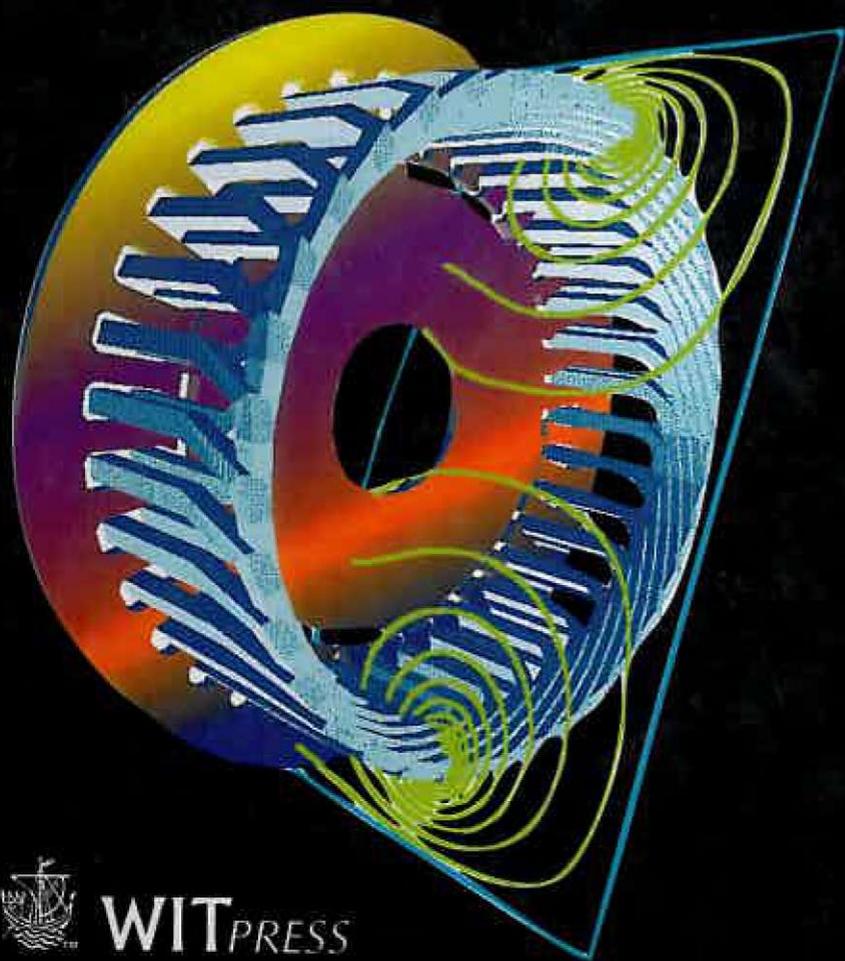# Numerical Modelling and Design of ELECTRICAL MACHINES and DEVICES

K. Hameyer and R. Belmans

WITPRESS

NUMERICAL MODELLING AND DESIGN OF ELECTRICAL MACHINES AND DEVICES

Series: Advances in Electrical and Electronic Engineering - Volume 1

This text provides an overview of numerical field computational methods and, in particular, of the finite element method (FEM) in magnetics. Detailed attention is paid to the practical use of the FEM in designing electromagnetic devices such as motors, transformers and actuators.

Based on the authors' extensive experience of teaching numerical techniques to students and design engineers, the book is ideal for use as a text at undergraduate and graduate level, or as a primer for practising engineers who wish to learn the fundamentals and immediately apply these to actual design problems.

Contents: Introduction; Computer Aided Design in Magnetics; Electromagnetic Fields; Potentials and Formulations; Field Computation and Numerical Techniques; Coupled Field Problems; Numerical Optimisation; Linear System Equation Solvers; Modelling of Electrostatic and Magnetic Devices; Examples of Computed Models.

Titles of related interest:

SOFTWARE FOR ELECTRICAL ENGINEERING ANALYSIS AND DESIGN IV
Editors: A. Konrad and C.A. Brebbia
Software Studies Series
ISBN: 1-85312-684-5   1999   250pp

MICROSIM II: SIMULATION AND DESIGN OF MICROSYSTEMS AND MICROSTRUCTURES
Editors: R.A. Adey and Ph. Renaud
ISBN: 1-85312-501-6   1997   296pp

SENSOR DYNAMICS
O.P. Skobelev
Volume 1 - Pressure Sensor Dynamics
ISBN: 1-85312-433-8; 1-56252-348-1 (US, Canada, Mexico)   1996
Disk/Manual
Volume 2 - Resistive Temperature Sensor Dynamics
ISBN: 1-85312-434-6; 1-56252-349-X (US, Canada, Mexico)   1996
Disk/Manual
Volume 3 - Acceleration, Vibration and Sho
ISBN: 1-85312-435-4; 1-56252-350-3 (US
Disk/Manual   SET ISBN: 1-85312

Numerical Modelling and Design of Electrical Machines and Devices

Advances in Electrical and Electronic Engineering

Editorial Board

J K Sykulski (Series Editor)
University of Southampton
University Road
Highfield
Southampton
Hants SO17 1BJ
UK

P Vas
Dept of Engineering
Unversity of Aberdeen
Kings College
Aberdeen
AB9 2UE
Scotländ

N Takahashi
Dept of Electrical Engineering
Faculty of Engineering
Okayama University
Okayama 700
Japan

K Zakrzewski
Instytut Maszyn Elektr i
Transformatorow
Politechnika Lodzka
Stefanowskiego 18/22
90-924 Lodz
Poland

T Tsiboukis
Dept of Electrical and Computer
Engineering
Aristotle University of
Thessaloniki
54006 Thessaloniki
Greece

Numerical Modelling and Design of Electrical Machines and Devices

Kay Hameyer

&

Ronnie Belmans

Katholieke Universiteit, Belgium
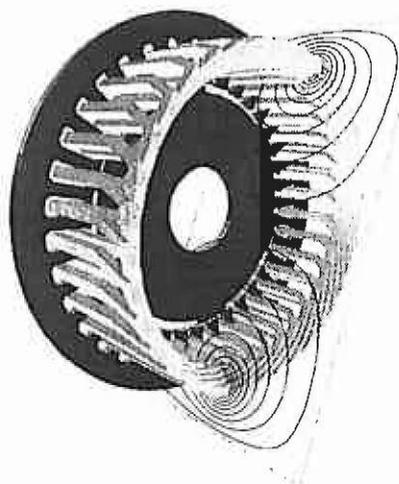
WITPRESS **Southampton, Boston**

Kay Hameyer
Ronnie Belmans

Contents

Preface

This book provides an overview of numerical field computational methods. Particular attention is paid to the finite element method (FEM) for the design of electrical machines and other magnetic devices.

It is based on the authors' experience in teaching numerical techniques to undergraduates, graduates and doctoral students in courses at their own and at various international universities, e.g. guest courses at COPPE Universidade Federal do Rio de Janeiro (Brazil), Université de Batna (Algeria) and the RWTH Aachen (Germany). The numerical techniques are introduced to engineers from industry in an annual *Short Course in Magnetics* organised at the Katholieke Universiteit Leuven.

This book is intended to be the basic reading material for such courses on numerical field computations as given nowadays at the Electrical Energy division in the Electrical Engineering Department of the Katholieke Universiteit Leuven and as guest courses at other universities.

The book describes the theory and techniques of modelling and simulating electromagnetic devices. While its primary focus is on the techniques applicable to the modelling of electrical machines and the electromechanical energy transducers, it also illustrates the usefulness of knowing the physical background of the specific problems. Accordingly, the right problem definitions can be applied and computed results can be interpreted and verified in a proper way.

Particular attention is paid to the FEM in designing electromagnetic devices, such as motors, actuators and transformers. This means that only frequencies below 10 kHz are considered. Several aspects of coupled fields are discussed in sections where the physical problem urges coupled solutions. The book has been written as a text book for undergraduates, graduates and engineers in practice who want to learn how to apply the fundamentals to solve electromagnetic design problems. Selected examples to develop skills to define and solve a field problem accurately are given at the end of this book. In parallel with the text for this book, an in-house software package for two dimensional and

axisymmetrical FEM problems is developed to investigate the various aspects of the numerical field computation and their usefulness.

Without the support of various people, this book would not appear in this shape. In the first instance we should like to thank especially Dr. U. Pahner for his contributions to the optimisation chapter. The texts from section 7.3.1 to 7.3.3 are slightly adapted parts of his excellent Ph.D. thesis.

The authors are grateful to all their Ph.D. students currently working in the area of field computations for their suggestions, criticism and active participation collecting this course material. Some topics introduced in this book are directly related to the current research work in the division ESAT/ELEN.

As a guest professor and regular visitor to our division, Prof. E.M. Freeman contributed to general ideas, which we have developed over the years as well.

It has been a pleasure working with Lance Sucharov and Clare Garcia of WIT Press.

Kay Hameyer and Ronnie Belmans
March 1999, Leuven

To
Erna

Mieke, Nathalie & Wim

1 Introduction

An original design and the step by step optimisation of physical technical devices is in practice often a trial and error process. During the design and construction of a device several expensive prototypes have to be built to monitor and check the mathematical approximations and the physical reality. This procedure is time consuming and expensive. Successful industrial developments demand shorter cycle times to fix or improve the economical competition of particular companies. To effectively compete in the market place nowadays, developed products of higher quality, improved efficiency and better functionality are recommended, leading to devices with very complex geometries. Furthermore, custom designs are becoming very important. The added value of standard massproduction devices is far lower. To solve the techno-economical demands, the idea is to replace the expensive prototyping by numerical simulations.

If an appropriate simulation model is found, various operating points can be simulated on a computer. Even the behaviour of the device for hazardous situations that cannot be measured inside a laboratory and the use of arbitrary even future materials can be studied. The appropriate choice of a calculation technique for an electromagnetic device is always closely linked to the complexity of the problem.

To develop a technical product, parasitic effects such as:
- ferromagnetic saturation
- increased leakage flux
- high operating temperatures
- irreversible flux losses by using permanent magnet materials at elevated temperatures
- coupling between different effects such as thermal-magnetic-mechanical-flow field problems and
- induced currents due to motion effects

have to be considered in the calculations accounting for sufficient accuracy. In devices with complex geometries, those effects can not be

treated by a classical analytical approach. Results with a high accuracy are required to predict the behaviour of the technical product. In this case the simulation of the electromagnetic fields and their effects by numerical models is suitable as an appropriate engineering tool. Using computer models and the appropriate numerical algorithms solves the physical problem. The numerical method has to fulfil specified demands such as:

- reliability
- robustness
- application range
- accuracy
- performance.

To see where the numerical simulation finds its place in the analysis of technical devices, Fig. 1.1 shows the links between the real technical device, the classical physical theory and the numerical simulation. This figure makes obvious that the numerical simulation is a connecting element between reality, measurements, and theoretical predictions. As a consequence, all numerical computations represent realistic activities in a fictive laboratory. This means that simulation results should be theoretically measurable in practice. The numerical simulation is in fact an experiment performed on the computer as a fictitious laboratory, where the engineer is using numerical tools to perform the experiments instead of measurement devices such as current, voltage, power, temperature and force meters.
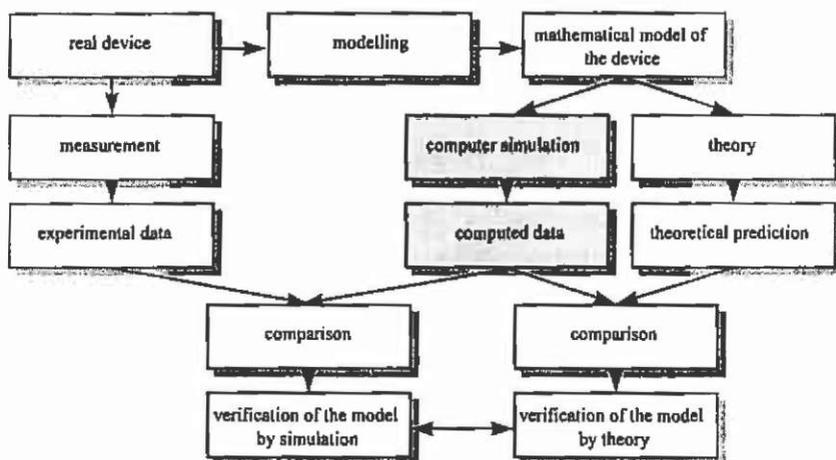
Fig. 1.1. Theory, experiment and simulation.

The numerical simulation influences the analytical theory where sometimes rough approximations or constants are used to consider physical effects such as ferromagnetic saturation or hysteresis. The verification of numerical solutions and results obtained by the analytical theory can lead to improved analytical models and vice versa. Both numerical simulation and analytical theory help to understand the physical reality and to improve technical predictions.

