m_i of element i and the molecular mass m_α of precipitate α . Then the f_i^T , f_i^D and p_α in equation [10.1] and the sequel must be replaced by f_i^T/m_i , f_i^D/m_i and p_α/m_α respectively. The activity a_i of element i in the matrix is now introduced as a function of its dissolved fraction f_i^D as follows:

$$a_i = \gamma_i f_i^D \quad [10.2]$$

In this equation, γ_i denotes the thermodynamic coefficient of element i . In the case of a dilute solution, i.e. for small fractions f_i^D , a first order approximation of these coefficients can be used:

$$\ln(\gamma_i) = \ln(\gamma_i^0) + \sum_{j=1}^N e_{ij} f_j^D \quad [10.3]$$

In this equation, γ_i^0 is Henry's coefficient and the e_{ij} the Wagner interaction parameters. Henry's coefficient can be made temperature-dependent using an Arrhenius law. According to equations [10.2] and [10.3], thermodynamic coefficients can be calculated from given activities, by a Newton-like method, together with their partial derivatives with respect to these activities. Diffusion is assumed to take place only in the matrix (diffusion inside precipitates is neglected). According to Fick's first

law, the flux density $\vec{\phi}_i$ of element i is assumed to be proportional to the gradient of its activity:

$$\vec{\phi}_i = -\frac{D_i}{\gamma_i} \overrightarrow{\text{grad}}(a_i) \quad [10.4]$$

In this equation, D_i is the diffusion coefficient of element i inside the matrix. It depends on the temperature through an Arrhenius function. It should be noted that equation [10.4] does not imply that cross-diffusion is neglected. Since the thermodynamic coefficients γ_i depend on the activities (equation [10.3]), the flux density $\vec{\phi}_i$ can be expressed as a function of the dissolved fractions by:

$$\vec{\phi}_i = \sum_{j=1}^N -D_{ij} \overrightarrow{\text{grad}}(f_j^D) \quad \text{with} \quad D_{ij} = D_i(\delta_{ij} + f_i^D e_{ij}) \quad [10.5]$$

Thus, cross-diffusion is governed by the Wagner interaction coefficients e_{ij} introduced in equation [10.3]. According to the above equations, Fick's second law relates the change rate of the total fraction of each element i to its activity as follows:

$$\dot{f}_i^T = \text{div} \left(\frac{D_i}{\gamma_i} \overrightarrow{\text{grad}}(a_i) \right) \quad [10.6]$$

In order to integrate equation [10.6] in a given domain Ω , it is necessary to define boundary conditions (on $\partial\Omega$) for the activities. A general expression for these boundary conditions is for each element i :

$$\frac{D_i}{\gamma_i} \overrightarrow{\text{grad}}(a_i) \cdot \vec{n} = J_i + h_i(a_i^e - a_i) \quad [10.7]$$

In this equation, \vec{n} is the outward normal unit vector of the boundary, J_i is a prescribed flux density, a_i^e is an external activity and h_i is a film transfer coefficient. Using this equation, we can obtain the following boundary conditions:

- a prescribed flux density J_i , with $h_i = 0$;
- a prescribed activity a_i^e , with $J_i = 0$ and a very high value of h_i ;
- a film transfer with $J_i = 0$.

When N chemical elements are considered in Ω , equations [10.6] and [10.7] form a set of differential equations, with boundary conditions, which can be solved by using different methods. Among these methods, the finite element technique appears to be

the most suitable for complex shapes of the domain Ω , whereas a finite difference approximation is sufficient for 1D calculations (i.e. when diffusion is assumed to occur in only one direction). In this book, we focus on the finite element approach used to approximate the solution of equations [10.6] and [10.7], which appear to be very similar to those obtained in thermal analyses.

10.1.2. Finite element formulation

The finite element method applies to a weak formulation of the boundary value problem [10.6]-[10.7], which can be written for each chemical element. According to this formulation, it is necessary to find $a_i : t \in]0, T[\rightarrow a_i(x, y, z, t) \in V$ such that $a_i(\cdot, t = 0) = a_i^0$ and:

$$\begin{aligned} \forall \psi_i \in V, \quad & \int_{\partial\Omega} \psi_i (J_i + h_i(a_i^e - a_i)) ds \\ & - \int_{\Omega} \overrightarrow{\text{grad}}(\psi_i) \frac{D_i}{\gamma_i} \overrightarrow{\text{grad}}(a_i) dv - \int_{\Omega} \psi_i \dot{f}_i^T dv = 0 \end{aligned} \quad [10.8]$$

In this formulation, $]0, T[$ is a given time interval, a_i^0 is an initial condition and V is a suitable functional space containing the test functions ψ_i . In order to build a finite element approximation for a_i , the domain Ω is divided into elements Ω^e connected by nodes. In each element Ω^e , a spatial approximation of a_i and its gradient is used:

$$\forall \vec{x} \in \Omega^e, \quad \begin{cases} a_i(\vec{x}) = \sum_p N_p^e(\vec{x}) a_{ip} \\ \overrightarrow{\text{grad}}(a_i(\vec{x})) = \sum_p \overrightarrow{\text{grad}}(N_p^e(\vec{x})) a_{ip} \end{cases} \quad [10.9]$$

In this equation, a_{ip} is the activity of element i at node p of element Ω^e , and $N_p^e(\vec{x})$ is the shape function associated with node p of element Ω^e . Equation [10.9] can also be applied to the test functions ψ_i . This leads to an approximation of the functional space. Then, by applying equation [10.8], we must find $\{a\} : t \in]0, T[\rightarrow \{a(t)\}$ such that:

$$\{R(a(t), t)\} = \sum_{e=1}^m [A^e] \cdot \{R^e(a(t), t)\} = \{0\} \quad [10.10]$$

At any instant t , the residual vector $\{R(a(t), t)\}$ is a function of the unknown vector $\{a(t)\}$. Both have $N \times n$ components, where n is the number of nodes of the

whole structure. $\{R(a(t), t)\}$ is obtained by assembling all element residual vectors $\{R^e(a(t), t)\}$, the component of which reads, for chemical element i and node p :

$$R_{ip}^e = \int_{\partial\Omega^e \cap \partial\Omega} N_p^e (J_i + h_i(a_i^e - a_i)) ds - \int_{\Omega^e} \overrightarrow{\text{grad}}(N_p^e) \frac{D_i}{\gamma_i} \overrightarrow{\text{grad}}(a_i) dv - \int_{\Omega^e} N_p^e \dot{f}_i^T dv \quad [10.11]$$

It can be seen in equation [10.11] that the element residual vector $\{R^e\}$ involves the time derivative of the total fraction of each chemical element. We approximate this time derivative using the generalized trapezoidal rule. Therefore, it is assumed that for each interval of integration $[t, t + \Delta t]$:

$$\frac{f_i^T(t + \Delta t) - f_i^T(t)}{\Delta t} = (1 - \nu) \dot{f}_i^T(t) + \nu \dot{f}_i^T(t + \Delta t) \quad \text{with} \quad 0 \leq \nu \leq 1 \quad [10.12]$$

A value $\nu = 1/2$ in equation [10.12] corresponds to the trapezoidal rule, which is the most accurate algorithm. Nevertheless, in the case of discontinuous time derivatives, the implicit Euler scheme ($\nu = 1$) is generally preferred. According to this equation, the component of an element residual vector at node p , for chemical element i , and at time $t + \Delta t$, reads:

$$R_{ip}^e = \int_{\partial\Omega^e \cap \partial\Omega} N_p^e (J_i + h_i(a_i^e - a_i)) ds - \int_{\Omega^e} \overrightarrow{\text{grad}}(N_p^e) \frac{D_i}{\gamma_i} \overrightarrow{\text{grad}}(a_i) dv - \int_{\Omega^e} N_p^e \left(\frac{1}{\nu \Delta t} (f_i^T(t + \Delta t) - f_i^T(t)) - \frac{1 - \nu}{\nu} \dot{f}_i^T(t) \right) dv \quad [10.13]$$

Starting from known quantities at time t , the activities of all elements at all the nodes of the structure, at time $t + \Delta t$, result simultaneously from the solution of equation [10.10]. This solution is found by means of an iterative Newton procedure, which involves a global tangent matrix $[K]$ obtained by assembling tangent element matrices $[K^e]$. The components of these tangent element matrices are:

$$K_{ipjq}^e = - \frac{\partial R_{ip}^e}{\partial a_{jq}} = - \int_{\Omega^e} \frac{D_i}{\gamma_i^2} \frac{\partial \gamma_i}{\partial a_j} N_q^e \overrightarrow{\text{grad}}(N_p^e) \overrightarrow{\text{grad}}(a_i) dv + \int_{\Omega^e} \frac{D_i}{\gamma_i} \delta_{ij} \overrightarrow{\text{grad}}(N_p^e) \overrightarrow{\text{grad}}(N_q^e) dv$$

$$\begin{aligned}
& + \int_{\Omega^e} \frac{1}{\nu \Delta t} \frac{\partial f_i^T}{\partial a_j} N_p^e N_q^e dv \\
& + \int_{\partial\Omega \cap \partial\Omega^e} h_i \delta_{ij} N_p^e N_q^e ds
\end{aligned} \tag{10.14}$$

In this equation, the partial derivatives of the thermodynamic coefficient γ_i and the total fraction f_i^T of element i (at time $t + \Delta t$), with respect to the activity a_j of element j , have to be estimated. The former is derived from equation [10.3], whereas the latter is approximated by neglecting the variation of the precipitate fraction with respect to the activities. According to equations [10.1] and [10.2], this gives:

$$\frac{\partial f_i^T}{\partial a_j} \approx \frac{1}{\gamma_i} \left(\delta_{ij} - f_i^D \frac{\partial \gamma_i}{\partial a_j} \right) \tag{10.15}$$

The main problem when solving the above equations is to estimate the total fraction of a chemical element as a function of its dissolved fraction (or its activity). This is done using equation [10.1], which involves the fractions of precipitates p_α . The aim of the next section is to describe the procedure used to incorporate precipitation into this equation.

10.2. Calculation of precipitation

10.2.1. Mathematical formulation

Starting from instant t , and thus assuming the fractions of precipitates at this time to be known, we want to estimate these fractions at time $t + \Delta t$. For this purpose, we first calculate an “equilibrium fraction” of precipitate, P_α , that would form if no diffusion were present for a large period of time. Then we use a kinetic law, which represents the time-dependent evolution of the precipitate fraction from its value at time t , to its value at time $t + \Delta t$, towards the equilibrium value P_α .

The equilibrium fractions P_α are calculated by applying the law of mass action, which can be formulated as follows:

$$\begin{cases} P_\alpha = 0 & \text{if } \prod_{i=1}^N A_i^{s_{i\alpha}} < K_\alpha \\ P_\alpha \geq 0 & \text{if } \prod_{i=1}^N A_i^{s_{i\alpha}} = K_\alpha \end{cases} \tag{10.16}$$

According to this equation, precipitation occurs only when the solubility product K_α is reached by the “equilibrium activities” of the elements A_i . These activities are obtained from the total fraction f_i^T of element i by the following relationship:

$$A_i = \Gamma_i F_i^D \quad \text{with} \quad \begin{cases} F_i^D = f_i^T - \sum_{\alpha=1}^M s_{i\alpha} P_\alpha \\ \Gamma_i = \gamma_i^0 e^{\sum_{j=1}^N e_{ij} F_j^D} \end{cases} \quad [10.17]$$

The solubility products K_α and the stoichiometric coefficients $s_{i\alpha}$ in equation [10.16] are given as characteristics of the precipitates. The solubility products can be made temperature-dependent using the Arrhenius law. The Γ_i and F_i^D terms appearing in equation [10.17] are respectively the thermodynamic coefficient and the dissolved fraction of element i , both expressed in equilibrium, i.e. without diffusion during a large period of time. For convenience, the law of mass action [10.16] can be formulated for all the precipitates that may be produced in the material as:

$$\begin{cases} P_\alpha \geq 0 \\ Q_\alpha = \ln(K_\alpha) - \sum_{i=1}^N \left[\ln(\gamma_i^0) + \sum_{j=1}^N e_{ij} \left(f_j^T - \sum_{\beta=1}^M s_{j\beta} P_\beta \right) + \ln \left(f_i^T - \sum_{\beta=1}^M s_{i\beta} P_\beta \right) \right] \geq 0 \\ \sum_{\alpha=1}^M Q_\alpha P_\alpha = 0 \end{cases} \quad [10.18]$$

Equation [10.18] must be solved for a given set of total fractions f_i^T , and for given material characteristics γ_i^0 , e_{ij} , $s_{i\alpha}$ and K_α . It turns out that in these equations Q_α is a non-linear function of the P_α . Equation [10.18] can thus be considered as defining a non-linear complementarity problem [KAR 72, LEM 80]. Starting with the P_α values calculated at time t , and from $P_\alpha = 0$ ($\alpha = 1, \dots, M$) at $t = 0$, the following steps are used to solve this problem at time $t + \Delta t$:

- $Q_\alpha(P_1, \dots, P_M) = 0$ is solved for each precipitate for which $P_\alpha \neq 0$. This leads to a new set of P_α values. A Newton method is used, which verifies the condition that the F_i^D terms must remain positive.

- The $Q_\alpha(P_1, \dots, P_M)$ are calculated for each precipitate for which $P_\alpha = 0$. This leads to a new set of Q_α values.