1.1 Numerical solution process

In Fig. 1.2, the solution process for a system of partial differential equations is outlined.
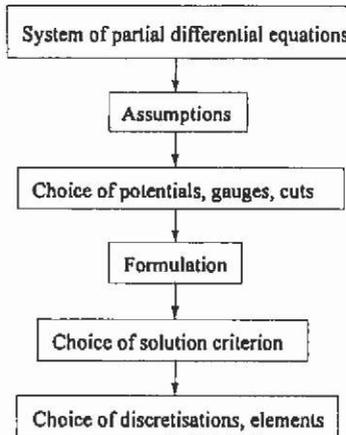
Fig. 1.2. Solution process for a system of partial differential equations.

The fields are described by differential equations. Assumptions concerning boundary conditions, material properties such as isotropy, dependencies in time, etc. have to be made before a computation of a field can be performed. For example in magneto-static fields, the time derivative is assumed to be zero and therefore no induced currents can be considered.

The choice of the potentials is based on these simplifications. For each problem type, the choice of an appropriate potential is different. The choice of a gauge is necessary to obtain a regular system of equations. Using the finite element method, the choice of the gauge also determines the choice of the element type. However, the user of a CAD program package that simulates magnetic, electric or thermal fields is usually not involved in choosing for such basic numerical properties.

The numerical method to solve the partial differential equation is understood as a solution criterion.

The appropriate solution method depends on the type of equation, such as parabolic, hyperbolic or elliptic.

For example, the choice of the elements for the finite element method depends on the differential equation, the potential formulation, and the solution method.

In a two-dimensional magnetostatic problem the unknowns are node potentials. Here, the magnetic vector potential is chosen because the nodal unknowns have only a single component A_Z. In this two-dimensional field problem, the Coulomb gauge is satisfied automatically.

The choice of method for solving a system of linear equations is dependent of the differential problem and its formulation. For example the magneto-static problem is an elliptic differential problem. The Laplace operator is symmetrically adjoint and positive definite. A system of equations with such properties can be solved by a conjugate gradient method.

To focus on the active parts performed by an design engineer, in principle, field computation is performed in three major steps: pre-processing, processing and post-processing. Fig. 1.3 shows a typical pattern for the FEM approach. The first step consists of the definition of the geometry of the electromagnetic device. Material properties, electrical current densities and boundary conditions are defined. All the activities have to be performed by the design engineer. Therefore, the pre-processing is time consuming. The estimated time expenditure for a two-dimensional problem is given in Fig.1.3. The processing, i.e. the solution of a very large system of equations is automatically done in the second step. Only parameters to control the solution process have to be defined by the design engineer. In the last part of the FEM procedure, the interesting field quantities are computed from the solution out of the processing. If the geometrical data can be parameterised, the pre- and post-processing can be automated as well. This represents an important prerequisite for the possibility of the combination of field computation and numerical optimisation.

Fig. 1.3. Solution processes during a field computation session.

2 Computer aided design in magnetics

For designing and constructing electromagnetic devices an accurate knowledge of the field quantities inside the magnetic circuit is necessary. In many cases the air gap is of particular importance (e.g. motors, switches, relays, contactors, actuators). Here the conversion from electrical to mechanical energy and vice versa takes place. In the air gap the field quantities such as flux density and field strength have to be calculated very accurately in order to be able correctly to asses the operational behaviour of the device.

Although Maxwell equations have been known for more than a century, in the past the task in calculating a magnetic circuit was to find as many assumptions and simplifications as possible. Then, results could be obtained with rather low numerical efforts. Using this approach, only devices or problems with a strongly simplified geometry could be studied. It was a design following simple rules, found empirically. Physical effects were considered by correction factors applied to the existing rules. In the following period of time this design through rules has changed into another design philosophy: design analysis. Here, computer models were used to solve the field problem. Analysis means the treatment of the field problem by numerical simulation.

With the ongoing developments in computer hard- and software and numerical research, difficulties concerning computational costs and numerical problems are continuously moving to the background. Today, efficient numerical solutions can be obtained for a wide range of problems beyond the scope of analytical methods. In particular the limitations imposed by the analytical methods, their restrictions to homogeneous, linear and steady state problems can be overcome using numerical methods.

In general, the procedure for analysing an electromagnetic device can be divided into three steps:

- pre-processing
- processing and

input data:
- geometry
- material
- boundary
  conditions

input data:
- error bound
- max. iteration steps

input data:
- evaluation locality for
  diagrams, colour plots, ...

- discretisation
- approximation
- parametrisation
- coupling:
  - ⇒ fields
  - ⇒ geometry
  - ⇒ circuits
  - ⇒ motion
  - ⇒ methods

- mesh adaptation
- num. method
- equation solver

- optimisation
- further modelling
  - ⇒ lumped parameter
- approximation of local field
  quantities
- field coupling

Fig. 2.1. Field analysis steps.

• the post-processing.

In the first step, the field problem is defined and prepared to be solved. The second step delivers the numerical solution of the physical problem. During the post-processing, the obtained solution is prepared to calculate the required field quantities or to evaluate forces and other macroscopic quantities. This threefold approach of defining, solving and evaluating is typical for every analysis procedure, numerical or analytical. The different techniques, data structures or algorithms used in the individual steps, influence and/or limit the overall procedure during the analysis of a field problem (Fig. 2.1).

To define a field problem, the input data describing the geometry of the domain of interest, the material representation and the boundary conditions are always required. Even with enhanced CAD drawing techniques, most of the analysis time will have to be spent on the pre-processing. Given error bounds will support a desired accuracy of the solution. Often, the user can not influence this step. During the post-processing, the solution must be prepared to study the local field effects. The post-processing represents an open-ended process, because the user of the analysis can evaluate the calculated solution in various ways for different aspects.

The methods and algorithms used in the single steps of the overall procedure can form an efficient analysis or design tool and determine the quality of the results of the analysis. For example a use of particular internal data structure can enable very quick search routines to obtain an efficient, fast and automated discretisation with parameterised geometries and materials. The various possible coupling mechanisms of different fields, circuit equations, methods such as FEM/BEM combinations, motion term or geometries yield into an accurate approximation of a realistic physical problem. The properties of the coefficient matrix decide which equation solver or algorithms must be used to solve the problem.

2.1   Finite element based CAD systems

CAD systems to treat two-dimensional field problems are in common use nowadays. Developments in hard- and software have made it possible to realise user-friendly and reliable systems running on different hardware platforms such as UNIX and on PCs. Commercially available software is oriented mainly to operate inside a PC Windows™ environment. A reason for this can be seen in the price/performance development of the PC market in the past years. Unix workstations remain more expensive when compared to the PC competitor and the performance of standard PCs is already comparable to UNIX workstations. A growing demand on PC-CAD systems for magnetics can be noticed. This tendency in the market can identify several reasons. The lifetime of hard- and software is decreasing. Having user-friendly software available enables the user to change the CAD system without large training efforts more quickly. It turns out that software systems are becoming consumables. Observing recent years, the CAD software price developments have been rather calm, but the performance characteristics of the software, such as solver speed and user-friendliness, are rapidly increasing. Using commercial software can solve more and more complicated problems:

- complex geometries in 2D/3D
- complexity of the analysis increases
- external circuits including capacitances and inductivities, voltage and current driven
- motion effects
- enhanced force computation, local field quantities
- enhanced mesh adaptation
- coupled field analysis.

2.1.1   Components and modules

A regular CAD system for magnetics includes various components to be able to solve the field problem in an appropriate way. To handle the problem in the three steps of pre-processing, solving and post-processing, modules such as:

- graphical drawing tool to generate the geometry
- mesh generator
- material library manager and modeller
- problem definition tool
- different solver modules for the various field types and formulations
- post-processor tools
- visualisation modules to evaluate the solution and
- file manager tool for the data transfer to other software modules

are pre-requisites of a FEM CAD system. User interfaces are recommended to have a maximum of process control for a minimum of efforts during the analysis. To obtain a high quality field analysis tool, the user-interfaces must allow sufficient interaction of the user with the process steps. The influence of the interface on the mesh, problem formulation and solution can be taken from Fig. 2.2. An open data base interface allows further manipulations of the data by other software products, for example to generate different graphical representations of the solution or to analyse FEM models generated by another software package.
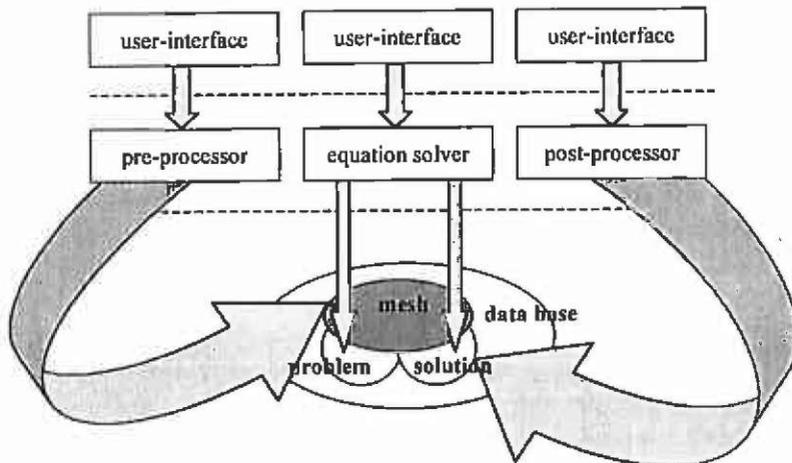
Fig. 2.2. Process control of a field analysis.

2.1.2 Graphical drawing tool for problem definition

To be able to model the technical device to be studied by field analysis, a graphical interface is necessary to generate a technical drawing. This user-interface must support primitives such as lines, arcs, circles and points to describe the geometry properly. If different software packages are used, interfaces to the other software should be available (IDEAS, PATRAN, AutoCAD DXF, …).

Using the generated technical drawing, boundary conditions are set and domains with different material properties are defined. Various labels represent the chosen material of particular domains and set edges of the geometry to characterise the given boundary conditions there. Commercial program packages are supporting libraries with various grades of non-linear ferromagnetic and hardmagnetic materials (Fig.2.3). To define own materials, special software modules can be used to include such data. The defined data can be controlled visually after their definition.

a)   b)
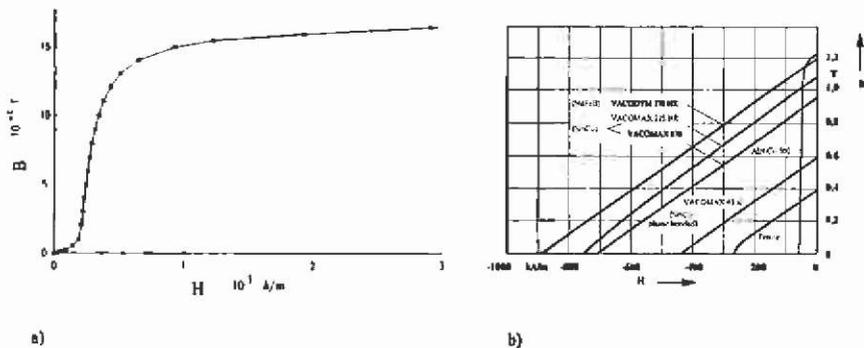
Fig. 2.3. Typical material representations: a) non-linear ferromagnetic and b) permanent magnet material characteristics.

If semi-automated mesh generators are used, mesh size definitions have to be given in this step of problem preparation. Therefore, lines, circles and arcs are selected and subdivided into several parts to form edges of the finite elements to be generated in the next step.

2.1.3 Mesh generation in general

A mesh generation module must supply a numerical discretisation of interior regions by the finite elements. Standard triangular elements are in common use in two-dimensional models (Fig. 2.4) and tetrahedrons are regularly used to model three-dimensional field problems.
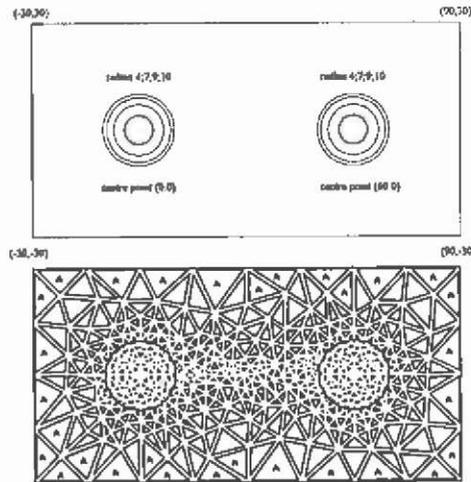
Fig. 2.4. Domain of interest of two conductors in air and the two-dimensional
FEM mesh.

The mesh generation is automated with a minimum of user interaction. Several a-priori criteria can be employed to guarantee a certain quality of the discretisation. The solution accuracy is strongly dependent on this mesh. If the strategy of an automated mesh adaptation is supported to enhance the quality of the discretisation in successive computation steps, only a minimum discretisation is recommended in the first mesh. More details can be found in the section on adaptive mesh refinement.