– If all the P_α and Q_α values obtained from the previous steps are non-negative, then the solution is reached, otherwise (i) the negative P_α are set to zero, (ii) the P_α corresponding to a negative Q_α are set to a non-vanishing value, and we go back to the first step.

The kinetic law introduced to account for the time-dependence of precipitation is based on a Johnson-Mehl-Avrami-Kolmogorov law, so that the precipitate fraction is given by:

$$p_\alpha = P_\alpha (1 - e^{-(ft)^n}) \quad [10.19]$$

In this equation, n and f are two parameters which characterize this kinetics. Thus, knowing $p_\alpha(t)$, the precipitate fraction at time $t + \Delta t$ is:

$$p_\alpha(t + \Delta t) = \begin{cases} P_\alpha (1 - e^{-(f(t_0 + \delta t))^n}) \\ \text{with } t_0 = \frac{1}{f} \left[\ln \left(\frac{P_\alpha}{P_\alpha - p_\alpha(t)} \right) \right]^{1/n} & \text{if } p_\alpha(t) < P_\alpha \\ p_\alpha(t) - (p_\alpha(t) - P_\alpha)(1 - e^{-(ft)^n}) & \text{if } p_\alpha(t) \geq P_\alpha \end{cases} \quad [10.20]$$

It should be noted that when the calculated equilibrium value $P_\alpha(t)$ is less than the current value $p_\alpha(t)$, the precipitate is currently being dissolved in the matrix and the exponent n of the kinetics is set to 1 (no nucleation period).

10.2.2. Numerical scheme

At any instant $t + \Delta t$, iterations are performed using a Newton-like method to determine the local activity a_i of each chemical element i in the domain. If $x^{(k)}$ denotes the value of any variable x obtained at the end of the k^{th} iteration, then the following scheme is used at each iteration (k) to incorporate precipitation into the diffusion model:

– The total fraction of chemical element i , $f_i^{T(k)}$, is calculated by means of equations [10.1] and [10.2] with the current values of the activities $a_i^{(k)}$ and thermodynamic coefficients $\gamma_i^{(k)}$, and with the precipitate fractions obtained at the end of the previous iteration $p_\alpha^{(k-1)}$ (for the first iteration, the precipitate fractions are taken from the previous time step).

– The equilibrium fractions of precipitates, $P_\alpha^{(k)}$, are calculated by means of equation [10.18] with the total fractions $f_i^{T(k)}$. Kinetics laws [10.20] then give the precipitate fractions $p_\alpha^{(k)}$.

– The total fraction of chemical element i , $f_i^{T(k)}$, is modified according to equation [10.1], with the new precipitate fractions $p_\alpha^{(k)}$, before estimating the residual vector of equation [10.10] and eventually going back to a new iteration.

According to the scheme presented above, convergence of the global Newton method cannot be achieved if the precipitate fractions calculated present a large variation between two successive iterations. Thus precipitation is treated numerically by a so-called fixed-point method.

10.3. Examples

10.3.1. Calculation of a phase diagram

In order to validate the method used to estimate the equilibrium fractions of precipitates P_α (equation [10.18]), the phase diagram of a micro-alloyed steel containing Ti , Mn , S and C as alloying elements can be estimated. Within this steel, the precipitates that may appear and form phases at 1, 200°C are TiS , TiC , $Ti_4C_2S_2$ and MnS . Table 10.1 gives the data used for the chemical elements, whereas Table 10.2 gives those associated with the precipitates.

Element	Ti	Mn	S	C
Atomic mass	47.900	54.938	32.064	12.011
Henry's coefficient	1.	1.	1.	1.
Wagner parameters	0.	0.	0.	0.
Total mass fraction (%)	0 to 0.1	0.2	0 to 0.01	0.003

Table 10.1. Data used for the chemical elements involved in the phase diagram

Precipitate	TiS	TiC	$Ti_4C_2S_2$	MnS
Molecular mass	79.964	59.911	279.75	87.002
Solubility product	4.835310^{-5}	5.355110^{-3}	3.9510^{-17}	6.403910^{-4}

Table 10.2. Characteristics of the precipitates involved in the phase diagram

Figure 10.1 gives the phase diagram calculated from the data in Tables 10.1 and 10.2. This diagram was obtained by solving equations [10.18] at various points, and by detecting the domains where precipitates appear. It can be seen in this figure that

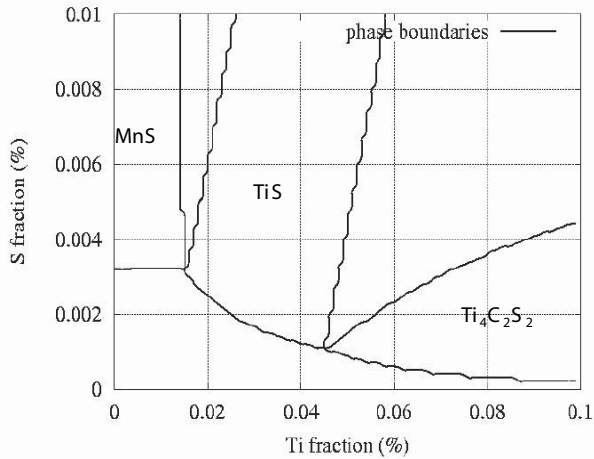


Figure 10.1. *Recalculated phase diagram of a micro-alloyed steel*

the TiC precipitate does not appear within the variation range used for the chemical elements. Moreover, the number of precipitates meets the Gibbs law. Finally, this diagram is in good agreement with the experiment. This validates equations [10.18], together with the numerical scheme used to solve them.

10.3.2. Carbon diffusion in a titanium steel

In order to validate the method used to model simultaneous diffusion and precipitation, we can calculate carbon diffusion in a steel containing titanium. The data used for the calculation are given in Table 10.3. They apply to a cylinder with a 10 mm radius. Diffusion of carbon and titanium is assumed to take place only in the radial direction, and the possibility of the appearance of a TiC precipitate is taken into account. It can be seen in this table that the calculations were performed with the following assumptions:

- Henry's coefficients of the elements are set to 1, and Wagner's coefficients vanish. This means that we assume that the activities of the elements and their dissolved fractions are equal. Moreover, it is assumed that the elements do not interact during diffusion.

- The initial mass fractions of carbon and titanium are set to 0.2% and 0.01% respectively. Then the steel specimen is placed into a furnace. The carbon activity in the atmosphere of the furnace is set to 1, and a film transfer coefficient of 0.0001 is used to represent the carbon flux density entering into the specimen during the treatment.

Element	<i>C</i>	<i>Ti</i>	Precipitate	<i>TiC</i>
Atomic mass	12.011	47.900	Molecular mass	59,911
Wagner parameters	0.	0.	Frequency factor (s^{-1})	0.1
Initial activity (%)	0.2	0.01	Exponent	1.
Film transfer coefficient	0.0001	0.		
Activity in the furnace (%)	1.	0.		
Diffusion coefficient (mm^2/s)	4.8710^{-7}	510^{-7}		

Table 10.3. Data used for the simulation of carbon diffusion in a titanium steel

– A kinetic law with a frequency factor of $0.1\ s^{-1}$ and an exponent of 1 is used for the *TiC* precipitate.

Figure 10.2 gives the distribution of the total, dissolved and precipitated, fractions of *Ti* along the radius, near the surface of the cylinder. It can be seen in this figure that carbon diffusion has led to the precipitation of *TiC* near the surface. Since the precipitate is made of *Ti* and *C*, the dissolved fraction of *Ti* is lowered, leading to a gradient of its activity, and thus to diffusion of this element. This diffusion then leads to an increase in the fraction of the *TiC* precipitate, so that in this case precipitation is governed by the mobility of titanium in the steel matrix, as well as by the kinetics of precipitation.

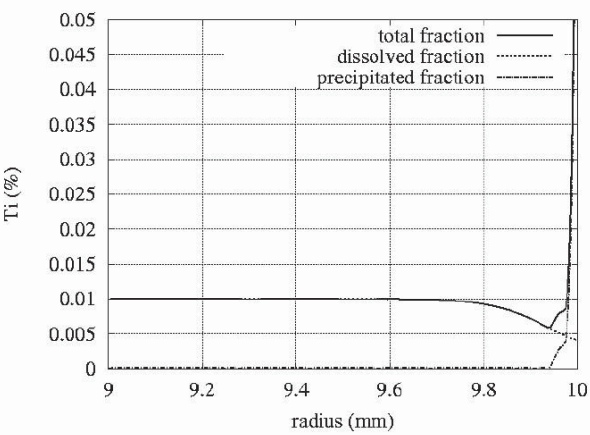


Figure 10.2. Variation of the fraction of *Ti* along the cylinder radius, after 4 hours' treatment

This page intentionally left blank

Chapter 11

Electrothermal Coupling

11.1. Electrokinetic modeling

Here we deal with the case of a high-intensity current applied to a material, causing the heating of that material by means of the Joule effect, thus coupling it with thermal aspects. In resistance welding applications, given the current frequency ($50Hz$), the dynamic effects (induced currents) are negligible and the electric currents can be calculated by means of an electrokinetic model.

Here we give the equations governing the electrokinetic phenomena involved in the material, their resolution by the finite element method and their coupling with thermal phenomena.

11.1.1. Weak formulation

In a domain Ω , applying an electric field \vec{E} produces an electric current whose density termed \vec{J} is given by the Ohm law, $\vec{J} = \underline{\sigma} \cdot \vec{E}$, where the symmetric tensor $\underline{\sigma}$ represents the electric conductivity of the material. When this conductivity is identical in all directions, then $\underline{\sigma} = \sigma \underline{I}$, where \underline{I} is the identity tensor and the scalar σ represents this conductivity. In electrokinetics, the electric field \vec{E} is such that $\text{rot}(\vec{E}) = \vec{0}$. It is therefore derived from a scalar potential V , for instance of the form $\vec{E} = -\overrightarrow{\text{grad}}(V)$. The current density in the material is thus proportional to the electric potential gradient:

$$\vec{J} = -\underline{\sigma}.\overrightarrow{\text{grad}}(V)$$

This equation is the material's electrokinetic behavior law. It is similar to Fourier's law of thermal transfer. If the current conservation inside any domain Ω is then written, without any possible volume current production, it is simple to determine the electrokinetic conservation equation:

$$\operatorname{div}(\vec{J}) = 0 \quad \text{at any point of } \Omega$$

Using these two equations, and applying current density or potential boundary conditions produces the following system to solve, which is very similar to the one encountered in steady state heating in Chapter 1:

$$\text{Find } V(\vec{x}) \text{ such that: } \begin{cases} \operatorname{div}(\underline{\sigma} \cdot \overrightarrow{\operatorname{grad}}(V)) = 0 & \text{for any } \vec{x} \in \Omega \\ V = V_d & \text{on } \partial\Omega_V \\ -\vec{J} \cdot \vec{n} = (\underline{\sigma} \cdot \overrightarrow{\operatorname{grad}}(V)) \cdot \vec{n} = j & \text{on } \partial\Omega_j \end{cases} \quad [11.1]$$

We define \vec{n} as the outward norm to the domain Ω , $\partial\Omega_V$ as the boundary of Ω on which the potential V is known, and $\partial\Omega_j$ as the boundary of Ω where the entry current density is known. Note that, as in thermal science, it is necessary to know at any point of the boundary $\partial\Omega$ of the domain Ω either the exiting current potential or density, so that $\partial\Omega_V \cup \partial\Omega_j = \partial\Omega$ and $\partial\Omega_V \cap \partial\Omega_j = \emptyset$. If the same procedure as that described in Chapter 1 of this book for steady state heating is followed, it is possible to give a weak formulation of problem [11.1]:

$$\left| \begin{array}{l} \text{Find } V \in E_V \text{ such that for any } V^* \in E_{V^*}: \\ \int_{\partial\Omega_j} V^* j \, ds - \int_{\Omega} \overrightarrow{\operatorname{grad}}^T(V^*) \cdot \underline{\sigma} \cdot \overrightarrow{\operatorname{grad}}(V) \, dv = 0 \end{array} \right. \quad [11.2]$$

We insert into this system the set E_V of solutions satisfying the essential boundary conditions $V = V_d$ on $\partial\Omega_V$, and the set E_{V^*} of test functions V^* equal to zero on that same boundary.

11.1.2. Modeling the coupling

The boundary value problem governing electrothermal coupling in a domain Ω consists of juxtaposing the boundary value problem related to heat diffusion with the problem related to electrokinetic phenomena. Both problems are coupled by:

- the power dissipated by the Joule effect, incorporated as an internal source term in the heat equation;
- the temperature dependence of electrical properties.