The generation of three-dimensional FEM models is extremely time consuming. Two different strategies can be followed:

- mesh extrusion (Fig. 2.5)
- solid modelling (Fig. 2.8).

The extrusion approach works with two-dimensional meshes extruded in the third direction. Rotations of axis-symmetrical geometries are possible to form the 3D model as well. A disadvantage is that not every contour can be modelled realistically. For example, conical surfaces represent a problem (Fig. 2.7). If the scalar potential formulation is used, excitation coils and windings can be introduced into the model in a following step (Fig. 2.6).

A solid modeller works mainly in two steps. First the surface of the geometry is discretised, and after this the volume is meshed in a second step. Various suggestions to generate solid meshes can be found in the literature (Tsukerman & Plaks [114]).
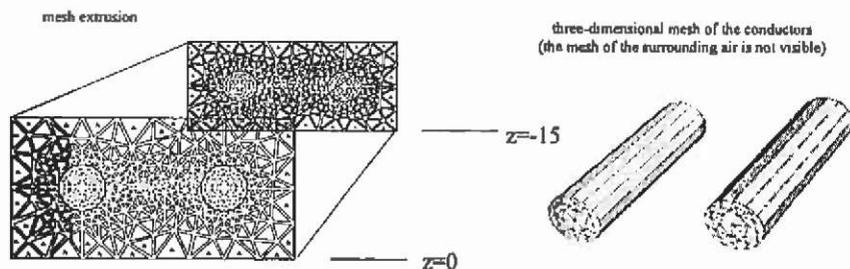
Fig. 2.5. Basic idea of the extrusion technique.
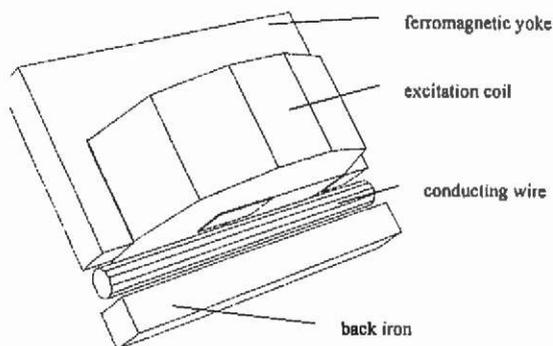
Fig. 2.6. A three-dimensional model of a wire heating device.

a)   b)
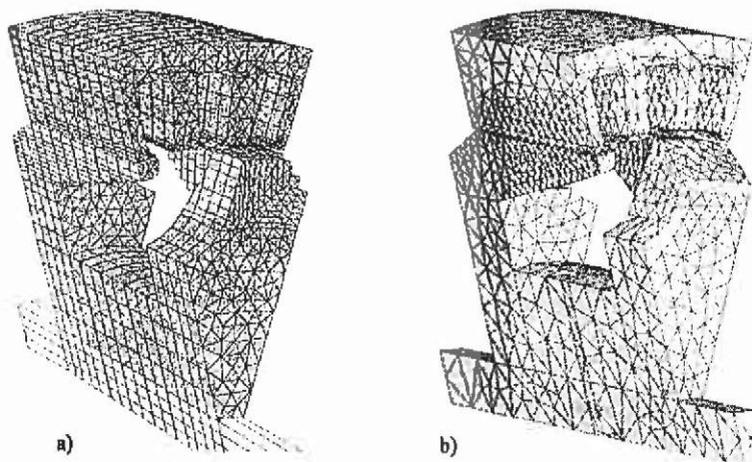
Fig. 2.7. Comparison of a) extrusion based and b) solid modelling of a claw-pole generator (source: IEM RWTH-Aachen).
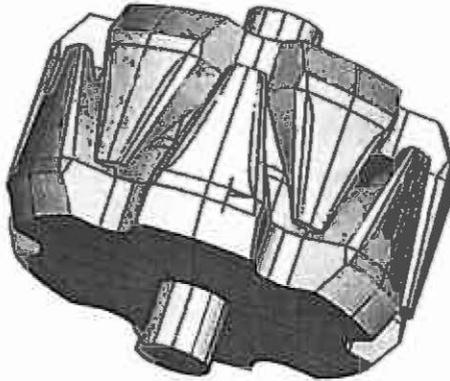
Fig. 2.8. Solid modelled 3D mesh of the rotor of a claw-pole generator.

(Source: IEM RWTH-Aachen)

After the mesh generation the field problem is defined and can be solved by the equation solver. This is mainly in commercial software packages performed without or with a minimum of user interactivity. The appropriate solver has to be chosen.

2.1.4   Post-processor tools

Several tools are recommended to evaluate the field solution. The potential solution has to be transformed into physical quantities such as flux density, field strength or forces. Therefore, numerical manipulations of the potential are necessary. The post-processor module must consist of a calculator to perform such manipulations.

To be able to evaluate the solution, various graphical representations of the solution can be of interest:

- colour plots of selected quantities
- plots of the lines of constant potential, flux plots
- diagrams showing quantities along defined contours.

To extract parameters out of the solution, a post-processor calculator can be used as well.

2.2   Design strategies

The development and design of electromagnetic devices reflects a complex process. Originating from an initial idea, the construction runs through different phases. This procedure is terminated when a final concept is selected and considered to be designed, subject to various targets and constraints. As a whole, the task of the design engineer is to find solutions for technical problems. On the way to the latter physical

and technical product, certain aspects have to be considered. Technological and material-dependent questions as well as cost effectiveness and ecological constraints have to be taken into consideration. A cut-set of the mentioned boundary conditions controls the feasibility of the final design. With emphasis on electromagnetic devices, Fig. 2.9 shows a simplified scheme of interdependencies of targets and constraints. This simple pattern clarifies that the design process is strongly dependent on the experience of the engineer and reflects an optimisation procedure with often contradictory aims. Therefore, the necessity of a systematic and strategic design with engineering tools is obvious. Here, solution strategies using modern numerical methods to accelerate and ensure a high-standard technical product in an overall design process are discussed.
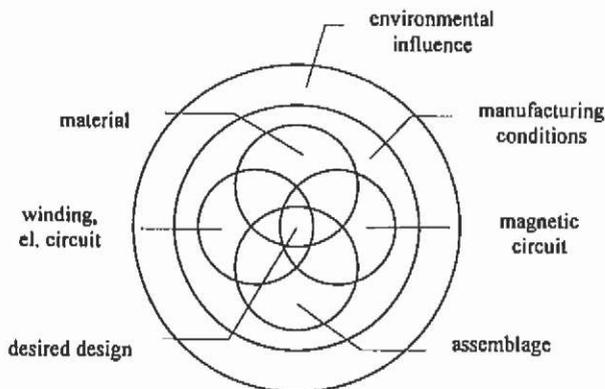
Fig. 2.9. Interdependencies in the design of electromagnetic devices.

Designing electromagnetic devices includes the calculation and analysis of the electromagnetic field distribution. From the local field quantities forces, torques and losses can be derived to make predictions concerning global quantities such as converted power and efficiency. For complicated geometries analytical field solutions are non-existent or very hard to obtain. Using numerical field computation techniques of a general application range, the microscopic field solution leads via a lumped parameter approach to the desired time-dependent behaviour of the device (Fig. 2.10).

The microscopic field solution itself delivers important knowledge regarding the material utilisation. Such results offer the opportunity to reduce material, weight and the costs of the latter product. To accelerate development, extensive field computations with various types of material can be performed avoiding expensive prototyping. It is even possible to

predict system behaviour before new materials are actually available on the market. With this knowledge, the design engineer can order special material to be developed at the material manufacturer or, vice versa, if the material supplier uses such numerical tools he can suggest and offer the right choice of material for a particular device.
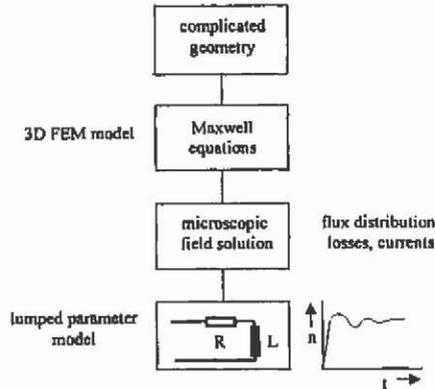
Fig. 2.10. Analysis scheme using the finite element method.

Lumped parameter models are essential for the development of control strategies for electromechanical devices such as electrical drive systems. To be able to perform real time control schemes, lumped parameter models are used to form an observer control. Here, very accurate field computations are recommended to determine the concentrated elements of such models.
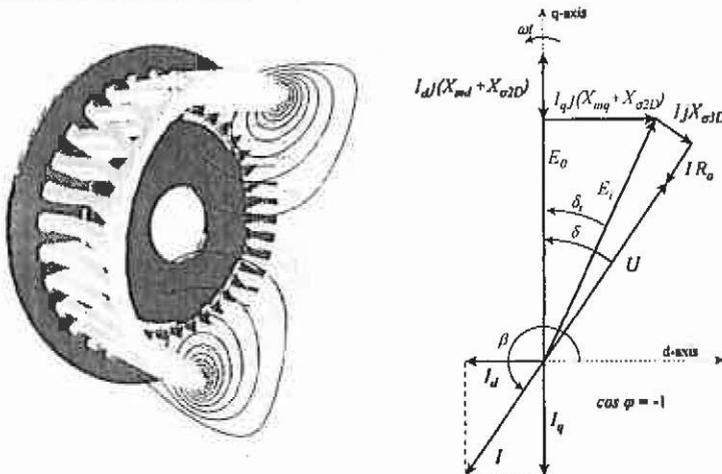
Fig. 2.11. FEM model of the end-winding area to compute the leakage reactance $X_{\sigma 3D}$ of a servomotor with d-q model in vector diagram representation.

For example, the computation of the leakage reactance of electrical motors can be performed using a three-dimensional FEM model (Fig. 2.11). The knowledge of this reactance is essential for the optimum control of a permanent magnet-excited servomotor. The vector diagram in Fig. 2.11 demonstrates the large influence of the leakage reactance $X_{\sigma 10}$ on the optimum control angle of this permanent magnet machine.

2.2.1 Knowledge-based design

The main aspect of the structured development of novel technical products is analysis followed by a detailed synthesis. Analysing means obtaining information on partial functions of the desired overall function, by investigating single elements and their mutual interactions. In this way overall links between various principles of the partial functions are found.

In Fig. 2.12, a structured and knowledge-based development process is illustrated in a simplified scheme. In this example, the final technical product has to be designed, able to fulfil three partial functions. Those individual functions to be connected to the overall task of the product are a linear motion, a continuous rotation and some reverse operation. After the analysis phase, in the synthesis step different physical working principles are selected and evaluated. The selection process is governed by simple qualitative rules. In this way the partial functions are evaluated with regard to their feasibility with respect to the given constraints and limitations (Fig. 2.12). In this step, the feasible principles are ranked qualitatively by weighted constraints and limitations.
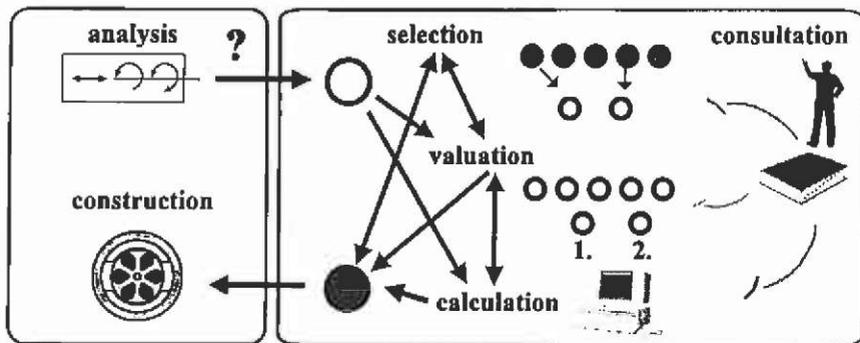
Fig. 2.12. Knowledge-based and structured design.

The process of synthesis (Fig. 2.13) leads from qualitative decisions to quantitative statements in a following design step. The whole process is accompanied by the consultation of experts and expert knowledge (Fig. 2.12). A detailed investigation and ranking, i.e. the precise calculation of the operating conditions, leads by a comparison to the final technical product. In this loop of iterations, between validation and the performance of detailed predictions of qualified concepts, a numerical optimisation combined with field computation methods is found as an important and powerful engineering tool for the design of electromagnetic devices.