The power dissipated by the Joule effect per material volume unit is of the form $P = \vec{E} \cdot \vec{J}$. It follows that the volumetric source term Q involved in the heat problem can be written directly according to the potential field $V(\vec{x})$, the solution of the electrokinetic problem:

$$Q = \overrightarrow{\text{grad}}^T(V) \cdot \underline{\sigma} \cdot \overrightarrow{\text{grad}}(V) \quad [11.3]$$

The temperature field $T(\vec{x})$ and the electric potential field $V(\vec{x})$ are therefore obtained by coupling the two related weak formulations. The thermal calculation being performed in the transient state, and hence with a time-dependent temperature field, the electrokinetic potential will also be time-dependent. The temperature field $T(\vec{x}, t)$ and the potential field $V(\vec{x}, t)$ will be obtained at any instant t between 0 and t_f by solving the following problem:

$$\left| \begin{array}{l} \text{Find } T: t \in [0, t_f[\longrightarrow T(\vec{x}, t) \in E_T, \\ \text{Find } V: t \in [0, t_f[\longrightarrow V(\vec{x}, t) \in E_V \text{ such that:} \\ \text{at } t = 0: T(\vec{x}, 0) = T_0(\vec{x}) \text{ and } V(\vec{x}, 0) = V_0(\vec{x}) \\ \text{for } t \in]0, t_f[, \text{ for any } \psi \in E_\psi \text{ and for any } V^* \in E_{V^*}: \\ \int_{\Omega} \psi(Q - \rho \dot{H}) dv + \int_{\partial\Omega_q} \psi q ds - \int_{\Omega} \overrightarrow{\text{grad}}^T(\psi) \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(T) dv = 0 \\ \int_{\partial\Omega_j} V^* j ds - \int_{\Omega} \overrightarrow{\text{grad}}^T(V^*) \cdot \underline{\sigma} \cdot \overrightarrow{\text{grad}}(V) dv = 0 \end{array} \right. \quad [11.4]$$

Over time, the natural boundary conditions provide the heat flux density q entering via $\partial\Omega_q$ and the current density j entering via $\partial\Omega_j$. Similarly, the essential boundary conditions are those making it possible to define the sets E_T and E_V according to the temperatures T_d and the potentials V_d prescribed, as well as the sets E_ψ and E_{V^*} of the weighting functions:

$$E_T = \{T \in H^1(\Omega) / T = T_d \text{ on } \partial\Omega_T\}$$

$$E_\psi = \{\psi \in H^1(\Omega) / \psi = 0 \text{ on } \partial\Omega_T\}$$

$$E_V = \{V \in H^1(\Omega) / V = V_d \text{ on } \partial\Omega_V\}$$

$$E_{V^*} = \{V^* \in H^1(\Omega) / V^* = 0 \text{ on } \partial\Omega_V\}$$

In problem [11.4], it must be noted that:

- the volumetric mass ρ , the thermal conductivity (symmetric) tensor $\underline{\lambda}$ and the electrical conductivity (symmetric) tensor $\underline{\sigma}$ depend on temperature T ;
- the enthalpy H is a function of temperature and time. This term also makes it possible to insert latent heat effects related to phase changes (see Chapter 9);
- the volumetric source term Q is given by relationship [11.3]. It represents the power dissipated by the Joule effect and depends at each instant on the electric potential field;
- as was seen in the previous parts of this book, the heat flux density q can take different forms to express a convection or radiation exchange with the external environment;
- the current density j can be expressed from the electric resistance R of an interface, possibly a function of the temperature, with a medium of potential V_∞ in the form $J = (1/R)(V_\infty - V)$.

11.1.3. Solving the coupled problem

In order to solve coupled problem [11.4], we could use a staggered approach where the electrokinetic and thermal problems would be treated independently. However, the importance of the coupling between such phenomena implies us to prefer an approach where the problems are treated together. To obtain the finite element formulation of this problem, a Galerkin approach is used, with the temperatures $T(\vec{x}, t)$ and potentials $V(\vec{x}, t)$ as unknown fields.

Consider a single mesh for the electrokinetic and thermal problems, including n nodes and m elements Ω^e . Inside each element, the temperature and the electric potential use the same approximation scheme:

$$\forall \vec{x} \in \Omega^e, \forall t \in [0, t_f], \quad \begin{cases} T(\vec{x}, t) = \langle N^e(\vec{x}) \rangle \cdot \{T^e(t)\} \\ V(\vec{x}, t) = \langle N^e(\vec{x}) \rangle \cdot \{V^e(t)\} \end{cases} \quad [11.5]$$

The problem now has $2n$ unknowns: the temperature and the electric potential at all the mesh nodes. In this instance an implicit Euler algorithm to discretize the time derivative of the enthalpy H of [11.4]. The finite element method then leads us to solve at each time step a system of $2n$ equations with $2n$ unknowns, which can be written as:

$$\begin{Bmatrix} \{R_T\} \\ \{R_V\} \end{Bmatrix} = \begin{Bmatrix} \{0\} \\ \{0\} \end{Bmatrix} \quad [11.6]$$

In the above equation, $\{R_T\}$ represents the residual vector at instant t related to the thermal problem and $\{R_V\}$ the vector related to the electrokinetic problem. They are both obtained by assembling element residuals:

$$\{R_T\} = \sum_{e=1}^m [A^e]^T \cdot \{R_T^e\}$$

$$\{R_V\} = \sum_{e=1}^m [A^e]^T \cdot \{R_V^e\}$$

The element residuals are easily obtained from the results of previous chapters, they are written:

$$\begin{aligned} \{R_T^e\} = & \int_{\Omega^e} \{N^e\} \overrightarrow{\text{grad}}^T(V) \cdot \underline{\sigma} \cdot \overrightarrow{\text{grad}}(V) dv + \int_{\partial\Omega^e \cap \partial\Omega_q} \{N^e\} q ds \\ & - \int_{\Omega^e} \{N^e\} \rho \frac{H - H_0}{\Delta t} dv - \int_{\Omega^e} \left\{ \overrightarrow{\text{grad}}^T(N^e) \right\} \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(T) dv \end{aligned} \quad [11.7]$$

$$\{R_V^e\} = \int_{\partial\Omega^e \cap \partial\Omega_j} \{N^e\} j ds - \int_{\Omega^e} \left\{ \overrightarrow{\text{grad}}^T(N^e) \right\} \cdot \underline{\sigma} \cdot \overrightarrow{\text{grad}}(V) dv \quad [11.8]$$

The resolution of system [11.6] can be performed with a Newton-Raphston method. This method involves a tangent matrix, where the different sub-matrices are obtained by assembling element matrices, as follows:

$$\{K_T\} = \begin{bmatrix} [K_{TT}] & [K_{TV}] \\ [K_{VT}] & [K_{VV}] \end{bmatrix} \quad \text{with} \quad \left\{ \begin{aligned} [K_{TT}] &= \sum_{e=1}^m [A^e]^T \cdot [K_{TT}^e] \cdot [A^e] \\ [K_{TV}] &= \sum_{e=1}^m [A^e]^T \cdot [K_{TV}^e] \cdot [A^e] \\ [K_{VT}] &= \sum_{e=1}^m [A^e]^T \cdot [K_{VT}^e] \cdot [A^e] \\ [K_{VV}] &= \sum_{e=1}^m [A^e]^T \cdot [K_{VV}^e] \cdot [A^e] \end{aligned} \right.$$

Note that the global matrix $[K_T]$ is non-symmetric. This is practically always the case when several coupled problems are treated together. The components of the element matrices from the previous equations are written:

$$\begin{aligned}
(K_{TT}^e)_{ij} &= - \int_{\Omega^e} N_i^e \overrightarrow{\text{grad}}^T(V) \cdot \frac{d\sigma}{dT} \cdot \overrightarrow{\text{grad}}(V) N_j^e dv - \int_{\partial\Omega^e \cap \partial\Omega_q} N_i^e \frac{dq}{dT} N_j^e ds \\
&\quad + \int_{\Omega^e} N_i^e \frac{1}{\Delta t} \frac{d(\rho H)}{dT} N_j^e dv + \int_{\Omega^e} \overrightarrow{\text{grad}}^T(N_i^e) \cdot \frac{d\lambda}{dT} \cdot \overrightarrow{\text{grad}}(T) N_j^e dv \\
&\quad + \int_{\Omega^e} \overrightarrow{\text{grad}}^T(N_i^e) \cdot \underline{\lambda} \cdot \overrightarrow{\text{grad}}(N_j^e) dv \\
(K_{TV}^e)_{ij} &= -2 \int_{\Omega^e} N_i^e \overrightarrow{\text{grad}}^T(V) \cdot \underline{\sigma} \cdot \overrightarrow{\text{grad}}(N_j^e) dv \\
(K_{VT}^e)_{ij} &= - \int_{\partial\Omega^e \cap \partial\Omega_j} N_i^e \frac{\partial j}{\partial T} N_j^e ds + \int_{\Omega^e} \overrightarrow{\text{grad}}^T(N_i^e) \cdot \frac{d\sigma}{dT} \cdot \overrightarrow{\text{grad}}(V) N_j^e dv \\
(K_{VV}^e)_{ij} &= - \int_{\partial\Omega^e \cap \partial\Omega_j} N_i^e \frac{\partial j}{\partial V} N_j^e ds + \int_{\Omega^e} \overrightarrow{\text{grad}}^T(N_i^e) \cdot \underline{\sigma} \cdot \overrightarrow{\text{grad}}(N_j^e) dv
\end{aligned}$$

11.2. Resistance welding

The passage of electric currents through a conductive material generates power dissipation via the Joule effect. This principle is particularly used in resistance welding processes. They are very commonly used in the car industry because of their capacity to assemble thin metal sheets.

Figure 11.1 introduces the general principle of spot resistance welding. Two copper electrodes are placed on both sides of the sheets to be assembled and are strongly pressed on them to facilitate contact. A high intensity current melts the metal at the sheet interface. The current used is generally a 50 Hz alternative current. As a result,

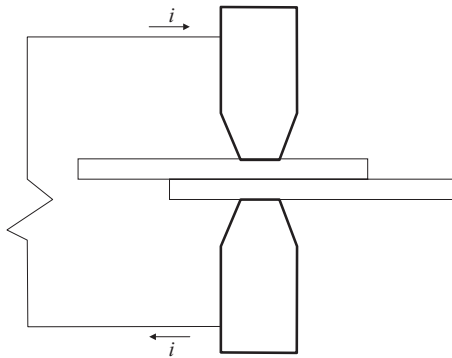


Figure 11.1. Spot resistance welding – principle diagram

the dynamic effects (induced currents) are negligible and the electric currents can be calculated by means of an electrokinetic model.

The simulation of spot resistance welding is of particular interest to study the size of the molten core according to the operational procedure or analyze the electrode lifetime [FEU 06]. Resistance welding involves strongly coupled mechanical, thermometallurgical and electrokinetic phenomena. The electrodes for instance help hold together the sheets being assembled during the solidification phase of the liquid core. The thermometallurgical aspects are covered in Chapter 9. Our study is here restricted to the thermal and electrokinetic aspects; treating the mechanical problems would require developments outside the scope of this book.

First, the model implementation for simulating this process is described. A thermal and electric contact management method is given. Then the results obtained are analyzed to demonstrate the capability of the finite element method to simulate this type of process and thus offer an insight into the local temperatures reached, the material structure after welding, etc.

11.2.1. *Implementing the model*

Figure 11.2 illustrates the mesh used to simulate the welding of three metal sheets 1.2 mm, 0.8 mm and 0.8 mm thick. Note the presence of a large number of contacts between each electrode and a metal sheet on the one hand, and the three metal sheets welded together on the other hand.

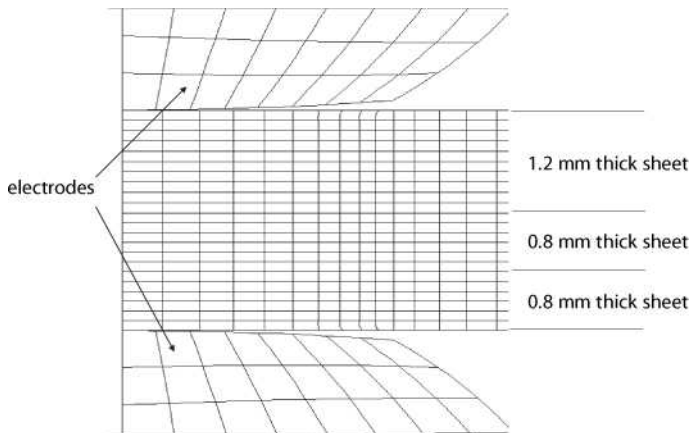


Figure 11.2. *Mesh used*

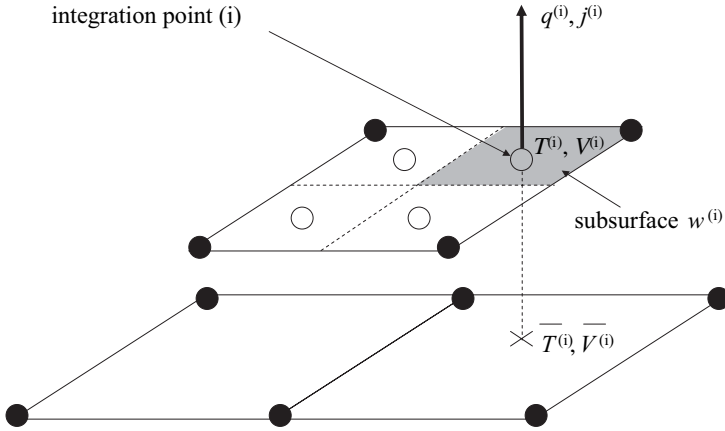


Figure 11.3. *Contact management*

Figure 11.3 illustrates the method used to manage the electric and thermal contacts. It is based upon a Gauss numerical integration scheme [BAT 96] applied to contact elements [JAO 98, BEL 99].

Consider two meshed surfaces in contact and evaluate the current density and the heat flux density entering locally (i.e. at each integration point of the elements) into each of the surfaces. Consider therefore an integration point (i) of an element belonging to one of the opposite surfaces, and term $w^{(i)}$, $T^{(i)}$ and $V^{(i)}$, respectively the weight of the integration point (i) in the Gauss scheme, the temperature and the electric potential at that point (Figure 11.3). Moreover, $\bar{T}^{(i)}$ and $\bar{V}^{(i)}$ represent respectively the temperature and the electric potential on the element opposite the integration point (belonging to the other surface).