The quantification and ranking of the working principles is governed by the choice of materials or other components such as electronic circuits. Their interdependency on the studied principle can be distinguished into an object and a rule world. The various, for instance ferromagnetic, permanent magnet, conductive or dielectric materials, and respectively components such as the electronic hardware have, considered inside an object world, particular properties and characteristics. To employ such object properties in order to obtain a physical working principle fulfilling a desired function, appropriate rules determining the function of the object have to be considered. In both object and rule world, constraints are found to govern a decision to consider the principle further in the ranking or to reject it. Numerical techniques can help to employ the rules accurately to the studied object.
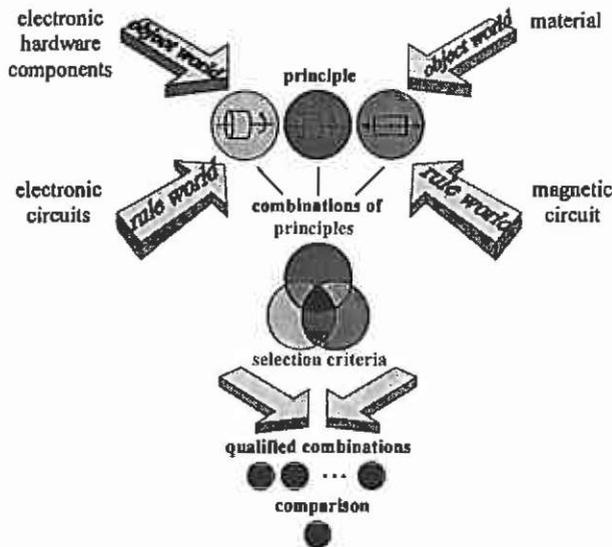
Fig. 2.13. Process of synthesis.

3 Electromagnetic fields

The notation and basic laws of the electric and magnetic fields are explained in this section. It is not intended to present the complete electromagnetic field theory. Only a limited set is given, necessary to understand the types of physical problems treated in this book, enabling the modelling of technical devices to be studied by numerical simulations of such fields. In this book, fields used for energy conversion are discussed only. The high frequency fields that, except those for microwave heating, are used to transfer information are not considered.

3.1 Quasi stationary fields

In general, two classes of electromagnetic fields can be distinguished, the time independent static and time varying fields (Fig. 3.1). They can be scalar and vector fields. A typical scalar field for example is the electrostatic potential distribution $\varphi(x, y, z)$ between charged electrodes
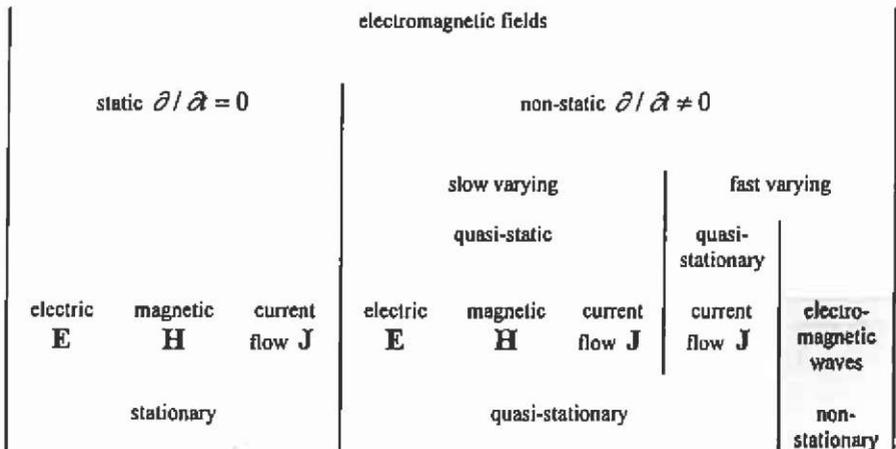
Fig. 3.1. Classification of electromagnetic fields.

and the magnetic field strength $\mathbf{H}(x, y, z)$; surrounding a current
carrying conductor is a typical vector field. In the group of slow varying
fields we can find the same types.

We have to distinguish between the slow and fast varying electrical
current flow field with regard to the geometrical dimensions of the
current carrying conductor. The slow varying fields are understood to be
fields not leading to current redistributions. This means that there are no
eddy current effects as the dimensions of the current carrying conductor
are smaller than the penetration depth of the field. The current at those
frequencies is distributed as in the DC case, uniformly over the whole
surface of the conductor. Eddy current effects are considered in the fields
with fast varying time dependency, due to the low frequency treated as
quasi-stationary. High frequency fields as focussed in antenna problems,
leading to the electromagnetic waves, are not considered in this book.

Most of the physical issues in electrical energy engineering can be
described by quasi-static phenomena. Slowly varying and periodic fields
up to 10 kHz are considered to be quasi-stationary. Electrical energy
devices such as electrical motors and actuators, induction furnaces and
high-voltage transmission lines are operated at low frequency.
Exceptions are microwave devices for electroheat applications, where
inherently the displacement current is not negligible.

Typical examples of quasi-static electromagnetic fields are the
fields excited by coils in rotating electrical machines, transformers and
inductors. Inside these conductors the displacement current is negligible
and the magnetic field $\mathbf{H}$ outside the coil is exclusively excited by the
free current density $\mathbf{J}$. For those quasi-static fields, AMPERE'S law is
applicable (Binns et al.[13]).

$$\nabla \times \mathbf{H} = \mathbf{J} \tag{3.1}$$

To decide whether the displacement current can be neglected or not,
depends on the wavelength λ of the problem considered in the frequency
domain. If it is large, when compared to the physical dimensions of the
problem l, the displacement current is negligible. To consider this
phenomenon in the time domain, the rise time T_a of a step function must
be large inside the problem compared to the transit time l/v. Field
problems are quasi-static if eq.(3.2) is valid.

$$T_a \gg l/v$$
$$\lambda \gg l \tag{3.2}$$

In general $T_a \approx 5...10 l/v$, respectively $\lambda \approx 5...10 l$ is sufficient.

For this class of problem, the interesting fields vary slowly and can
be periodic. Then three categories of problems are distinguished:

- static
- slowly varying transient
- time-harmonic eddy current.

In time-harmonic problems sinusoidal varying field quantities are assumed. In theory, a time-harmonic solution is only valid for a linear system as a sinusoidal excitation does not yield a single frequency response in the non-linear case.

3.2   Boundary value problem

Many scientific engineering or physical problems lead to boundary value problems. The describing differential equations have to be solved in a volume satisfying particular conditions on its boundary Γ (Fig. 3.2). Therefore, the definition of a boundary value problem is necessary and evident. The proper definition of a numerical model is important to obtain correct results and assumes a good understanding of the underlying physical background of the field problem.
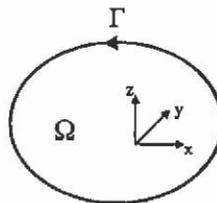
Fig. 3.2. Boundary value problem.

For obtaining the solution of the boundary value problem, it can be formulated in the form:

Find a function $u \in V$, so that for all $v \in V$
$$a(u,v) = f(v).$$
(3.3)

The boundary value problem is defined by a differential equation $a(u, v)$ feasible in the volume Ω. u is the exact solution of the problem that has to be found.

V is a set of continuously differentiable functions in Ω with for example $v=0$ at the boundary Γ. The type of functions v and differential equation that can be employed is subject to the method or approximation used to solve the field problem.

Assuming that the appropriate differential equation for a particular physical problem is known, the definition of the numerical model is dependent on the correct choice of boundary conditions.

If it is considered that the desired field solution is a local potential distribution $u = \phi(x)$ in a local co-ordinate system and time depending $x = (x_1, x_2, ..., x_n, t)$, different types are possible and can be defined.

3.2.1   Initial and boundary conditions

Boundary conditions have to be applied to a field problem to ensure a well posed problem with a unique solution. Particular attention has to be paid to the Dirichlet and Neumann boundary conditions. Applying these boundaries in an appropriate way reduces the size of the field problem significantly. On the one hand, therefore, the accuracy of the solution can be improved with the same computational expenses; on the other hand an enlarged domain can be studied resulting in the same solution accuracy.

Mainly finite element program packages are limited in the number of elements to approximate the geometry of the problem or the computational efforts must be limited in order to obtain acceptable computation times. Therefore, the correct and appropriate application of the boundary conditions is the key to defining field problems and to allow an accurate solution in an efficient way.

3.2.1.1 Starting conditions When a differential problem covers the time domain, the starting conditions are quantities valid at simulation start-time, for example the velocity, flux couplings with windings in electrical machines, field exciting currents or voltages that have to be set and defined in order to find the solution of a transient problem formulation.

3.2.1.2 Dirichlet boundary condition A Dirichlet boundary condition sets the unknown function to a known function on the boundary of the differential problem.

$$\phi(x) = g(x) = const. \tag{3.4}$$

Fig. 3.4 shows in a typical example the application of the Dirichlet boundary condition of an electromagnetic problem. A ferromagnetic circuit is shown consisting of a U-shaped permanent magnet, an air gap and ferromagnetic back iron. Lines of constant vector potential represent the flux lines. Physically, the field is assumed to be zero at a sufficiently large distance from the magnetic circuit. Therefore, a Dirichlet boundary condition, the potential set to zero, is applied at the entire boundary of the problem. Due to eq.(3.4) it is impossible that flux lines can cross the boundary Γ. Fig. 3.4b shows the potential lines of an electrostatic field with Dirichlet boundary. Here, the potential distribution excited by the charged plates of a capacitor is computed.

$\phi(x) = 0 = const.$

a)

b)

$\phi(x) = 0 = const.$

c)

d)

e)

f)
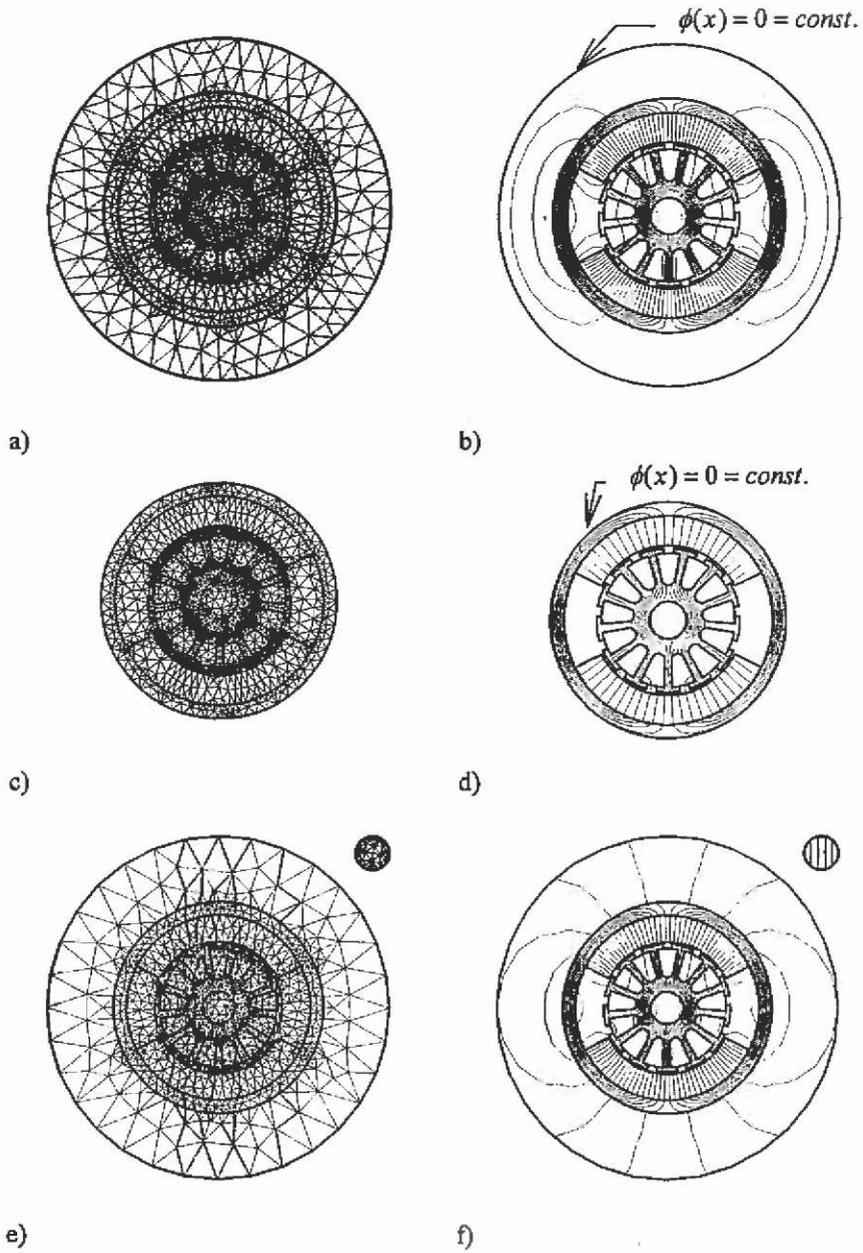
Fig. 3.3. Dirichlet boundary condition applied to the FEM model of an electrical machine. a) Mesh of the model accounting for flux relief, b) the flux lines, c) the mesh neglecting the flux relieve of the machine, d) its flux plot, e) the mesh applying the Kelvin transformation, f) its flux plot.

The Dirichlet boundary is an essential boundary condition as it does not characterise the space V . It is sometimes called the boundary condition of first kind.
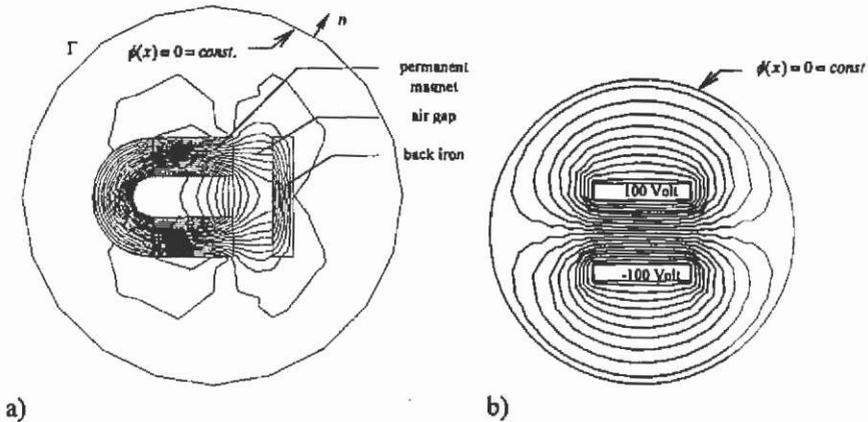
Fig. 3.4. Dirichlet boundary condition a) for an electromagnetic and b) for an electrostatic problem.

For the analyst of a field problem, a crucial question is how far away the Dirichlet boundary condition has to be applied from the field exciting sources to restrict the field-domain on the one hand and to have an accurate overall solution of the near and far fields respectively on the other hand. When analysing electromagnetic fields in the presence of ferromagnetic material and small air gaps, such as in electrical machines, an outer diameter of roughly 20% above the characteristic diameter of the device can be applied to compute the field inside the device accurately (Fig. 3.3a). Flux lines can not pass the Dirichlet boundary. If the flux relief inside a ferromagnetic core due to saturation can be neglected, the outer diameter of the back iron yoke of electrical machines can be represented by the Dirichlet boundary condition (Fig. 3.3c). If the flux outside the machine yoke can be neglected, the number of elements in the numerical model and thus the computation time, decreases.

If the far field is analysed, a diameter of up to 5 or 6 times the characteristic dimensions of the device should be used or special transformations, such as the Kelvin-transformation, an open boundary condition (Fig. 3.3e/f), can be employed to terminate the field in the transformed infinite distance.

3.2.1.3 Neumann boundary condition The next important boundary condition is the Neumann boundary condition. Here, the known value of

the derivative of the unknown function in the normal direction of the boundary Γ is prescribed.

$$\frac{\partial \phi(x)}{\partial n} = g(x) = const. \tag{3.5}$$

If the derivative in normal direction is constant, lines of constant potential can pass the outer boundary Γ of the studied domain Ω.

The most important property of this type of boundary condition is that by knowing symmetries of the field, and applying the Neumann boundaries there, the numerical model can be reduced to obtain the same solution of the problem. In this way, the problem size, the time to prepare the field problem and the computational efforts can be reduced significantly. On the other hand, if less than the complete geometry has to be defined and discretised, a higher accuracy is achievable for the overall solution of the problem without extra effort. Therefore, particular attention must be paid to this boundary condition.

Fig. 3.5a shows the electromagnetic field problem with applied Neumann boundary at the symmetry line of the U-shaped permanent magnet. This is the only symmetry inside this model and yields a problem reduction of 50% with respect to the accuracy of the problem solved in Fig. 3.4.

The Dirichlet boundary remains of course at the outer diameter of the domain studied.
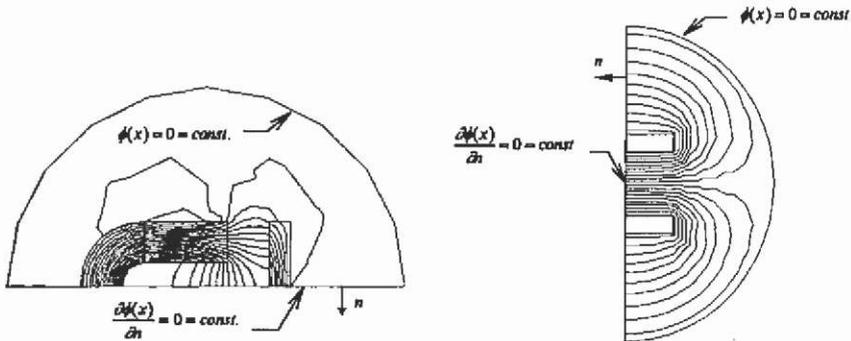
Fig. 3.5. Neumann boundary condition applied at the line of symmetry.

By looking at the electrostatic example of the capacitor in Fig. 3.5, an additional symmetry in the potential distribution attracts attention. Knowing the potentials at the electrodes of the capacitance, +100 V at one side and -100 V at the other, the Dirichlet boundary with a constant potential of 0 Volt can be applied, reducing the problem size a second time (Fig. 3.6).
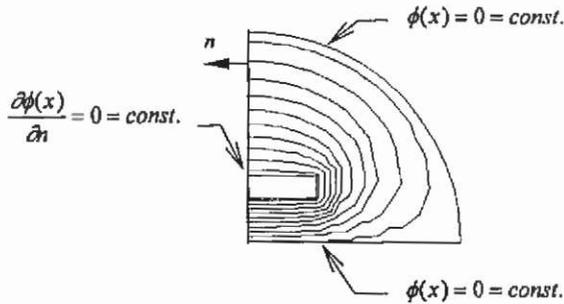
Fig. 3.6. Lines of symmetry with appropriate boundary conditions.

The Neumann boundary is a natural boundary condition as it does not influence the definition of the space V. It is automatically satisfied at the boundary and is sometimes called the boundary condition of second kind (Zienkiewicz & Taylor[122]).

3.2.1.4 Mixed boundary condition A mixed boundary is a combination of the two last boundary conditions (Dirichlet and Neumann).

$$a\phi(x) + b\frac{\partial\phi(x)}{\partial n} = g(x) \tag{3.6}$$

It is called a Robin or Cauchy boundary condition or boundary condition of the third kind. This type of boundary condition (Comini et al.[27]) can define convective boundaries in heat conduction problems. There, the heat flux as function of the temperature is prescribed at Γ and the temperature of the surrounding medium is known.

3.2.1.5 Binary or periodic boundary conditions Until now only symmetries in the geometry were considered to lead to the application of the Dirichlet or Neumann boundary conditions. Especially in cylinder symmetric devices, such as rotating electrical machines, not only symmetries in the geometry but also in the magnetic field distribution are present. Under load conditions the air gap field of an electrical machine repeats periodically every double pole pitch. At no-load operation, it repeats itself every pole pitch. This field periodicity can be used to define another type of boundary condition to reduce the size of the numerical model. The local potentials in such boundaries depend on the solution of the field problem and thus inherently occur always in pairs. One boundary is computed and the opposite one is linearly linked to this value.
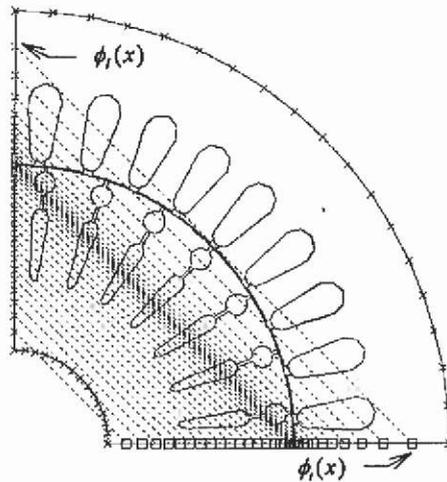
Fig. 3.7. Periodic boundary condition applied to a 4-pole induction motor model.

In Fig. 3.7 the dotted lines indicate the pairs of the boundaries. Obviously, the numerical discretisation of the model at those boundaries must be identical. This type of periodic boundary condition has the form:

$$k\phi_t(x) + \phi_t(x) = m. \tag{3.7}$$

If m=0 and k equals 1 or -1 this boundary is called binary boundary condition.

3.2.1.6 Far-field boundary condition The differential formulation of the field equations in the finite element method has its disadvantages for computing open and unbounded physical fields. The whole field domain, theoretically until infinity from the field sources, must be discretised to be able to compute the far field. For example, if the electromagnetic field in the vicinity of a high voltage transmission line is analysed, the air and ground have to be modelled. To model the infinity, a Kelvin-transformation can be used to map the infinite space to a finite space, forcing their solution to be identical (Fig. 3.8). Using this technique reduces the problem size and computational expenses significantly. Fig. 3.8 shows the circular domain of interest of an electrostatic problem with a high voltage tower in its centre. The small circle above is the FEM model that approximates the infinite space. There, one of the centre nodes is set to zero potential whereas the nodes at its circumference have the same potential as the diameter nodes of the large circle. The lines connecting both circles illustrate this link.
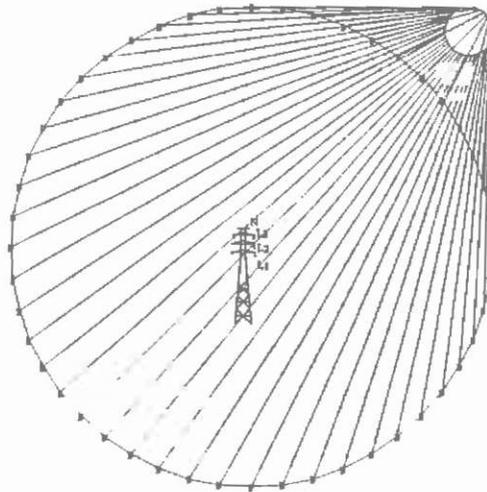
Fig. 3.8. Open or far-field boundary condition applied to the simulation of the electromagnetic field of a high voltage transmission line.

3.3 Field equations in partial differential form

Every electromagnetic phenomenon can be attributed to the seven basic equations, the four Maxwell equations of the electro-dynamic and those equations of the materials. The latter can be
- isotropic or an-isotropic
- linear or non-linear
- homogenous or non-homogenous.

The Maxwell equations are linked by interface conditions. Together with the material equations they form the complete set of equations describing the fields completely.

In this section the Maxwell equations, necessary for the calculation of electromagnetic fields, are discussed in their differential form. The seven equations describe the behaviour of the electromagnetic field in every point of a field domain. All electric and magnetic field vectors $\mathbf{E}$, $\mathbf{D}$, $\mathbf{B}$, $\mathbf{H}$, and $\mathbf{J}$ and the space charge density ρ are in general functions of time and space. The conducting current density can be distinguished by a material/field dependent part $\mathbf{J}_c$ and by an impressed and given value $\mathbf{J}_0$. It is assumed that the physical properties of the material permitivity ε, permeability μ and conductivity σ are independent of the time. Furthermore it is assumed that those quantities are piecewise homogenous.

Three groups of equations can be distinguished:

group 1:

(i)		$\nabla \times \mathbf{E} = -\dfrac{\partial \mathbf{B}}{\partial t}$	(3.8)
(ii)		$\nabla \times \mathbf{H} = \mathbf{J} + \dfrac{\partial \mathbf{D}}{\partial t}$	(3.9)

group 2:

(iii)	$\nabla \cdot \mathbf{B} = 0$	(3.10)
(iv)	$\nabla \cdot \mathbf{D} = \rho$	(3.11)

group 3 for isotropic material:

(v)	$\mathbf{D} = \varepsilon \mathbf{E}$	(3.12)
(vi)	$\mathbf{B} = \mu \mathbf{H}$	(3.13)
(vii)	$\mathbf{J} = \mathbf{J}_0 + \mathbf{J}_c = \mathbf{J}_0 + \sigma \mathbf{E}$	(3.14)

In the literature published in different languages different operators are used. Here for the identity of the operators (∇: *Nabla* or *del*)

$\nabla \cdot \mathbf{X} \triangleq div\, \mathbf{X}$

$\nabla \times \mathbf{X} \triangleq curl\, \mathbf{X} \triangleq rot\, \mathbf{X}$

$\nabla \mathbf{X} \triangleq grad\, \mathbf{X}$

is chosen. If ε, μ and σ are constant in a domain their position in the equations can be exchanged for the geometrical vector operators $\nabla \cdot$ and ∇. If they are constant in time, their position can be exchanged for the time derivatives $\partial / \partial t$.

The three groups of equations are called the main equations, the laws of conservation and the material equations. The first equation (i), eq.(3.8) is known as the law of Faraday-Lenz with $\mathbf{E}$ the electric field strength and $\mathbf{B}$ the flux density. Eq. (ii), (3.9) is known as Ampere's law with magnetic field strength $\mathbf{H}$, conducting current density $\mathbf{J}$ and $\mathbf{D}$ the electric flux density. The term $\partial \mathbf{D} / \partial t$, the displacement current density, is neglected from now on, as already argued. Equations (iii), (3.10) and (iv), (3.11) describe the constitutive properties of the magnetic flux density and the displacement current with the space charge density ρ. All six field quantities $\mathbf{E}$, $\mathbf{D}$, $\mathbf{H}$, $\mathbf{B}$, $\mathbf{J}$ and ρ are dependent on each other.