The current density $j^{(i)}$ entering through the subsurface $w^{(i)}$ related to the integration point (i) can be written according to the contact's electric resistance R as:

$$j^{(i)} = \frac{1}{R} (\bar{V}^{(i)} - V^{(i)})$$

The thermal flux density $q^{(i)}$ received by the subsurface $w^{(i)}$ related to the integration point (i) can be decomposed into two parts:

$$q^{(i)} = K (\bar{T}^{(i)} - T^{(i)}) + p^{(i)}$$

The first term uses a thermal exchange coefficient K (inverse of a contact thermal resistance) to illustrate the thermal exchanges at the interface. Of course, coefficient

K depends on the properties of that interface, and in that expression it is assumed that the interface has no thermal capacitance.

The second term of the previous equation represents the thermal power dissipated by the Joule effect at the interface. Given the expression of $j^{(i)}$, the total power dissipated at the interface will be $P^{(i)} = (V^{(i)} - \bar{V}^{(i)})^2/R$, and each opposite surface will receive only part of that power. If $f^{(i)}$ is the fraction received by the subsurface $w^{(i)}$, then the thermal flux density received by that surface is written:

$$q^{(i)} = K(\bar{T}^{(i)} - T^{(i)}) + \frac{f^{(i)}}{R}(\bar{V}^{(i)} - V^{(i)})^2$$

The expressions of $q^{(i)}$ and $j^{(i)}$ above are then inserted into the expressions of the global residuals to cancel to solve the coupled problem.

The Gaussian method makes it possible to take into good consideration the thermal flux q and current j local densities. However, corrective terms are inserted to satisfy the global flux conservation [FEU 04].

11.2.2. Results

Figure 11.4 gives the temperature fields calculated at the end of the heating process in the case of three soft steel sheets welded together. Note the presence of a liquid

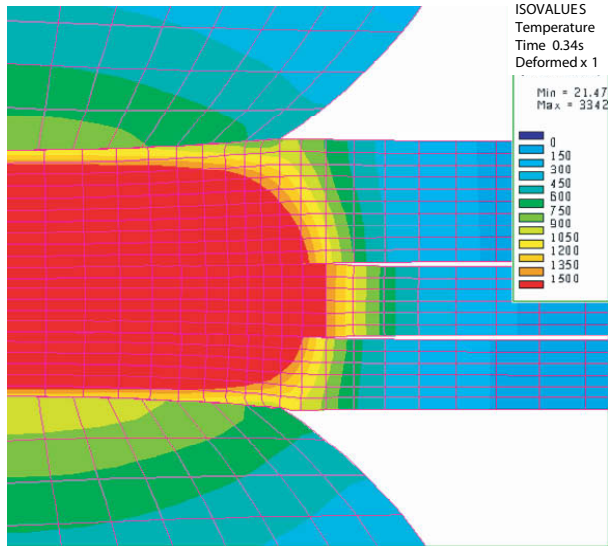


Figure 11.4. *Plotted temperature contour curves*

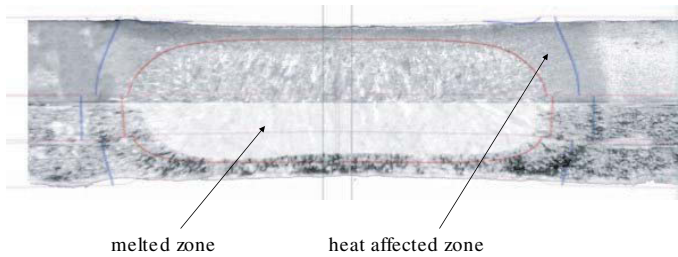


Figure 11.5. Comparing experiment with calculation

core in the center, which is in good agreement with the experiment, as illustrated by Figure 11.5, in which the temperature contour corresponding to the steel *liquidus* was superimposed on top of a micrographic section.

During the calculation process, thermometallurgical coupling was performed with the method described in Chapter 9. In Figure 11.5, the temperature contour corresponding to the steel transformation point A_{c1} (initial austenitizing temperature) was superimposed on top of the micrographic section. Once again there is a fairly good agreement with the experiment.

This finite element numerical simulation of a complex process, involving thermometallurgical phenomena and contact problems, demonstrates that the method is highly appropriate for this type of approach. Numerical simulations make it possible on the one hand to make progress in the command of the process, i.e. the knowledge of relevant parameters, and on the other hand, to perform tests that enable us to optimize the heating sequence according to the liquid core desired.

Chapter 12

Magnetothermal Coupling

12.1. Introduction

Any electric conductor placed in a time-varying magnetic field is the base of eddy-currents that dissipate power through the Joule effect. This power is concentrated at the surface of the component, in a very small thickness which depends on the electric conductivity and magnetic permeability of the material, but also on the frequency of the currents. Induction hardening processes take advantage of this effect because it is therefore possible to control the thickness of the treated layer.

Induction hardening of steel components is performed in two steps. First, an induction heating stage puts the surface of the piece in an austenitic state. Then a cooling stage involving heat conduction through the piece and an external shower enables metallurgical transformations in the heat affected zone in order to form hard structures.

The numerical simulation of this process is of major interest to control and quantify the thickness of the treated layer, the superficial hardness and the residual distortions in order to avoid a final machining which could annihilate the benefits of the treatment. Moreover, such processes often lead to compressive residual stresses at the surface which improve the fatigue behavior of the mechanical component. The complete modeling of the induction hardening process rests upon the solution of coupled electromagnetic, thermal, metallurgical and mechanical problems, as shown in Figure 12.1. In what follows, we will focus on the modeling of induction heating.

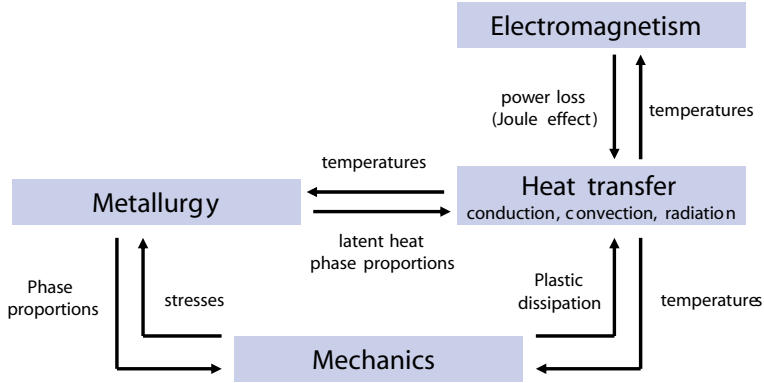


Figure 12.1. *Couplings between physical phenomena involved by the induction hardening process*

The modeling of induction heating rests upon the resolution of a magnetodynamic problem and a problem of heat transfer coupled by:

- the power dissipated by the Joule effect on the one hand;
- the dependence of the electromagnetic properties with temperature on the other hand.

12.2. Magnetic vector potential formulation for magnetodynamics

Magnetodynamic phenomena are governed by the following Maxwell's equations:

$$\overrightarrow{\text{curl}}(\vec{H}) = \vec{J} \quad (\text{Maxwell-Ampère's equation}) \quad [12.1]$$

$$\overrightarrow{\text{curl}}(\vec{E}) = -\frac{\partial \vec{B}}{\partial t} \quad (\text{Maxwell-Faraday's equation}) \quad [12.2]$$

$$\text{div}(\vec{B}) = 0 \quad (\text{Maxwell-Gauss's equation}) \quad [12.3]$$

with the following constitutive equations:

$$\vec{H} = \nu(\|\vec{B}\|, T) \vec{B} \quad (\text{Magnetization law}) \quad [12.4]$$

$$\vec{J} = \sigma(T) \vec{E} \quad (\text{Ohm's law}) \quad [12.5]$$

In the equations above, \vec{H} , \vec{J} , \vec{E} and \vec{B} represent the magnetic field, the electric current density, the electric field and the magnetic flux density vector respectively. In equation [12.4], ν is the magnetic reluctivity (the inverse of the magnetic

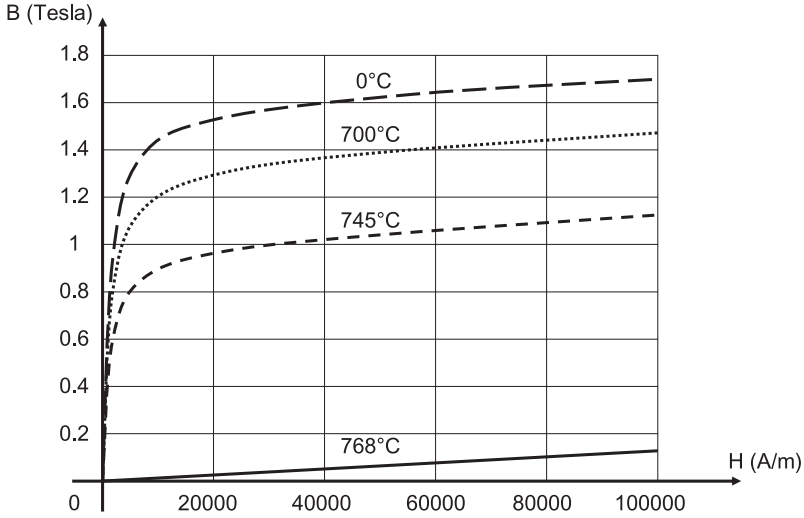


Figure 12.2. Typical magnetization curves $B(H, T)$ for steel

permeability). Taking account of the magnetic field magnitude involved by induction heating applications, hysteresis loops are neglected. In equation [12.5], σ is the electric conductivity. We can note in Figures 12.2 and 12.3 that the electromagnetic properties strongly depend on temperature T .

From equation [12.3], let us now introduce the magnetic vector potential \vec{A} such that:

$$\vec{B} = \text{curl}(\vec{A}) \quad [12.6]$$

To ensure the uniqueness of the magnetic vector potential, an additional equation on $\text{div}(\vec{A})$ must be introduced, for example:

$$\text{div}(\vec{A}) = 0 \quad (\text{Coulomb gauge}) \quad [12.7]$$

Equation [12.2] then gives:

$$\vec{E} = -\frac{\partial \vec{A}}{\partial t} - \text{grad}(V_0) \quad [12.8]$$

where V_0 is the electric scalar potential.

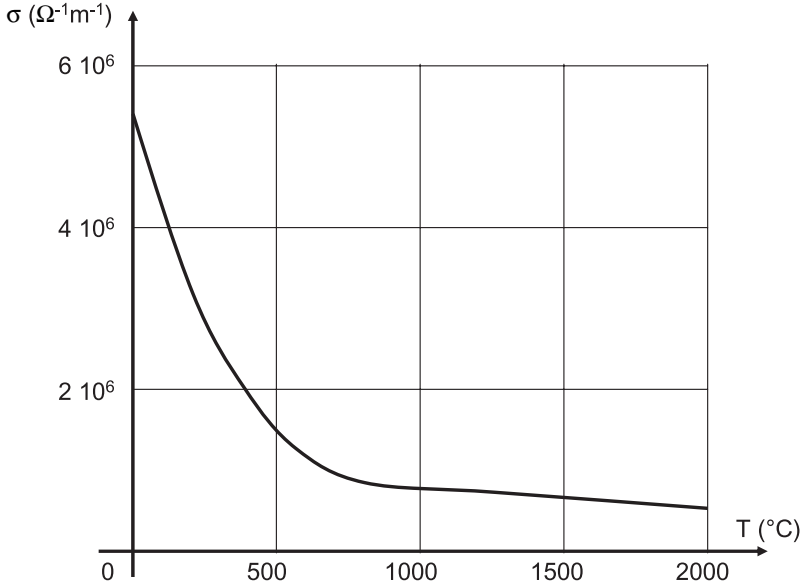


Figure 12.3. Typical temperature-dependent electric conductivity $\sigma(T)$ for steel

Here again, in order to define V_0 uniquely, we must introduce the following additional equation derived from equation [12.1]:

$$\operatorname{div}(\vec{J}) = 0 \quad [12.9]$$

which, with equations [12.5], [12.8] and [12.7], gives: $\operatorname{div}((\sigma \overrightarrow{\operatorname{grad}}(V_0))) = 0$.

The vector potential formulation supposes that equation [12.9] is satisfied *a priori*. It can be shown that this equation is automatically verified for 2D and axisymmetric cases as defined below. Therefore, using equations [12.4], [12.5], [12.6] and [12.8], equation [12.1] finally becomes:

$$\sigma \frac{\partial \vec{A}}{\partial t} + \overrightarrow{\operatorname{curl}}(\nu \overrightarrow{\operatorname{curl}}(\vec{A})) + \sigma \overrightarrow{\operatorname{grad}}(V_0) = 0 \quad [12.10]$$

So, the magnetic vector potential must be the solution of equations [12.7] and [12.10]. We must note that equations [12.7] and [12.10] hold in the whole space. Due to air, the spatial domain on which the solution of both equations must be found is not bounded except for symmetry reasons. Let us now consider the case of a geometry obtained by the translation of a cross-section along the z -axis direction and suppose

that the source current ($\vec{J}_0 = -\sigma \overrightarrow{\text{grad}}(V_0)$) is also directed along the z -axis and only depends on x and y coordinates:

$$\vec{J}_0 = j_0(x, y, t) \vec{e}_z \quad [12.11]$$

In this case, it can be shown that the problem is reduced to a 2D problem because the vector potential now has only one non-zero component:

$$\vec{A} = a(x, y, t) \vec{e}_z \quad [12.12]$$

Similarly, in the case of axisymmetric geometries and a circumferential source current:

$$\vec{J}_0 = j_0(r, z, t) \vec{e}_\theta \quad [12.13]$$

the problem is reduced to an axisymmetric problem and the vector potential can be written:

$$\vec{A} = a(r, z, t) \vec{e}_\theta \quad [12.14]$$

We can note that the gauge condition (equation [12.7]) is automatically satisfied in both cases. Therefore, the magnetic vector potential now has to be the solution of equation [12.10] only [MEU 08].

12.3. Coupled finite element-boundary element method

The idea of coupling the finite element method (FEM) and the boundary element method (BEM) to benefit from the advantages of both methods is rather old [ZIE 77, ATL 78]. It finds here a rather natural application [FET 97, PAS 03b]. Let us subdivide the physical space into a domain Ω_1 made up of conductors (possibly traversed by sources of current) and a domain Ω_2 constituted by air (Figure 12.4).

Let us call Γ the common border between Ω_1 and Ω_2 . The problem to be solved in the air (Ω_2) is reduced to:

$$\overrightarrow{\text{curl}}(\nu_0 \overrightarrow{\text{curl}}(\vec{A})) = 0 \quad [12.15]$$

$$\text{div}(\vec{A}) = 0 \quad [12.16]$$

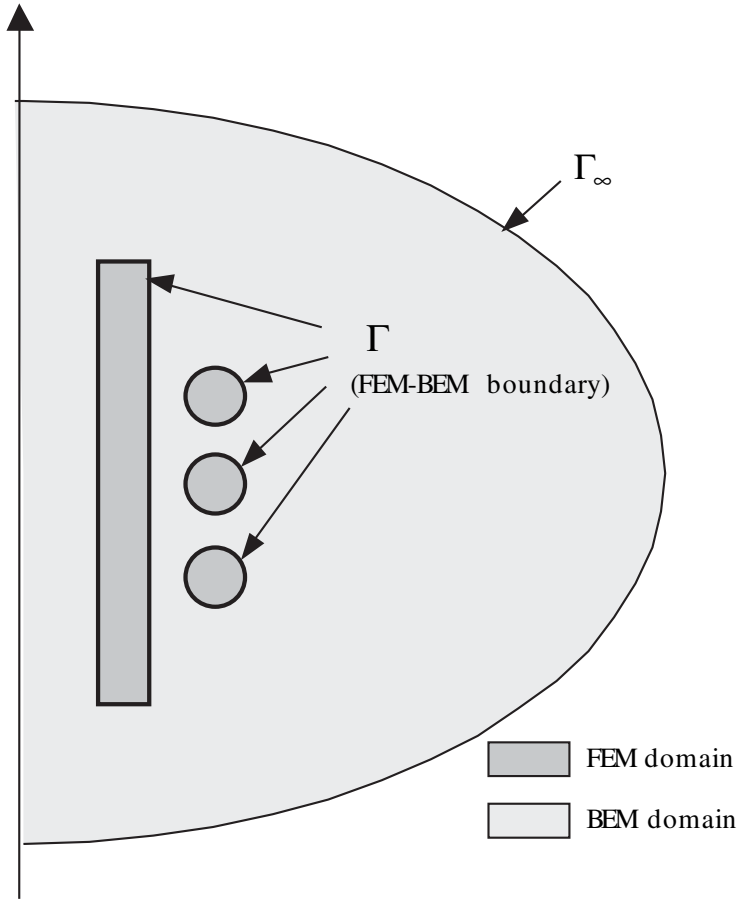


Figure 12.4. *Coupling between FEM and BEM domains*

with the following boundary conditions:

$$\vec{A} = 0 \text{ to infinity on } \Gamma_\infty \quad [12.17]$$

$$\text{continuity of } \vec{A} \text{ and } \text{div}(\vec{A}) \text{ on } \Gamma \quad [12.18]$$

$$\vec{h}_1 = \vec{H} \times \vec{n}_1 = -\vec{h}_2 = -\vec{H} \times \vec{n}_2 \text{ on } \Gamma \quad [12.19]$$

where \vec{n}_1 and \vec{n}_2 are the units normal to Γ outside Ω_1 and Ω_2 respectively and $\nu_0 = \frac{1}{\mu_0}$.

Consider a mesh of Ω_1 including n nodes and a compatible mesh of the border Γ including b nodes. Moreover, suppose that \vec{h}_1 in equation [12.19] is defined from nodal values $\{H_1\}$ (b components) by using the shape functions of the elements located on the border Γ .

12.3.1. Finite element formulation

The weak formulation associated with equations [12.10] and [12.7] is obtained by using the following relation, where the \vec{A}^* vectors are weighting functions associated with \vec{A} :

$$\begin{aligned} \int_{\Omega_1} \vec{A}^* \cdot \overrightarrow{\text{curl}}(\nu \overrightarrow{\text{curl}}(\vec{A})) dv &= \int_{\Omega_1} \overrightarrow{\text{curl}}(\vec{A}^*) \cdot \nu \overrightarrow{\text{curl}}(\vec{A}) dv \\ &+ \int_{\Gamma} \vec{A}^* \cdot (\vec{n}_1 \wedge \nu \overrightarrow{\text{curl}}(\vec{A})) ds \end{aligned} \quad [12.20]$$

The following weak formulation can thus be easily derived:

Find $\vec{A} : t \in]0, T[\mapsto \vec{A}(\vec{x}, t) \in V$ such that

$\vec{A}(\vec{x}, 0) = \vec{A}_0(\vec{x})$ and for any $\vec{A}^* \in V$:

$$\begin{aligned} \int_{\Omega_1} \vec{A}^* \cdot \sigma \frac{\partial \vec{A}}{\partial t} dv + \int_{\Omega_1} \overrightarrow{\text{curl}}(\vec{A}^*) \cdot \nu \overrightarrow{\text{curl}}(\vec{A}) dv \\ = \int_{\Omega_1} \vec{A}^* \cdot \vec{J}_0 dv + \int_{\Gamma} \vec{A}^* \cdot \vec{h}_1 ds \end{aligned} \quad [12.21]$$

In the above equation, $]0, T[$ is a given time interval (one or more periods of the source current), $\vec{A}_0(\vec{x})$ an initial condition and V the functional space of admissible functions i.e. satisfying equation [12.7] and the conditions of regularity necessary to give a sense to equation [12.21].

Let us now consider axisymmetric geometries and the circumferential source current. In this case, the source current and the magnetic vector potential are given by equations [12.13] and [12.14] respectively. We can note that we also have $\vec{h}_1 = h_1 \vec{e}_\theta$ and $\vec{h}_2 = h_2 \vec{e}_\theta$. The gauge condition is automatically satisfied and we obtain $\overrightarrow{\text{curl}} \vec{A} = \frac{1}{r} \text{grad}(ra) \times \vec{e}_\theta$. This formula encourages us to seek a nodal

approximation of ra rather than a . The finite element discretization of variational equation [12.21] then leads to the following system of equations:

$$[C] \cdot \{\dot{A}\} + \{R\} - [T] \cdot \{H_1\} = \{J_0\} \quad [12.22]$$

$\{A\}$ contains the nodal values of ra , $[C]$ is a $n \times n$ matrix, $\{R\}$ is a n -component vector, $[T]$ is a $n \times b$ rectangle matrix and $\{J_0\}$ is a n -component vector containing the nodal values equivalent to the source current. $[C]$, $\{R\}$, $[T]$ and $\{J_0\}$ are obtained from an assembly process from the following element vectors and matrices:

$$\begin{aligned} [C^e] &= \int_{\Omega^e} \frac{\sigma}{r^2} \{N^e\} \cdot \langle N^e \rangle dv \\ \{R^e\} &= \int_{\Omega^e} \left\{ \overrightarrow{\text{curl}}^T(N^e) \right\} \cdot \nu \left\langle \overrightarrow{\text{curl}}(N^e) \right\rangle \cdot \{A^e\} dv \\ [T^e] &= \int_{\partial\Omega^e \cap \Gamma} \frac{1}{r} \{N^e\} \cdot \langle N^e \rangle ds \\ \{J_0^e\} &= \int_{\Omega^e} \frac{J_0}{r} \{N^e\} dv \end{aligned} \quad [12.23]$$

where $\langle \overrightarrow{\text{curl}}(N^e) \rangle = \langle \frac{1}{r} \overrightarrow{\text{grad}}(N_1^e) \times \vec{e}_\theta, \dots, \frac{1}{r} \overrightarrow{\text{grad}}(N_n^e) \times \vec{e}_\theta \rangle$ is the matrix giving $\overrightarrow{\text{curl}}(\vec{A})$ from nodal values of ra . Note that $\{R^e\}$ can be a non-linear function of $\{A\}$ since ν is $\|\vec{B}\|$ -dependent, and hence depends on \vec{A} via equation [12.4].

12.3.2. Boundary element formulation

Equation [12.15] multiplied by a scalar weight function $\psi(r, z)$, then integrated over Ω_2 leads, after some mathematical handling and taking account of equations [12.16] and [12.17], to the following integral equation:

$$\begin{aligned} \int_{\Omega_2} \nu_0 \Delta \psi \vec{A} dv &= \int_{\Gamma} (\nu_0 (\overrightarrow{\text{grad}}(\psi) \cdot \vec{n}_2) \vec{A} - \psi \nu_0 \overrightarrow{\text{curl}}(\vec{A}) \times \vec{n}_2) ds \\ &\quad - \int_{\Gamma} (\nu_0 (\overrightarrow{\text{grad}}(\psi) \cdot \vec{A}) \vec{n}_2 - \nu_0 (\vec{A} \cdot \vec{n}_2) \overrightarrow{\text{grad}}(\psi)) ds \end{aligned} \quad [12.24]$$

Considering the axisymmetric case, we obtain $\overrightarrow{\text{grad}}(\psi) \cdot \vec{A} = 0$ and $\vec{A} \cdot \vec{n}_2 = 0$. Moreover, taking equation [12.19] into account, equation [12.24] thus becomes:

$$\int_{\Omega_2} \nu_0 \Delta \psi a dv = \int_{\Gamma} (\nu_0 (\overrightarrow{\text{grad}}(\psi) \cdot \vec{n}_2) a - \psi h_2) ds \quad [12.25]$$

Considering weighting functions ψ_p as the solution of Poisson's equation for a unit load located at point \vec{x}_p : $\Delta\psi_p + \delta(\vec{x} - \vec{x}_p) = 0$ (δ Dirac's distribution) with $\psi_p = 0$ at infinity ([BRE 92]), equation [12.25] becomes at each node i of the border Γ :

$$\nu_0 \frac{c_i}{2\pi} a_i = \int_{\Gamma} \psi_i h_2 ds - \int_{\Gamma} \nu_0 \frac{\partial \psi_i}{\partial n} a ds \quad [12.26]$$

In this equation, c_i is the inside angle at node i ($c_i = \pi$ for a very smooth boundary). Using the mesh of Γ previously introduced for the discretization of equation [12.26], we finally obtain:

$$[H] \cdot \{A_2\} - [G] \cdot \{H_2\} = 0 \quad [12.27]$$

where $[H]$ and $[G]$ are $b \times b$ component matrices with $H_{ij} = \nu_0 \frac{c_i}{2\pi r_i} \delta_{ij} + \int_{\Gamma} \nu_0 \frac{\partial \psi_i}{\partial n} \frac{1}{r} N_j^e ds$ and $G_{ij} = \int_{\Gamma} \psi_i N_j^e ds$, and where $\{A_2\}$ is a b -component vector containing the nodal values of ra at the nodes of Γ .

12.3.3. FEM-BEM coupling

Equation [12.19] gives:

$$\{H_1\} = -\{H_2\} \quad [12.28]$$

Equation [12.22] combined with equations [12.27] and [12.28] becomes:

$$[C] \cdot \{\dot{A}\} + \{R\} + [T] \cdot [G]^{-1} \cdot [H] \cdot \{A_2\} = \{J_0\} \quad [12.29]$$

The non-symmetric matrix $[T] \cdot [G]^{-1} \cdot [H]$ appears as the contribution of air to the system of equations [12.29]. The system of differential equations [12.29] can thus be solved by a method of direct integration (implicit Euler or central differences for example) together with a Newton-Raphson method to find the solution at each time-step.

12.4. A harmonic balance method for the magnetodynamic problem

In practice, the source current is always sinusoidal. However, due to the non-linearity of the magnetization law, induced electromagnetic fields contain higher order harmonics. Rather than solve equation [12.29] by a direct method step by step over time, a harmonic balance method can be used [YAM 88a, YAM 88b, LU 95,

TER 84]. This method consists of calculating the first terms of the development in a Fourier series of the magnetic vector potential:

$$\vec{A}(\vec{x}, t) = \sum_{k=1,3,\dots} (\vec{A}_{kc}(\vec{x}) \cos(k\omega t) + \vec{A}_{ks}(\vec{x}) \sin(k\omega t)) \quad [12.30]$$

where $\omega = \frac{2\pi}{\tau}$ represents the pulsation of the source current and τ , the period.