3.3.1 Motion

Electric and magnetic fields form a unit with phenomena depending on the point of view of the observer. An observer at rest looking at a moving charge perceives an electric field caused by the charge and an additional magnetic field. The observer moving with the same speed as the charge does notices only the electric field. The field quantities can be

represented in different co-ordinate systems by different electric and magnetic field quantities. The Lorentz-transformation can be used to link both resting and moving systems. The field strength in a resting system x,y,z and of a uniformly in x-direction moving system x',y',z' can be given by:

$$E'_\parallel = E$$

$$B'_\parallel = B$$

$$E'_\perp = \frac{(\mathbf{E} + \mathbf{v} \times \mathbf{B})_\perp}{\sqrt{1 - v^2/c^2}} \tag{3.15}$$

$$B'_\perp = \frac{(\mathbf{B} - (\mathbf{v}/c^2) \times \mathbf{E})_\perp}{\sqrt{1 - v^2/c^2}}$$

A distinction is made between the parallel and perpendicular direction of motion. The vectors for the field quantities are calculated by:

$$\mathbf{E}' = \begin{bmatrix} E'_x \\ E'_y \\ E'_z \end{bmatrix} = \begin{bmatrix} E_x \\[2ex] \dfrac{E_y - vB_z}{\sqrt{1 - v^2/c^2}} \\[3ex] \dfrac{E_z + vB_y}{\sqrt{1 - v^2/c^2}} \end{bmatrix} \tag{3.16}$$

$$\mathbf{B}' = \begin{bmatrix} B'_x \\ B'_y \\ B'_z \end{bmatrix} = \begin{bmatrix} B_x \\[2ex] \dfrac{B_y + E_z v/c^2}{\sqrt{1 - v^2/c^2}} \\[3ex] \dfrac{B_z - E_y v/c^2}{\sqrt{1 - v^2/c^2}} \end{bmatrix} \cdot \tag{3.17}$$

Using the Lorentz transformation for the charge density ρ' and the vector of the conducting current density **J**' for example for a particle beam moving in the x-direction with

$$\rho' = \frac{\rho - J_x v/c^2}{\sqrt{1 - v^2/c^2}} \quad , \tag{3.18}$$

it can be written

$$\mathbf{J}' = \begin{bmatrix} J'_x \\ J'_y \\ J'_z \end{bmatrix} = \begin{bmatrix} \dfrac{J_x - v\rho}{\sqrt{1 - v^2/c^2}} \\ J_y \\ J_z \end{bmatrix} \qquad . \tag{3.19}$$

In a similar way[1], the scalar potential φ' and vector potential $\mathbf{A}'$ of a moving charge in the x-direction is transformed by:

$$\varphi' = \frac{\varphi - vA_x}{\sqrt{1 - v^2/c^2}} \qquad , \tag{3.20}$$

$$\mathbf{A}' = \begin{bmatrix} A'_x \\ A'_y \\ A'_z \end{bmatrix} = \begin{bmatrix} \dfrac{A_x - v/c^2 \cdot \rho}{\sqrt{1 - v^2/c^2}} \\ A_y \\ A_z \end{bmatrix} \qquad . \tag{3.21}$$

In practically all electrical engineering problems of technical importance, observed phenomena concerning motion are slow in an electromagnetic sense, when the speed $\mathbf{v}$ is compared to the speed of light c. A possible exception is a high-energy particle accelerator in pulsed-power-technology. Observing an uniformly moving system x',y',z' from a resting co-ordinate system x,y,z and assuming $v \ll c$, the transformation is simplified (Schwab [101]) to:

$$\mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B}$$
$$\mathbf{B}' = \mathbf{B} - \frac{\mathbf{v}}{c^2} \times \mathbf{E} \qquad . \tag{3.22}$$

Applying the quantities within the appropriate co-ordinate system, the Maxwell equations remain valid. The field equations from groups 1 and 2 are Lorenz invariant. The same can be stated for the forces on charges caused by electric and magnetic fields. Forces depend on the frame of reference and can be of electric or magnetic origin. The Lorentz force is:

---

[1] The various potentials are introduced in a later section.

$$\mathbf{F} = \mathbf{F}_{el} + \mathbf{F}_{mag} = Q\mathbf{E} + Q(\mathbf{v} \times \mathbf{B}) \quad . \tag{3.23}$$

In contrast to the Lorentz invariance of the field equations, the material relation (vii), (3.14) changes for the moving system and the resting observer to:

$$\mathbf{J}' = \sigma(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{3.24}$$

and respectively for the resting system with the moving observer using $\mathbf{J}'=\mathbf{J}-\mathbf{v}\rho$, to:

$$\mathbf{J} - \mathbf{v}\rho = \sigma(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \quad . \tag{3.25}$$

3.3.2 Interface conditions
Technical devices are constructed using piecewise homogenous materials. The boundary of such materials can be identified as a surface inhomogeneity. To consider this boundary, the associated interface conditions for the electric and magnetic field are discussed in the following section.

To derive those interface conditions, the integral form of the Maxwell equations is used.

Table 3.1. Maxwell equations for quasi-stationary fields.

	differential form	integral form	
(i)	$\nabla \times \mathbf{E} = -\dfrac{\partial \mathbf{B}}{\partial t}$	$\oint_c \mathbf{E} \cdot d\mathbf{r} = -\dfrac{d\phi}{dt}$	(3.26)
(ii)	$\nabla \times \mathbf{H} = \mathbf{J}$	$\oint_c \mathbf{H} \cdot d\mathbf{r} = I$	(3.27)
(iii)	$\nabla \cdot \mathbf{B} = 0$	$\oint_S \mathbf{B} \cdot d\mathbf{S} = 0$	(3.28)
(iv)	$\nabla \cdot \mathbf{D} = \rho$	$\oint_S \mathbf{D} \cdot d\mathbf{S} = Q$	(3.29)

Here, ϕ is the magnetic flux, I the conducted current, Q the charge, C indicates the contour integral and S the surface integral.

3.3.2.1 Normal component Maxwell equation (iii), (3.28) in integral form is used to derive the interface conditions at the boundary of different materials for the normal component (Fig. 3.9). Using $d\mathbf{S}_1 = -d\mathbf{S}_2 = \mathbf{n}dS$ with n the unit vector in normal direction yields:

$$\lim_{w \to 0} \oint_S \mathbf{B} \cdot d\mathbf{S} = \int_S \mathbf{B}_1 \cdot d\mathbf{S}_1 + \int_S \mathbf{B}_2 \cdot d\mathbf{S}_2 = \int_S (\mathbf{B}_1 - \mathbf{B}_2)\mathbf{n} \cdot dS = 0 \quad . \tag{3.30}$$
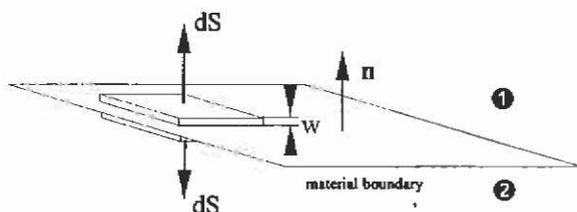
Fig. 3.9. Interface between material ❶ and ❷ with different properties.

The surface S can be of arbitrary shape and the integral only vanishes if the integrand is zero. This yields the interface conditions for the electromagnetic field. The components of the magnetic flux density **B** are continuous at material boundaries even if they have different ferromagnetic properties.

$$(\mathbf{B}_1 - \mathbf{B}_2)\cdot \mathbf{n} = 0$$

or (3.31)

$$B_{n1} = B_{n2} \quad .$$

With respect to the finite element method, this means that the normal component of the flux density must be constant at the boundary between finite elements (Fig. 3.10).
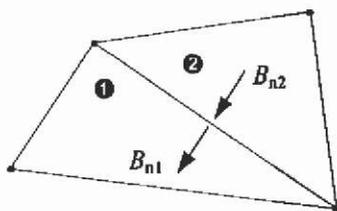
Fig. 3.10. Normal component of the flux density at the interface of two triangular finite elements.

Similar to the normal component of the magnetic flux density, the normal component of the displacement current density can be derived. Here, Maxwell equation (iv) in integral form eq.(3.29) is evaluated in the same way yielding:

$$\lim_{w\to 0}\oint_S \mathbf{D}\cdot dS = \int_S \mathbf{D}_1\cdot dS_1 + \int_S \mathbf{D}_2\cdot dS_2 = \int_S (\mathbf{D}_1 - \mathbf{D}_2)\mathbf{n}\cdot dS$$
$$= \int_V \rho\cdot dV = \int_V dQ = \int_S \rho_s\cdot dS \quad . \tag{3.32}$$

Where ρ is the space charge density, V indicates the volume integral and $\rho_s = dQ/dS$ the surface charge density,

$$(\mathbf{D}_1 - \mathbf{D}_2) \cdot \mathbf{n} = \rho_s$$

or   (3.33)

$$D_{n1} = D_{n2} + \rho_S \quad .$$

With the presence of a surface charge density, the normal component of the displacement current density is discontinuous at the interface of a material boundary. Without surface charge density, the normal component of the displacement current density is continuous at boundaries.
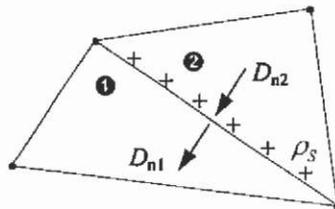
Fig. 3.11. Normal component of the displacement current density at the interface of two triangular finite elements.

To discuss the interface conditions for the conducting current density J, the Maxwell equation (ii), eq.(3.27) is considered. With (vi), (3.13), the same conditions found for the magnetic field can be applied for the current density as well:

$$(\mathbf{J}_1 - \mathbf{J}_2) \cdot \mathbf{n} = 0$$

or   (3.34)

$$J_{n1} = J_{n2} \quad .$$
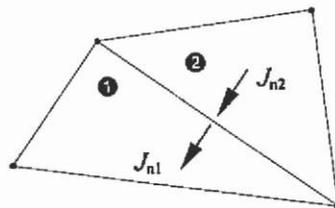
Fig. 3.12. Normal component of the current density.

3.3.2.2 Tangential component The interface conditions valid for the tangential component of the electrical field E can be derived from the integral form of the first Maxwell equation (i), (3.26).
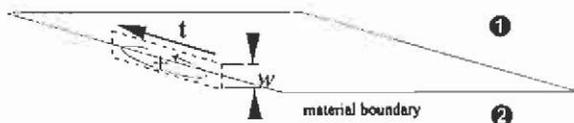
Fig. 3.13. Interface and path of integration between material ❶ and ❷ with different material properties.

Applying the path of integration on the contour drawn in Fig. 3.13 and counting positive as indicated, yields:

$$\lim_{w\to 0}\oint_c \mathbf{E}\cdot d\mathbf{r} = \int_c \mathbf{E}_1 \cdot d\mathbf{r}_1 + \int_c \mathbf{E}_2 \cdot d\mathbf{r}_2 = \int_c (\mathbf{E}_1 - \mathbf{E}_2)\mathbf{t}\cdot dr$$

$$= -\lim_{w\to 0}\int_s \frac{\partial \mathbf{B}}{\partial t}\cdot d\mathbf{S} = 0 \quad . \tag{3.35}$$

It is assumed that $d\mathbf{r}_1 = -d\mathbf{r}_2 = \mathbf{t}dr$, with $\mathbf{t}$ the unit vector in tangential direction. With a finite $\mathbf{B}$ and $w\to 0$ the surface integral over $\mathbf{B}$ vanishes and this yields:

$$(\mathbf{E}_1 - \mathbf{E}_2)\cdot \mathbf{t} =$$
$$\mathbf{n}\times(\mathbf{E}_1 - \mathbf{E}_2) = 0 \tag{3.36}$$

or

$$E_{t1} = E_{t2} \quad .$$
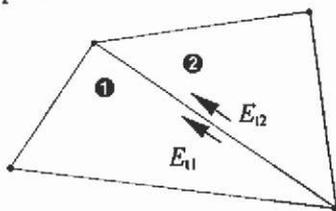
Fig. 3.14. Continuous tangential component of the electric field strength.

The tangential component of the electric field strength is continuous at interfaces of a boundary with different material properties.

Analogous to the electric field strength and employing the integral form of Maxwell equation (ii) eq.(3.27) yields:

$$\lim_{w\to 0}\oint_c \mathbf{H}\cdot d\mathbf{r} = \int_c \mathbf{H}_1 \cdot d\mathbf{r}_1 + \int_c \mathbf{H}_2 \cdot d\mathbf{r}_2 = \int_c (\mathbf{H}_1 - \mathbf{H}_2)\mathbf{t}\cdot dr$$

$$= \lim_{w\to 0}\int dI = 0 + \int J_\sigma dr \quad . \tag{3.37}$$

Here J_σ is a possible surface current density.

$$\int_c \left[(\mathbf{H}_1 - \mathbf{H}_2) \cdot \mathbf{t} - J_\sigma \right] dr = 0 \quad . \tag{3.38}$$

Herewith, the interface conditions for the tangential component of the magnetic field strength **H** can be written by:

$$(\mathbf{H}_1 - \mathbf{H}_2) \cdot \mathbf{t} =$$
$$\mathbf{n} \times (\mathbf{H}_1 - \mathbf{H}_2) = J_\sigma \tag{3.39}$$
$$\text{or}$$
$$H_{t1} = H_{t2} + J_\sigma \quad .$$
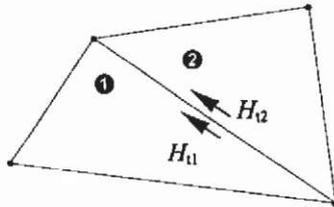
Fig. 3.15. Continuous tangential component of the magnetic field strength (J_σ=0).

With vanishing surface current density the tangential component of the magnetic field strength is continuous at interfaces.

4 Potentials and formulations

The Maxwell equations represent the physical properties of the fields. To solve them, mainly the differential form of the equations and mathematical functions, the potentials, satisfying the Maxwell equations, are used. The proper choice of a potential depends on the type of field problem. In this section, the various scalar and vector potentials are introduced.

The electric vector potential for the displacement current density will not be introduced here, because it is only important for the calculation of fields in charge-free and current-free regions such as hollow wave-guides or in surrounding fields of antennas.

Various potential formulations are possible for the different field types. Their appropriate definition ensures the accurate transition of the field problem between continuous and discrete space.

Using these artificial field quantities reduces the number of differential equations. Considering a problem described by n differential equations, a potential is chosen in such a way that one of the differential equations is fulfilled. This potential is substituted in all other differential equations, the resulting system of differential equations reduces to $n\text{-}1$ equations. It is distinguished between magnetic and electric vector respectively scalar potentials:

Table 4.1. Definition of the potentials.

	potentials	
	scalar	vector
electric	$\mathbf{E} = -\nabla V - \dfrac{\partial \mathbf{A}}{\partial t}$	$\mathbf{J} = \nabla \times \mathbf{T}$
magnetic	$\mathbf{H} = \mathbf{T} - \nabla \phi$ $\mathbf{H} = -\nabla \phi$	$\mathbf{B} = \nabla \times \mathbf{A}$

4.1   Magnetic vector potential

By using the vector identity

$$\nabla \cdot (\nabla \times \mathbf{A}) = 0 \tag{4.1}$$

and applying eq.(iii), (3.10), introduces the magnetic vector potential $\mathbf{A}$. The magnetic flux density is derived as the *curl* of another vector field:

$$\mathbf{B} = \nabla \times \mathbf{A} \quad . \tag{4.2}$$

The magnetic vector potential is suitable in regions with and without conducting currents. The vector field $\mathbf{A}$ is assigned right handed to the direction of the magnetic field $\mathbf{B}$ (Fig. 4.1).
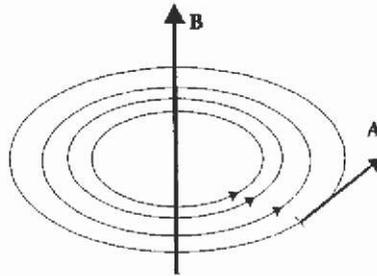
Fig. 4.1. Geometrical assignment of the vector potential A with the magnetic field vector B.

A static magnetic problem is described by

$$\nabla \times \mathbf{H} = \mathbf{J}_0$$
$$\nabla \cdot \mathbf{B} = 0 \quad . \tag{4.3}$$

By using the magnetic vector potential $\mathbf{A}$, the system of differential equations is reduced to

$$\nabla \times (\frac{1}{\mu} \nabla \times \mathbf{A}) = \mathbf{J}_0 \quad . \tag{4.4}$$

Applying the vector calculus $\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$ to eq.(4.4) yields:

$$\nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} = \mu \mathbf{J}_0 \tag{4.5}$$

with $\nabla \cdot \mathbf{A} = 0$ and by assuming a constant permeability μ, leads to the A-formulation of a magneto-static field, a Poisson equation:

$$\nabla^2 \mathbf{A} = -\mu \mathbf{J}_0 \quad . \tag{4.6}$$

To consider quasi-stationary fields, for example necessary for eddy current calculations, the magneto-dynamic formulations have to be employed. In addition to Ampere's law, the Faraday law (i), (3.26) has to

be considered to evaluate the contribution to the field by the eddy currents:

$$\nabla \times \mathbf{E} = -\frac{\partial}{\partial t}\mathbf{B} = -\frac{\partial}{\partial t}\nabla \times \mathbf{A} \quad . \tag{4.7}$$

Now employing Ohm's law to calculate the eddy currents $\mathbf{J}_e$ yields:

$$\mathbf{J}_e = -\sigma \frac{\partial}{\partial t}\mathbf{A} \quad . \tag{4.8}$$

Ampere's law can now be rewritten, yielding the A-formulation for the quasi-stationary magnetic field in the time domain:

$$\nabla \times (\frac{1}{\mu}\nabla \times \mathbf{A}) + \sigma \frac{\partial}{\partial t}\mathbf{A} = \mathbf{J}_0 \quad . \tag{4.9}$$

Substituting again $\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A}$ and assuming $\nabla \cdot \mathbf{A} = 0$, results in a similar A-formulation in the time domain for the transient magnetic field:

$$\nabla^2\mathbf{A} - \mu\sigma\frac{\partial}{\partial t}\mathbf{A} = -\mu\mathbf{J}_0 \quad . \tag{4.10}$$

Assuming sinusoidal excitation currents with an angular frequency ω and thus substituting

$$\frac{\partial}{\partial t}\mathbf{A} = j\omega\mathbf{A} \tag{4.11}$$

yields the A-formulation in the frequency domain to solve eddy current problems.

$$\nabla^2\mathbf{A} - j\omega \cdot \mu\sigma\mathbf{A} = -\mu\mathbf{J}_0 \tag{4.12}$$

This equation is the A-formulation to describe time-harmonic problems. The time dependent components of the vector potential $A(t) = \hat{A} \cdot \cos(\omega t + \varphi)$ are expressed by:

$$A = \hat{A} \cdot e^{j(\omega t + \varphi)} \quad . \tag{4.13}$$

The current is expressed in analogy in its complex representation.

4.2 Electric vector potential for conducting current

For the calculation of eddy current problems the electric vector potential is often employed. The current density fulfils the zero divergence condition. Therefore, and analogous to **A**, an electric vector potential **T** can be defined:

$$\mathbf{J} = \nabla \times \mathbf{T} \quad . \tag{4.14}$$

4.3   Electro-static scalar potential

By using the vector identity $\nabla \times (\nabla V) = 0$, eq.(i), (3.26) can be rewritten:

$$\mathbf{E} = -\nabla V \quad . \tag{4.15}$$

The negative sign is arbitrary and applied to have a close similarity to the definition of the magnetic scalar potential. Employing $\mathbf{E}$ to eq.(iv), (3.29) yields:

$$\nabla \cdot \nabla V = \nabla^2 V = -\frac{\rho}{\varepsilon} \quad . \tag{4.16}$$

This is an equation of the Poisson type.

4.4   Magnetic scalar potential

By analogy to the electric field, the magnetic field strength is calculated as the gradient of a scalar potential. It must be distinguished between current-carrying and current-free regions.

4.4.1   Current-free regions

The magneto-static problem without conducting currents can be formulated in terms of the magnetic scalar potential ϕ.

With the vector identity $\nabla \times (\nabla \phi) = 0$ and Ampere's law, a magnetic scalar potential ϕ can easily be defined by evaluating:

$$\mathbf{H} = -\nabla \phi \quad . \tag{4.17}$$

This potential formulation is not suitable for problems inside regions with conducting currents. A typical application for this type of potential is the calculation of a magnetic shielding.

With the zero divergence condition of the magnetic flux density the ϕ-formulation of the scalar magnetic potential is introduced:

$$\begin{aligned} \nabla \cdot (\mu \nabla \phi) &= 0 \\ \mu \nabla^2 \phi &= 0 \quad . \end{aligned} \tag{4.18}$$

This formulation is a Laplace equation. ϕ is a scalar and $\mathbf{A}$ in the vector potential formulation of the magnetic field is a vector quantity. By using the same numerical discretisation, the scalar potential problem consists of a third of unknown when compared to the formulation using the vector potential. This example makes clear that an appropriate choice of the potential formulation has significant influence on the size of the problem.

This definition of the scalar magnetic potential, $\mathbf{H} = -\nabla \phi$, causes problems in multiple connected domains.
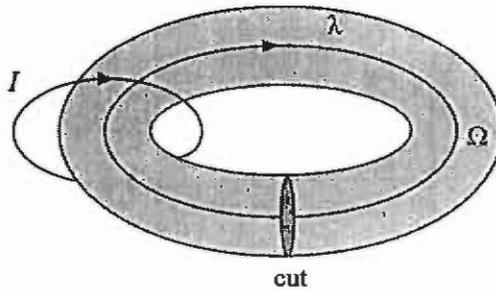
Fig. 4.2. Multiple connected domain.

The domain is free of conducting current. There is a current I carrying conductor leading through the opening in the domain (Fig. 4.2). Applying Ampere's law on the contour λ inside the domain yields:

$$\int_\lambda \mathbf{H} \cdot dl = \int_\lambda -\nabla \phi \cdot dl = \phi_A - \phi_B \quad . \tag{4.19}$$

The potential difference $\phi_A - \phi_B$ can only be non-zero if ϕ is discontinuous inside the domain. Therefore, a discontinuity, a cut, is defined in the way that by considering the cut as an outside boundary, the domain is not further multiple connected.

$$\nabla \cdot \nabla \phi = \nabla^2 \phi = I \quad . \tag{4.20}$$

Therefore, the boundary condition applied to the cut is

$$\phi_A - \phi_B = I \quad . \tag{4.21}$$

This is a periodic boundary condition.

4.4.2 Current-carrying regions
In the case of current-carrying regions it is not possible to define a magnetic scalar potential. However, by again using an arbitrary vector field $\mathbf{T}$, it is possible to define a similar potential (Silvester & Ferrari[103]).

The electric vector potential $\mathbf{T}$ and Ampere's law eq.(ii), (3.27) yield:

$$\nabla \times (\mathbf{H} - \mathbf{T}) = \mathbf{0} \quad . \tag{4.22}$$

With the vector identity

$$\nabla \times (\nabla \phi) = \mathbf{0} \quad , \tag{4.23}$$

the gradient of the scalar magnetic potential ϕ is now defined by:

$$\mathbf{H} - \mathbf{T} = -\nabla \phi \quad . \tag{4.24}$$

The zero divergence condition of the magnetic field $\nabla \cdot \mathbf{B} = 0$ and the material equation combining $\mathbf{B}$ and $\mathbf{H}$ yields an equation of Poisson type:

$$\nabla \cdot (\mu \nabla \phi) = \nabla \cdot (\mu \mathbf{T}) \quad . \tag{4.25}$$

A disadvantage in solving this problem is that the solution of the magnetic field problem has to be obtained in three steps:

1. Determine the auxiliary potential function $\mathbf{T}$.
2. Solve the Poisson equation $\nabla \cdot (\mu \nabla \phi) = \nabla \cdot (\mu \mathbf{T})$ to find the magnetic scalar potential.
3. Evaluate $\nabla \phi$ and $\mathbf{T}$ to obtain the required overall solution of $\mathbf{H}$.

$$\nabla \cdot (\mu \nabla \phi) = \mu \nabla^2 \phi = \nabla \cdot (\mu \mathbf{T}) \tag{4.26}$$

is called a $T\phi$-formulation. The magnetic scalar potential has the dimension [A]. This potential formulation is in common use in magneto-static and diffusion problems. To determine $\mathbf{T}$, Biot-Savart's law can be evaluated (Hafner[47]).

By using the finite element method, $\mathbf{T}$ can be determined in the following way:

1. Create a tree of mesh edges.
2. $\mathbf{T} = 0$ for all tree-edges.
3. Apply Ampere's law to each element and determine $\mathbf{T}$ for all co-tree edges.

A tree in the topology of the finite element mesh is defined as a set of edges reaching all nodes of the mesh but forming no loops in this mesh. The associated co-tree is the set of the remaining edges.