Equation [12.10] is then projected on to the basis of chosen functions of time. The equations to solve are now written:

$$\text{for } l = 1, 3, \dots \quad \begin{cases} \int_0^\tau \left(\sigma \frac{\partial \vec{A}}{\partial t} + \overrightarrow{\text{curl}}(\nu \overrightarrow{\text{curl}}(A)) - \vec{J}_0 \right) \cos(l\omega t) dt = 0 \\ \int_0^\tau \left(\sigma \frac{\partial \vec{A}}{\partial t} + \overrightarrow{\text{curl}}(\nu \overrightarrow{\text{curl}}(A)) - \vec{J}_0 \right) \sin(l\omega t) dt = 0 \end{cases} \quad [12.31]$$

We obtain a new problem, where the unknowns are now the fields $\vec{A}_{kc}(\vec{x})$ and $\vec{A}_{ks}(\vec{x})$ for $k = 1, 3, \dots, m$, m being the order of the last harmonic considered. The FEM applied to equations [12.31] in the axisymmetric cases leads us to solve the following system of non-linear equations:

$$\begin{Bmatrix} \vdots \\ \{R_{lc}(\{A_{lc}\}, \{A_{ls}\}, \dots)\} \\ \{R_{ls}(\{A_{lc}\}, \{A_{ls}\}, \dots)\} \\ \vdots \end{Bmatrix} = 0 \quad [12.32]$$

In these equations, the element terms are calculated as follows for each element e of Ω_1 :

$$\begin{aligned} \{R_{lc}^e(\dots)\} &= - \int_{\Omega}^e \{N^e\} \frac{l\sigma\omega}{r^2} \langle N^e \rangle \cdot \{A_{ls}^e\} dv + \int_{\Omega}^e \frac{1}{r} \{N^e\} J_{0lc}^e dv \\ &\quad - \sum_{k=1,3,\dots} \int_{\Omega}^e \left\{ \overrightarrow{\text{curl}}^T(N^e) \right\} \cdot \nu_{lc,kc} \cdot \left\langle \overrightarrow{\text{curl}}(N^e) \right\rangle \cdot \{A_{kc}^e\} dv \\ &\quad - \sum_{k=1,3,\dots} \int_{\Omega}^e \left\{ \overrightarrow{\text{curl}}^T(N^e) \right\} \cdot \nu_{lc,ks} \cdot \left\langle \overrightarrow{\text{curl}}(N^e) \right\rangle \cdot \{A_{ks}^e\} dv \end{aligned} \quad [12.33]$$

$$\begin{aligned}
\{R_{ls}^e(\cdots)\} &= \int_{\Omega}^e \{N^e\} \frac{l\sigma\omega}{r^2} \langle N^e \rangle \cdot \{A_{lc}^e\} dv + \int_{\Omega}^e \frac{1}{r} \{N^e\} J_{0ls}^e dv \\
&\quad - \sum_{k=1,3,\dots} \int_{\Omega}^e \left\{ \overrightarrow{\text{curl}}^T(N^e) \right\}^T \cdot \nu_{ls,kc} \cdot \left\langle \overrightarrow{\text{curl}}(N^e) \right\rangle \cdot \{A_{kc}^e\} dv \\
&\quad - \sum_{k=1,3,\dots} \int_{\Omega}^e \left\{ \overrightarrow{\text{curl}}^T(N^e) \right\}^T \cdot \nu_{ls,ks} \cdot \left\langle \overrightarrow{\text{curl}}(N^e) \right\rangle \cdot \{A_{ks}^e\} dv
\end{aligned} \tag{12.34}$$

For example, we have $\nu_{ls,ks} = \int_0^\tau \nu(\|\overrightarrow{\text{curl}}(A)\|) \cos(\omega t) \sin(k\omega t) dt$. As mentioned previously, $\{A_{kc}\}$ and $\{A_{ks}\}$ contain the nodal values of ra_{kc} and ra_{ks} respectively.

The generalization of the coupling between the finite element method and the boundary element method does not create any particular problem. Equations [12.15]-[12.19] must now be solved for each harmonic considered. We thus obtain $m+1$ systems of equations of the [12.27] type:

$$\begin{aligned}
[H] \cdot \{A_{kc2}\} - [G] \cdot \{H_{kc2}\} &= 0 \\
[H] \cdot \{A_{ks2}\} - [G] \cdot \{H_{ks2}\} &= 0
\end{aligned} \tag{12.35}$$

The coupling is obtained by considering each element of Γ :

$$\begin{aligned}
\{R_{lc}^e\} &= \int_{S^e} \frac{1}{r} \{N^e\} \cdot \langle N^e \rangle ds \cdot \{H_{lc1}\} \\
\{R_{ls}^e\} &= \int_{S^e} \frac{1}{r} \{N^e\} \cdot \langle N^e \rangle ds \cdot \{H_{ls1}\}
\end{aligned} \tag{12.36}$$

with $\{H_{kc2}\} = -\{H_{kc1}\}$ and $\{H_{ks2}\} = -\{H_{ks1}\}$ for $k = 1, 3, \dots, m$.

12.5. Coupling magnetodynamics with heat transfer

12.5.1. Iterative coupling

The main difficulty arising from the numerical simulation of couplings between magnetodynamics and heat transfer is related to time constants of both phenomena, which differ considerably. Indeed, the frequency of currents usually utilized with induction heating applications are in the order of magnitude of a few tens of kHz.

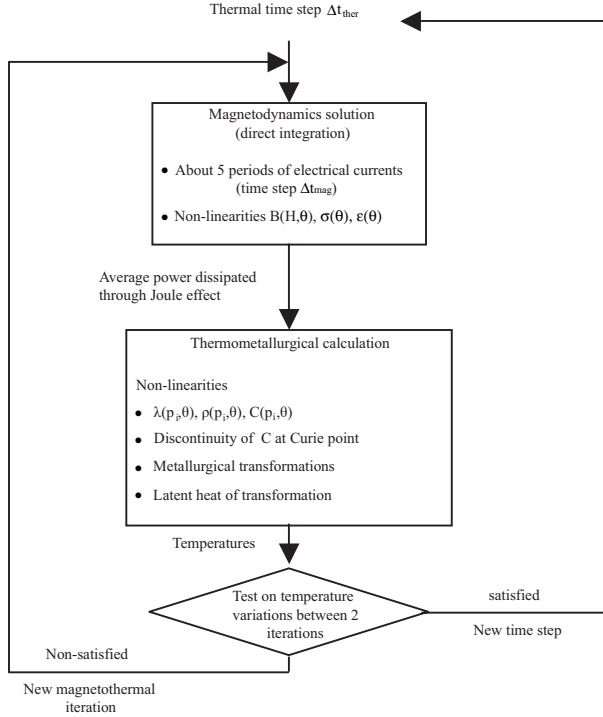


Figure 12.5. *Iterative procedure for coupling magnetodynamics with heat transfer*

Compared to these frequencies, thermal phenomena vary very slowly. It is therefore impossible to solve the problem using a direct method where the magnetic vector potential field and the temperature field are computed together simultaneously at each time step. Such an approach would need time steps in the order of 10^{-6} s (time step needed for the calculation of electromagnetic fields). This would lead to unreasonable computational times and, moreover, to serious numerical problems, the very small temperature variations during a time step being potentially drowned in calculation errors. On the other hand, it is completely justified to solve the problem using an iterative procedure where magnetodynamics and thermal problems are solved successively, as shown in Figure 12.5 [CHA 97, BER 00].

The power dissipated through the Joule effect resulting from magnetodynamics analysis is then considered on average over a period of the source current in the thermal analysis:

$$Q(\vec{x}) = \frac{1}{\sigma\tau} \int_0^\tau \vec{J}(\vec{x}, t)^2 dt \quad [12.37]$$

where $\vec{J} = \vec{J}_0 - \sigma \frac{\partial \vec{A}}{\partial t}$ and τ represent the total electric current density and the period of the source current respectively.

It is clear that such a method leads to costly calculations that are not always compatible with industrial objectives. In addition to non-linearities associated with magnetodynamics and thermal problems, the modeling of the induction heating process introduces two other non-linearities that must be accounted for:

– The non-linearity and the temperature dependence of the magnetization law that can be represented for example by the Frohlich-Kenelly model [BOS 78, CHA 97]:

$$\vec{B} = \left(\mu_0 + \frac{B_0}{\|\vec{H}\| + H_\alpha} \right) \vec{H} \quad [12.38]$$

where μ_0 is the void permeability, and B_0 and H_α are two parameters depending on temperature.

– The transition of the behavior of steel at the Curie point. The Curie point is the temperature above which ferromagnetic materials such as steel lose their ferromagnetic ability. Below the Curie point, the material is ferromagnetic and the magnetization curve is very non-linear, as shown in Figure 12.2. Above the Curie point, the material is paramagnetic and we simply obtain $\vec{B} = \mu_0 \vec{H}$. In addition to the magnetic effects, the passage of the Curie point is accompanied by an intense variation of the specific heat.

12.5.2. A direct method for magnetothermal coupling

The use of a harmonic balance method enables us to treat the magnetothermal coupling by a direct method [FET 91, PAS 03a]. Indeed, we only have to define a finite element including $m + 2$ degrees of freedom: the cosine and sine parts of the $\frac{m+1}{2}$ harmonics of the magnetic vector potential and the temperature. The element residual vector of such a finite element is given by equations [12.33] and [12.34] for the magnetodynamics part. For the thermal part, we obtain:

$$\begin{aligned} \{R_T^e\} = & \int_{\Omega^e} \{N^e\} \frac{1}{2\sigma} \sum_{k=1,3,\dots}^m (J_{kc}^2 + J_{ks}^2) dv + \int_{\Omega^e} \{N^e\} \rho \frac{H - H_0}{\Delta t} dv \\ & - \int_{\Omega^e} \left\{ \overrightarrow{\text{grad}}^T(N^e) \right\} \cdot \lambda \cdot \overrightarrow{\text{grad}}(T) dv \end{aligned} \quad [12.39]$$

Note that in equation [12.39], H_0 and H represent the enthalpy (and not the magnetic field) at time t and $t + \Delta t$ respectively. Combined with the coupling between

finite elements and boundary elements, we obtain a method specially suited to the modeling of surface induction hardening processes involving moving parts.

12.6. Application: induction hardening of a steel cylinder

Now consider the superficial heat treatment of a cylindrical steel bar. The heating is ensured by a toric copper inductor fed with tension (frequency = 50 Hz) and the cooling by a shower, the device moving axially along the cylinder in order to treat the whole surface (Figure 12.6).

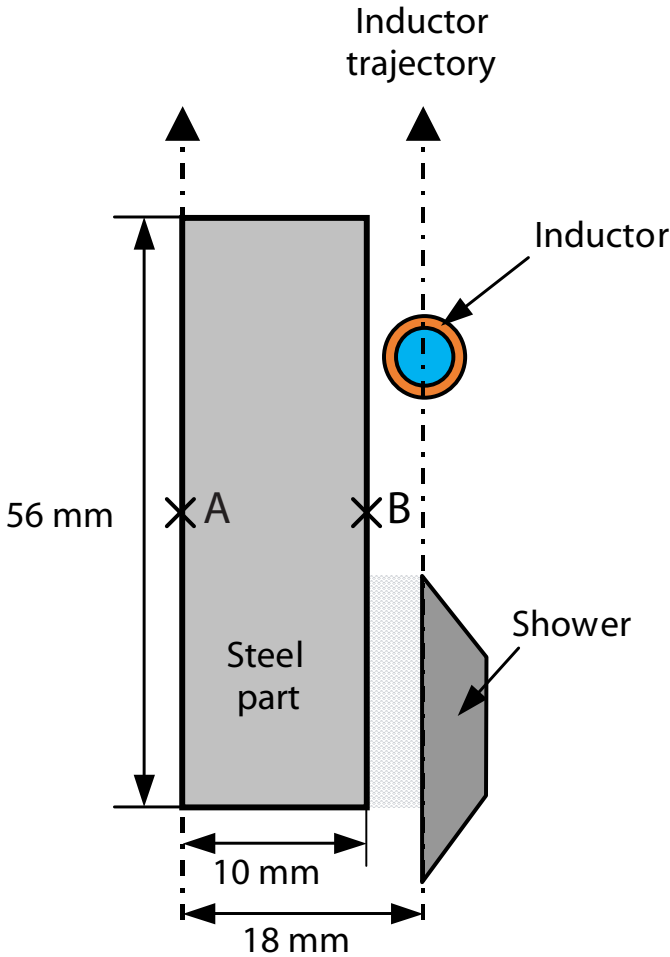


Figure 12.6. Induction hardening device

The device has a symmetry of revolution so that the phenomena can be studied on a meridian cut in an axisymmetric option of calculation.

The mesh of the cylinder includes 1,200 elements and 1,271 nodes and the air is taken into account by the boundary element method as described above. The inductor is fed during 3.2 s during which it moves over a 64 mm length (length of the cylinder). The thermometallurgical analysis uses a time step of 0.05 s corresponding to a 1 mm displacement of the inductor. Each displacement of the inductor requires a new calculation of the BEM quantities associated with the air. The shower follows the inductor at a distance of 20 mm from it. Its action continues after cutting off the inductor current during 5.8 s, then the cylinder is entirely cooled in a bath. Steel at 20°C consists of ferrite. During the heating phase, this ferrite can be transformed into austenite which will be able to be transformed into ferrite, pearlite, bainite or martensite during cooling. The conditions of simulation as well as the physical properties of materials and the coefficients of the laws used are detailed in [PAS 03a] and will not be recalled here.

Five different simulations were carried out. These simulations differ only in the heating phase. In all cases, the simulation of cooling is based on a traditional thermometallurgical analysis. The first simulation uses an indirect method of resolution of the coupling between the magnetodynamic and thermometallurgical phenomena as described in section 12.5.1. The time integration of the magnetodynamic phenomena is based on the trapezoidal rule and is carried out over 3 periods of the inductor current (approximately 100 time steps per period) in order to reach the periodic state. The second and the third simulations are carried out with the direct method described in section 12.5.2. The harmonic balance method is used here with 1 harmonic for simulation 2 and 2 harmonics for simulation 3. Simulations 4 and 5 use an indirect approach with a harmonic balance method with 1 (simulation 4) or 2 harmonics (simulation 5).

Figure 12.7 gives the distributions of temperature and austenite proportion obtained during the displacement of the inductor. The difference in temperature calculated by simulations 1 and 3 on a line joining points A and B (Figure 12.6) remains lower than 27.5°C for maximum temperatures reached about 1,300°C. If we consider simulation 2, the difference passes to 50.5°C which remains acceptable on account of the resolution time saved with this method. Indeed, simulation 2 seems to be the most efficient insofar as it is the one that provides by far the best compromise between accuracy of results and computation time (see Table 12.1). Table 12.1 shows in addition that, with the harmonic balance method, the choice of a direct resolution method for the magnetothermal coupling is preferable to that of an indirect method.

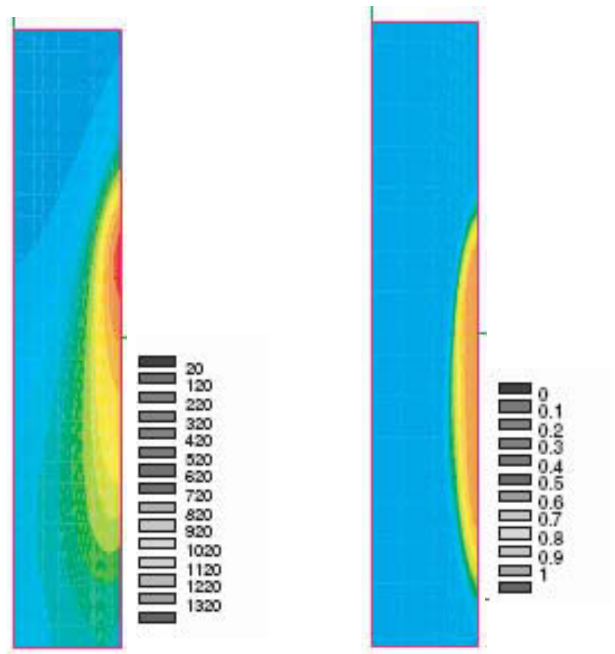


Figure 12.7. *Temperature (left) and austenite proportion (right) distributions at time 2.15 s*

Simulation	1	2	3	4	5
Magnetothermal coupling	indirect method	direct method	direct method	indirect method	indirect method
Magnetodynamics model	Transient calculation	1 harmonic	2 harmonics	1 harmonic	2 harmonics
	Over 3 periods				
CPU	17 h	1 h 35 m	8 h 11 m	4 h 31 m	20 h 58 m

Table 12.1. *Times of resolution for the different simulations (PC Pentium IV 1.7 GHz RAM = 512 MO)*

Bibliography

- [ATL 78] ATLURI S.N., GRANDEL J.J., Boundary element methods (BEM) and combination of BEM-FEM, Report num. GIT-ESM-SA-78-16, Center for the Advancement of Computational Mechanics, Georgia Institute of Technology, GA, 1978.
- [AVR 40a] AVRAMI M., “Kinetics of phase change. I: general theory”, *J. Chem. Phys.*, vol. 7, p. 103–112, 1940.
- [AVR 40b] AVRAMI M., “Kinetics of phase change. II: transformation-time relations for random distribution of nuclei”, *J. Chem. Phys.*, vol. 8, p. 212–224, 1940.
- [AVR 41] AVRAMI M., “Kinetics of phase change. III: granulation, phase change and microstructure”, *J. Chem. Phys.*, vol. 9, p. 117–184, 1941.
- [BAT 96] BATHE K.J., *Finite Element Procedures*, Prentice Hall, 1996.
- [BEL 99] BELLET M., JAOUEN O., “Finite element approach of thermomechanics of solidification processes”, in OHNAKA I., YASUDA H., Eds., *Int. Conf. on Cutting Edge of Computer Simulation of Solidification and Casting*, The Iron and Steel Institute of Japan (ISIJ), p. 173–190, November 14–16 1999.
- [BÉN 84] BÉNARD J., MICHEL A., PHILIBERT J., TALBOT J., *Métallurgie générale*, Masson, 2nd edition, 1984.
- [BER 86] BERGHEAU J.M., GALAUP C., “Influence du maillage en E.F. sur la cinétique de solidification d’une pièce moulée”, in *Proceedings of E-MRS Conference, Strasbourg*, p. 39–48, 1986.
- [BER 98] BERGHEAU J.M., LALOT V., “Finite element modeling of heat exchangers”, in *Advanced Computational Method in Heat Transfer V*, Computational Mechanics Publications, p. 369–378, 1998.
- [BER 00] BERGHEAU J.M., CONRAUX P., “FEM-BEM coupling for the modelling of induction heating processes including moving parts”, *J. Shanghai Jiao Tong University*, vol. E-5, num. 1, p. 91–99, 2000.

- [BER 01] BERGHEAU J.M., POTIER F., “Finite element modeling of coupled radiative and diffusive heat transfers in non participating media including symmetry and periodicity conditions”, *Numerical Heat Transfer, Part B*, vol. 40, num. 3, p. 229–247, 2001.
- [BER 03] BERGHEAU J.M., DEVAUX J., DURANTON P., FORTUNIER R., LARREUR M., LEBLOND J.B., “Numerical simulation of superficial diffusion and precipitation of chemical elements with application to some nitriding process”, in BREBBIA C.A., DE HOSSON J.T.M., NISHIDA S.I., Eds., *Surface Treatment VI*, p. 297–306, 2003.
- [BES 01] BESSON J., CAILLETAUD G., CHABOCHE J.L., FOREST S., *Mécanique non linéaire des matériaux*, Hermes, 2001.
- [BOR 90] BORNSIDE D.E., KINNEY T.A., BROWN R.A., “Finite element/newton method for the analysis of Czochralski crystal growth with diffuse grey radiative heat transfer”, *Int. J. Numer. Meth. Eng.*, vol. 30, p. 133–154, 1990.
- [BOS 78] BOSSAVIT A., Chauffage d’un cylindre d’acier par induction. Modèles numériques, Report, note EDF/DER, code HI2825-02, 1978.
- [BRE 90] BREITBACH G., “Solution of radiative problems using variational based finite element method”, *Int. J. Numer. Meth. Eng.*, vol. 29, p. 1701–1714, 1990.
- [BRE 92] BREBBIA C.A., DOMINGUEZ J., *Boundary Elements – An Introductory Course*, Computational Mechanics Publications, 2nd edition, 1992.
- [BRO 82] BROOKS A., HUGHES T.J.R., “Upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on incompressible Navier-Stokes equations”, *Computer Methods in Applied Mechanics and Engineering*, vol. 33, p. 199–259, 1982.
- [CAR 59] CARSLAW H.S., JAEGER J.C., *Conduction of Heat in Solids*, Clarendon Press, Oxford, 1959.
- [CHA 97] CHABOUDEZ C., CLAIN S., GLARDON R., MARI D., RAPPAZ J., SWIERKOSZ M., “Numerical modelling in induction heating for axisymmetric geometries”, *IEEE Transactions on Magnetics*, vol. 33, num. 1, p. 739–745, 1997.
- [CHU 82] CHUNG T.J., KIM J.Y., “Radiation view factors by finite elements”, *ASME J. Heat Transfer*, vol. 104, p. 792–895, 1982.
- [CIA 70] CIARLET P.G., “Discrete maximum principle for finite difference operators”, *AEQ Math*, vol. 4, p. 338–352, 1970.
- [CIA 82] CIARLET P.G., *Introduction à l’analyse numérique matricielle et à l’optimisation*, Masson, 1982.
- [CLA 02] CLAUDINON S., LAMESLE P., ORTEU J.J., FORTUNIER R., “Continuous *in situ* measurement of quenching distortions using computer vision”, *J. Mat. Proc. Tech.*, vol. 122, p. 69–81, 2002.

- [COM 74] COMINI G., DEL GIUDICE S., LEWIS R.W., ZIENKIEWICZ O.C., "Finite element solution of non-linear heat conduction problems with special reference to phase change", *Int. J. Numerical Methods in Engineering*, vol. 8, num. 613–624, 1974.
- [COM 94] COMINI G., DEL GIUDICE S., NONINO C., *Finite Element Analysis in Heat Transfer, Basic Formulation and Linear Problem*, Taylor & Francis, 1994.
- [DAL 86] DALHUIJSEN A.J., SEGAL A., "Comparison of finite element techniques for solidification problems", *Int. J. Numer. Meth. Eng.*, vol. 23, p. 1807–1829, 1986.
- [DEN 76] DENNIS JR. J.E., "A brief survey of convergence results for quasi-Newton methods", *SIAM-AMS Proceedings*, vol. 9, p. 185–199, 1976.
- [DEN 97] DENIS S., "Prévision des contraintes résiduelles induites par traitement thermique et thermochimique", *Revue française de métallurgie, CIT/sciences et génie des matériaux*, vol. 2, p. 157–176, 1997.
- [DES 87] DESBIOLLES J.-L., DROUX J.-J., RAPPAZ J., RAPPAZ M., "Simulation of solidification of alloys by the finite element method", *Computer Physics Reports*, vol. 6, p. 371–383, 1987.
- [DRO 90] DROUX J.J., Simulations numériques bidimensionnelles et tridimensionnelles de processus de solidification, PhD thesis, Ecole Polytechnique Fédérale de Lausanne, Switzerland, 1990.
- [EME 91] EMERY A.F., JOHANSSON O., LOBO M., ABROUS A., "A comparative study of methods for computing the diffuse radiation viewfactors for complex structures", *ASME J. Heat Transfer*, vol. 113, p. 413–422, 1991.
- [ENG 91] ENGELMAN M., JAMNIA M.A., "Grey-body surface radiation coupled with conduction and convection for general geometries", *Int. J. Numer. Meth. Fluids*, vol. 13, p. 1029–1053, 1991.
- [FER 85] FERNANDES F.M., DENIS S., SIMON A., "Mathematical model coupling phase transformations and temperature evolution during quenching of steels", *Mat. Sci. Technol.*, vol. 10, p. 838–844, 1985.
- [FET 91] FETACHI M., DEVELEY G., "Magneto-thermal behavior finite element analysis for ferromagnetic materials in induction heating devices", *IEEE Transactions on Magnetics*, vol. 27, num. 6, p. 5235–5237, 1991.
- [FET 97] FETZER J., KURZ S., LEHNER G., "The coupling of boundary elements and finite elements for non-destructive testing applications", *IEEE Transactions on Magnetics*, vol. 33, num. 1, p. 677–681, 1997.
- [FEU 04] FEULVARCH E., ROBIN V., BERGHEAU J.M., "Resistance spot welding simulation: a general finite element simulation of electrothermal contact conditions", *J. Mat. Proc. Technology*, vol. 153–154, p. 436–441, 2004.

- [FEU 06] FEULVARCH E., ROGEON PH., CARRÉ P., ROBIN V., SIBILIA G., BERGHEAU J.M., “Resistance spot welding process: experimental and numerical modelling of the weld growth mechanisms with consideration of contact conditions”, *Numerical Heat Transfer, Part A*, vol. 49, p. 345–367, 2006.
- [FEU 07] FEULVARCH E., BERGHEAU J.M., “An implicit fixed grid method for the finite element analysis of heat transfer involving phase changes”, *Numerical Heat Transfer, Part B*, vol. 51, p. 585–610, 2007.
- [FOR 95] FORTUNIER R., LEBLOND J.B., PONT D., “Recent advances in the numerical simulation of simultaneous diffusion and precipitation of chemical elements in steels”, in HAWBOLT E.B., YUE S., Eds., *Phase Transformations During the Thermal/Mechanical Processing of Steel*, p. 357–365, 1995.
- [FOR 00a] FORTUNIER R., LEBLOND J.-B., BERGHEAU J.-M., “Computer simulation of thermochemical treatments: Modelling diffusion and precipitation in metals”, *J. Shanghai Jiao Tong University*, vol. E-5, num. 1, p. 303–309, 2000.
- [FOR 00b] FORTUNIER R., LEBLOND J.-B., BERGHEAU J.-M., “A numerical model for multiple phase transformations in steels during thermal processes”, *J. Shanghai Jiao Tong University*, vol. E-5, num. 1, p. 213–220, 2000.
- [GAL 15] GALERKIN B.G., *Series Solutions of Some Problems of Elastic Equilibrium of Rods and Plates*, Wjestnik Ingenerow, Petrograd, 1915.
- [HIL 78] HILL R., “Aspects of invariance in solid mechanics”, in *Advances in Applied Mechanics*, vol. 18, Academic Press, p. 1–75, 1978.
- [HUG 82] HUGHES T.J.R., BROOKS A., “A theoretical framework for Petrov-Galerkin methods with discontinuous weighting functions: application to the streamline-upwind procedure”, in GALLAGHER R.H., NORRIE D.H., ODEN J.T., ZIENKIEWICZ O.C., Eds., *Finite Elements in Fluids*, vol. 4, John Wiley & Sons, p. 47–65, 1982.
- [INO 85] INOUE T., WANG Z., “Coupling between stress, temperature and metallic structures during processes involving phase transformations”, *Mat. Sci. Technol.*, vol. 19, p. 845–850, 1985.
- [JAO 98] JAOUEN O., Modélisation tridimensionnelle par éléments finis pour l’analyse thermomécanique du refroidissement de pièces soudées, PhD thesis, Ecole des Mines, Paris, 1998.
- [KAR 72] KARAMARDIAN S., “The complementarity problem”, *Math. Progr.*, vol. 2, p. 107–129, 1972.
- [KEA 88] KEAVEY M.A., “An isoparametric boundary solution for thermal radiation”, *Comm. App. Numer. Meth.*, vol. 4, p. 639–646, 1988.
- [KOI 59] KOISTINEN D.P., MARBURGER R.E., “A general equation prescribing the extent of the austenite-martensite transformation in pure iron-carbone alloys and plain carbon steels”, *Acta Metallurgica*, vol. 7, p. 59–60, 1959.

- [LEB 84] LEBLOND J.B., DEVAUX J.C., “A new kinetic model for anisothermal metallurgical transformations in steels including effect of austenite grain size”, *Acta Materialia*, vol. 32, p. 137–146, 1984.
- [LEB 86] LEBLOND J.B., MOTTET G., DEVAUX J.C., “A theoretical and numerical approach to the plastic behavior of steels during phase transformation, I: Derivation of general relations, II: Study of classical plasticity for ideal-plastic phase”, *Journal of the Mech. and Phys. of Solids*, vol. 34, num. 4, p. 395–432, 1986.
- [LEB 89a] LEBLOND J.B., “Mathematical modelling of transformation plasticity in steels. II: Coupling with strain hardening phenomena”, *International Journal of Plasticity*, vol. 5, p. 573–591, 1989.
- [LEB 89b] LEBLOND J.B., DEVAUX J., DEVAUX J.C., “Mathematical modelling of transformation plasticity in steels. I: Case of ideal-plastic phases”, *International Journal of Plasticity*, vol. 5, p. 551–572, 1989.
- [LEM 80] LEMKE C.E., “A survey of complementarity theory”, in COTTLE R.W., GIANESSI F., LIONS J.-L., Eds., *Variational Inequalities and Complementarity Problems*, Wiley, p. 213–239, 1980.
- [LEM 81] LEMMON E.C., “Multidimensional integral phase change approximations for finite element conduction codes”, in LEWIS R.W., MORGAN K., ZIENKIEWICZ O.C., Eds., *Numerical Methods in Heat Transfer*, John Wiley & Sons, p. 201–213, 1981.
- [LEV 44] LEVENBERG K., “A method for the solution of certain non-linear problems in least squares”, *Quart. Appl. Math.*, vol. 2, p. 164–168, 1944.
- [LEW 87] LEWIS R.W., ROBERTS P.M., “Finite element simulation of solidification problems”, *Applied Scientific Research*, vol. 44, p. 61–92, 1987.
- [LEW 96] LEWIS R.W., MORGAN K., THOMAS H.R., SEETHARAMU K.N., *The Finite Element Method in Heat Transfer Analysis*, John Wiley & Sons, 1996.
- [LOB 93] LOBO M., EMERY A.F., “The discrete maximum principle in finite-element thermal radiation analysis”, *Numer. Heat Transfer, Part B*, vol. 24, p. 209–227, 1993.
- [LOB 95] LOBO M., EMERY A.F., “Use of the discrete maximum principle in finite-element analysis of combined conduction and radiation in non-participating media”, *Numer. Heat Transfer, Part B*, vol. 27, p. 447–465, 1995.
- [LU 95] LU J., LI Y., SUN C., YAMADA S., “A parallel-computation model for nonlinear electromagnetic field analysis by harmonic balance finite element method”, in *IEEE First International Conference on Algorithms and Architectures for Parallel Processing*, vol. 2, p. 780–787, 1995.
- [MEU 08] MEUNIER G., *The Finite Element Method for Electromagnetic Modeling*, ISTE-Wiley, 2008.
- [MIN 06] MINKOWYCZ W.J., SPARROW E.M., MURTY J.Y., *Handbook of Numerical Heat Transfer*, John Wiley & Sons, 2006.

- [MOR 78] MORGAN K., LEWIS R.W., ZIENKIEWICZ O.C., “An improved algorithm for heat conduction problems with phase change”, *Int. J. Numerical Methods in Engineering*, vol. 12, p. 1191–1195, 1978.
- [PAS 03a] PASCAL R., Modélisation du traitement thermique superficiel par induction, PhD thesis, Ecole Centrale de Lyon, France, 2003.
- [PAS 03b] PASCAL R., CONRAUX P., BERGHEAU J.M., “Coupling between finite elements and boundary elements for the numerical simulation of induction heating processes using an harmonic balance method”, *IEEE Transactions on Magnetics*, vol. 39, num. 3, p. 1535–1538, 2003.
- [PON 94] PONT D., BERGHEAU J.-M., ROCHETTE M., FORTUNIER R., “Identification of a kinetic model for anisothermal metallurgical transformations in steels”, in *Inverse Problems in Engineering Mechanics*, Balkema, p. 151–156, 1994.
- [PUM 48] PUMPHEY W.I., JONES F.W., “Inter-relation of hardenability and isothermal transformation data”, *JISI*, vol. 159, p. 137–144, 1948.
- [RAP 98] RAPPAZ M., BELLET M., DEVILLE M., *Traité des matériaux – 10. Modélisation numérique en science et génie des matériaux*, Presses Polytechniques et Universitaires Romandes, 1998.
- [RED 94] REDDY R.W., GARTLING D.K., *The Finite Element Method in Heat Transfer and Fluid Dynamics*, CRC Press, 1994.
- [SHA 85] SHAPIRO A.B., “Computer implementation, accuracy, and timing of radiation view factor algorithms”, *ASME J. Heat Transfer*, vol. 107, p. 730–732, 1985.
- [SIE 81] SIEGEL R., HOWELL J.R., *Thermal Radiation Heat Transfer*, McGraw-Hill, 3rd edition, 1981.
- [SOR 90] SORIA A., PEGON P., “Quasi-Newton iterative strategies applied to the heat diffusion equation”, *Int. J. Numer. Meth Eng.*, vol. 30, p. 661–677, 1990.
- [TER 84] TERRAIL Y.D., SABONNADIÈRE J.C., MASSE P., COULOMB J.L., “Non-linear complex finite elements analysis of electromagnetic field in steady-state AC devices”, *IEEE Transactions on Magnetics*, vol. 20, num. 4, p. 549–552, 1984.
- [TOU 81] TOUZOT G., DHATT G., *Une présentation de la méthode des éléments finis*, Maloine éditeur, Paris, 1981.
- [WAE 94] WAECKEL F., Une loi de comportement thermo-métallurgique des aciers pour le calcul mécanique des structures, PhD thesis, Ecole Nationale Supérieure d’Arts et Métiers, Paris, 1994.
- [YAM 88a] YAMADA S., BESSHO K., “Harmonic field calculation by the combination of finite element analysis and harmonic balance method”, *IEEE Transactions on Magnetics*, vol. 24, num. 6, p. 2588–2590, 1988.

- [YAM 88b] YAMADA S., BESSHO K., LU J., “Harmonic balance finite element method applied to non-linear AC magnetic analysis”, *IEEE Transactions on Magnetics*, vol. 25, num. 4, p. 2971–2973, 1988.
- [ZIE 77] ZIENKIEWICK O.C., KELLY D.W., BETTESS P., “The coupling of the finite element method and boundary solution procedures”, *Int. J. Numer. Meth. Engng.*, vol. 11, p. 355–375, 1977.
- [ZIE 87] ZIENKIEWICK O.C., ZHU J.Z., “A simple error estimator and adaptive procedure for practical engineering analysis”, *Int. J. Num. Meth. In Eng.*, vol. 24, p. 337–357, 1987.
- [ZIE 91] ZIENKIEWICZ O.C., TAYLOR R.L., *The Finite Element Method*, McGraw-Hill, 1991.

This page intentionally left blank

Index

A

- accuracy
 - discretization error, 67
- activity, 232
- admissible
 - functions, 25
- assembly, 52, 110, 209
- Avrami kinetics, 215

B

- bandwidth, 62
- basis function, 46
- boundary conditions
 - Dirichlet, 24
 - essential, 28
 - Fourier, 24
 - Neuman, 24
- boundary element method
 - boundary elements, 260
 - FEM-BEM coupling, 261

C

- CCT diagram, 223
- central difference method, 116
- compact
 - storing, 62
- conform approximation, 47, 89
 - compatible elements, 89
- consistent capacitance matrix, 118
- contact electric resistance, 246

- contact thermal resistance, 250
- convection, 24, 208

D

- diffusion-convection, 169
 - instabilities, 171
 - SUPG method, 174, 209
 - upwind, 174
- direct time integration
 - Crank-Nicolson algorithm, 117
 - explicit algorithm, 116
 - implicit algorithm, 116
 - Runge-Kutta method, 217
 - semi-implicit algorithm, 116
 - trapezoidal rule, 116, 138
- discrete maximum principle, 128, 189
- divergence theorem, 22
 - integration by parts, 22

E

- electric contact, 249
- electromagnetic fields
 - electric current density, 254
 - electric field, 254
 - electric scalar potential, 255
 - magnetic field, 254
 - magnetic flux density vector, 254
 - magnetic vector potential, 255
- element conductance matrix, 50
- element load vector, 50

element matrix, 50
 capacitance matrix, 109
 stiffness matrix, 50
 tangent matrix, 148, 154, 161
 element residual, 49, 108
 enthalpy method, 160, 162

F

Fourier law, 22
 Fourier number, 123

G

Galerkin method, 30, 48, 107
 gas quenching, 228
 global
 numbering, 44

H

harmonic balance method, 261
 heat flux, 21, 155

I

induction heat treatment, 253, 266
 interpolation function, 46
 inverse method, 226
 isoparametric formulation
 actual element, 80
 geometric transformation, 80
 isoparametric element, 86
 reference element, 79
 IT diagram, 223

J

Jacobian, 80, 92
 Jacobian matrix, 91
 Joule effect, 254

L

Lagrange polynomials, 94
 linear system
 conjugate gradient method, 61
 Gauss method, 58
 Gauss-Seidel method, 60
 Jacobi method, 60
 over-relaxation method, 60

linear system
 Choleski method, 60
 load vector, 52
 local numbering, 44
 lumped matrix, 164

M

magnetodynamics, 254
 vector potential formulation, 256
 martensitic kinetics, 217
 Maxwell's equations, 254
 meshing, 43
 elements, 43
 nodes, 43
 modal method, 111
 eigenvalue problem, 112
 modal basis, 112
 modal superposition, 115

N

natural
 boundary conditions, 29
 nodal approximation, 43, 46, 71, 81, 85, 108, 198
 non-linear system
 BFGS method, 151
 convergence test, 145
 fixed point method, 149
 line search method, 152
 modified Newton method, 149
 Newton-Raphson method, 147
 Picard iterative method, 149
 quadratic convergence, 148
 quasi-Newton method, 150
 secant method, 150, 152
 substitution method, 149, 198
 under-relaxation method, 153
 numerical integration
 Gauss method, 93, 250
 Gauss point, 63
 Newton-Cotes method, 95

P

Peclet number, 173

Petrov-Galerkin method, 177

phase change

equivalent specific heat, 158

front tracking method, 157

latent heat, 158, 222

mushy zone, 158

solid state, 215

solidification, 157

state change, 157

physical properties

absorption coefficient, 191

diffusion coefficient, 233

electric conductivity, 243, 255, 256

emission coefficient, 191

emittance, 191

enthalpy, 105, 246

Henry's coefficient, 232

magnetic permeability, 255

magnetic reluctivity, 254

magnetization law, 254, 255

Ohm's law, 255

reflection coefficient, 191

specific heat, 108

thermal conductivity, 23, 109, 153, 221,
246

thermal diffusivity, 133

thermal effusivity, 131

volumetric enthalpy, 158

volumetric mass, 105

Wagner interaction, 232

precipitation, 232

prescribed temperature

elimination method, 54

Lagrange multipliers, 55

penalty method, 55

R

radiation, 24, 156

radiation matrix, 198

radiosity, 191

residual, 25, 51, 106

resolution methods

finite differences, 25

finite volumes, 27

point collocation, 26

sub-domain collocation, 26

result

accuracy, 66

S

setting algorithm, 226

shape functions, 46

skin mesh, 73

skyline storage, 62

smoothing, 64

solution error

accuracy, 67

sparse

storage, 62

stability, 121

Fourier method, 123

von Neumann condition, 124

Stefan constant, 24, 191

Stefan problem, 157

T

tangent matrix, 148, 211, 247

temperature contour, 64

thermal contact, 249

thermal contact resistance, 135

true approximation

continuity conditions, 47

V

view factor, 191, 194

Hottel's crossed strings, 194

numerical integration, 195

volumetric heat source, 155

volumetric source, 21, 109, 245

W

weighting functions, 25