$\mathbf{T}$ is constructed as a field built of edge elements $\mathbf{w}_{e,i}$:

$$\mathbf{T} = \sum_{i=1}^{n} h_{s,i} . \mathbf{w}_{e,i} \tag{4.27}$$

The coefficients $h_{s,i}$ of the elements associated with the tree edges are 0. The coefficients of the elements associated with the co-tree edges are calculated from Ampere's law (Fig. 4.3):

$$\int_{\Gamma} \mathbf{H} \, d\mathbf{r} = h_{s,1} . \mathbf{w}_{e,1} + h_{s,2} . \mathbf{w}_{e,2} + h_{s,3} . \mathbf{w}_{e,3} = I \tag{4.28}$$
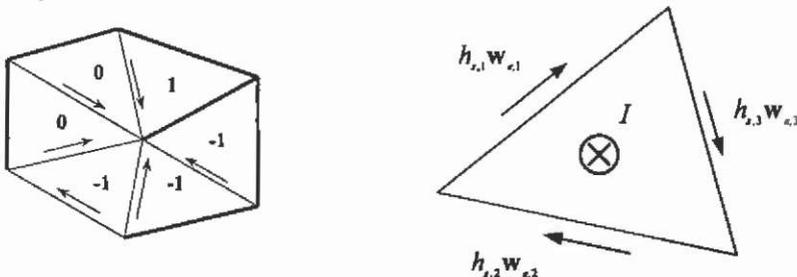
Fig. 4.3. Tree definition and finite element with imposed current.

With Table 4.2 a comparison of the properties of the magnetic vector and scalar potential is possible.

Table 4.2. Comparison between A- and ϕ-formulation.

	A-formulation	ϕ-formulation
potential formulation	vector potential $\nabla^2 \mathbf{A} = -\mu \mathbf{J}$	scalar potential $\mu \nabla^2 \phi = \nabla \cdot (\mu \mathbf{T})$
implicitly fulfilled equation	$\nabla \cdot \mathbf{B} = 0$	$\nabla \times \mathbf{H} = \mathbf{J}$
explicitly fulfilled equations	$\nabla \times \mathbf{H} = \mathbf{J}$ $\mathbf{B} = \mu \mathbf{H}$	$\nabla \cdot \mathbf{B} = 0$ $\mathbf{B} = \mu \mathbf{H}$
source field	$\mathbf{J}$	$\mathbf{T}$ has to be determined
additional condition	gauge	cut
element type	edge	node

4.5   A ϕ-formulation

It is sometimes interesting to use a hybrid Aϕ-formulation. The ϕ-formulation is used in regions without current whereas an A-formulation is used in regions with an applied current density. At the interface surfaces between regions with a different potential formulation, the conditions for $\mathbf{H}$ and $\mathbf{B}$ are applied:

$$(\mathbf{H}_1 - \mathbf{H}_2) \cdot \mathbf{t} = (\nabla \phi - \frac{1}{\mu} \nabla \times \mathbf{A}) \cdot \mathbf{t} = 0 \quad , \tag{4.29}$$

$$(\mathbf{B}_1 - \mathbf{B}_2) \cdot \mathbf{n} = (\mu \nabla \phi - \nabla \times \mathbf{A}) \cdot \mathbf{n} = 0 \quad .$$

The advantage of this approach compared to the ϕ-formulation, is that no source field $\mathbf{T}$ has to be constructed. The advantage of this approach compared to the A-formulation, is that regions without current are described by a scalar potential instead of a vector potential. A disadvantage of the hybrid formulation is the extra interface condition.

4.6   AV-formulation

Using Ampere's law and the magnetic vector potential and applying the appropriate material equation yields:

$$\nabla \times \mathbf{H} = \mathbf{J} = \mathbf{J}_0 + \sigma \mathbf{E} \quad ,$$

$$\nabla \times (\nabla \times \mathbf{A}) + \mu \sigma (\nabla V + \frac{\partial}{\partial t} \mathbf{A}) = \mu \mathbf{J}_0 \quad , \tag{4.30}$$

$$\nabla^2 \mathbf{A} - \mu \sigma \frac{\partial}{\partial t} \mathbf{A} = -\mu \mathbf{J}_0 - \mu \sigma \nabla V \quad . \tag{4.31}$$

This equation is called the AV-formulation of the magnetic field. The presence of the derivative in time indicates the magneto-dynamic field problem. By substituting $\frac{\partial}{\partial t}\mathbf{A} = j\omega\mathbf{A}$ this formulation is transferred from the time domain into the frequency domain:

$$\nabla^2\mathbf{A} - j\omega \cdot \mu\sigma\mathbf{A} = -\mu\mathbf{J}_o - \mu\sigma\nabla V \quad . \tag{4.32}$$
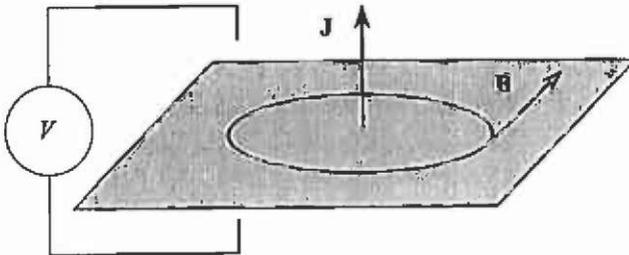
Fig. 4.4. Field problem in two-dimensional AV-formulation.

4.7   In-plane formulation

The inverse problem, to calculate a current density distribution for an imposed magnetic field can be obtained by using an in-plane formulation. Applying the Maxwell equation (3.8) and using the magnetic scalar and electric vector potential yields a particular $T\phi$ -formulation.

$$\nabla \times \mathbf{E} = -\frac{\partial}{\partial t}\mathbf{B} = -\frac{\partial}{\partial t}\mu(\mathbf{T} - \nabla\phi) \quad .$$

$$\mathbf{J} = \sigma\mathbf{E} = \nabla \times \mathbf{T} \quad \text{yields:} \tag{4.33}$$

$$\nabla \times \mathbf{E} = \nabla \times (\frac{1}{\sigma}\nabla \times \mathbf{T}) \quad .$$

Finally the in-plane $T\phi$ -formulation, valid in the time domain, is described by:

$$\nabla \times (\nabla \times \mathbf{T}) + \mu\sigma\frac{\partial}{\partial t}\mathbf{T} = \mu\sigma\frac{\partial}{\partial t}\nabla\phi \quad . \tag{4.34}$$

In the frequency domain, by assuming a sinusoidal magnetic field and substituting $\frac{\partial}{\partial t}\mathbf{T} = j\omega\mathbf{T}$, it can be written:

$$\nabla \times (\nabla \times \mathbf{T}) + j\omega \cdot \mu\sigma\mathbf{T} = j\omega \cdot \mu\sigma\nabla\phi \quad . \tag{4.35}$$

This formulation is very similar to the A-formulation of the magnetic field and can be solved with the same methods. This formulation can not be applied to non-conducting regions. There,

$$\nabla \cdot (\mu T) = \nabla \cdot (\mu \nabla \phi) \tag{4.36}$$

must be evaluated.
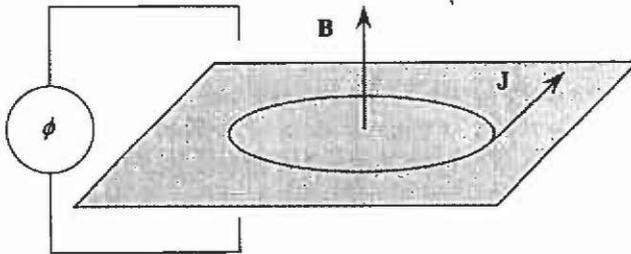
Fig. 4.5. In-plane $T\phi$-formulation.

4.8 AV-formulation with $v \times B$ motion term

For the moving system and the resting observer, as introduced earlier, the motion term is considered by applying:

$$J = \sigma E + \sigma(v \times B) \qquad . \tag{4.37}$$

Together with Ampere's law (3.9), this leads to the AV-formulation considering the motion term $v \times B$.

$$\nabla^2 A + \mu\sigma(v \times (\nabla \times A)) - \mu\sigma \frac{\partial}{\partial t} A = -\mu J_0 + \mu\sigma\nabla V \tag{4.38}$$

Considering the motion term causes the later coefficient matrix to be non-symmetric. This is important for the stability of the method used. To obtain a stable solution procedure, the discretised part of the convection term $\mu\sigma(v \times B)$ has to be smaller when compared to the term $\nabla^2 A$ (Hackbusch [46]).

4.9 Gauge conditions

With the known material equations, the magnetic flux density B can be expressed by the magnetic field strength H or vice versa; D and J respectively can be determined by the electric field strength E:

$$\nabla \times H = J = J_0 + \sigma E$$

$$\nabla^2 A - \nabla(\nabla \cdot A) - \mu\sigma \frac{\partial}{\partial t} A = -\mu J_0 + \mu\sigma\nabla V \quad , \tag{4.39}$$

and

$$\nabla \cdot \mathbf{D} = \rho = \nabla \cdot (\varepsilon \mathbf{E}) = -\nabla \cdot \varepsilon (\nabla V + \frac{\partial}{\partial t} \mathbf{A})$$

$$ \tag{4.40}$$

$$\nabla \cdot \nabla V + \frac{\partial}{\partial t} \nabla \cdot \mathbf{A} = -\frac{\rho}{\varepsilon} \quad .$$

The numerical difficulty with the equations is obvious; they are coupled. In both equations both potentials are found. Evaluating these equations to obtain the potentials shows that the solution obtained in this way for the vector potential $\mathbf{A}$ and the electric scalar potential V is not unambiguous. Therefore, a gauge has to be applied to define the potential solution uniquely. Different gauges are possible. Arbitrary additions to the coupled potentials $\mathbf{A}$ or V do not influence the field values:

$$\mathbf{B} = \nabla \times \mathbf{A} \quad , \tag{4.41}$$

$$\nabla V = -\mathbf{E} - \frac{\partial \mathbf{A}}{\partial t} \quad . \tag{4.42}$$

Those additional constants can be normalised by fixing an arbitrary point in the field region to zero. With this normalisation, the potential solution is not yet unique. Therefore, a divergence condition for $\mathbf{A}$ has to be given. The choice of this condition is arbitrary because it influences the values of $\mathbf{A}$ and V only, but not the value of the derived field quantities $\mathbf{B}$ and $\mathbf{H}$. The most simplest and frequently used gauge condition for static fields is the zero divergence condition, the Coulomb gauge:

$$\nabla \cdot \mathbf{A} = 0 \quad . \tag{4.43}$$

Applying the Coulomb gauge yields for the potential eqs.(4.39) and (4.40):

$$\nabla^2 \mathbf{A} - \mu\sigma \frac{\partial \mathbf{A}}{\partial t} = -\mu \mathbf{J}_0 + \mu\sigma \nabla V \quad , \tag{4.44}$$

and

$$\nabla \cdot \nabla V = \nabla^2 V = -\frac{\rho}{\varepsilon} \quad . \tag{4.45}$$

Employing the Coulomb gauge here simplifies these equations insignificantly. The potentials stay coupled. Here, it is interesting that the potential V satisfies and describes the type of a Poisson equation.

To de-couple the potential equations, the Lorentz gauge

$$\nabla \cdot \mathbf{A} = -\mu\sigma V \tag{4.46}$$

can be used. Applied to the equations (4.39) and (4.40) this yields:

$$\nabla^2 \mathbf{A} - \mu\sigma \frac{\partial}{\partial t} \mathbf{A} = -\mu \mathbf{J}_0 \quad ,$$

$$\nabla^2 V - \mu\sigma \frac{\partial}{\partial t} V = -\frac{\rho}{\varepsilon} \quad . \tag{4.47}$$

An advantage of using the Lorentz gauge is the fact that the decoupled potential equations satisfy differential equations of the same form. Assuming the case $\rho = 0$, the Lorentz gauge can be written by:

$$\nabla \cdot \nabla V + \frac{\partial}{\partial t} \nabla \cdot \mathbf{A} = 0 \quad . \tag{4.48}$$

For the source-free case, $\rho = \mathbf{J}_0 = 0$, the Coulomb gauge is interesting. The equation for the scalar potential becomes a Laplace equation and if the scalar potential is normalised to zero in infinite distance, the solution of the Laplace equation is trivial $\phi = 0$. Therefore, only one equation remains:

$$\nabla^2 \mathbf{A} - \mu\sigma \frac{\partial}{\partial t} \mathbf{A} = 0 \quad . \tag{4.49}$$

The interesting field vectors $\mathbf{E}$ and $\mathbf{B}$ can directly been determined by the vector potential:

$$\mathbf{B} = \nabla \times \mathbf{A} \quad ,$$

$$\mathbf{E} = -\frac{\partial}{\partial t} \mathbf{A} \quad . \tag{4.50}$$

The Lorentz gauge is in common use in solving wave equations. There, the displacement current density is considered. The inhomogeneous wave equations are given at this place for completion:

$$\nabla^2 \mathbf{A} - \mu\sigma \frac{\partial}{\partial t} \mathbf{A} - \mu\varepsilon \frac{\partial^2}{\partial t^2} \mathbf{A} = -\mu \mathbf{J}_0 \quad ,$$

$$\nabla^2 V - \mu\sigma \frac{\partial}{\partial t} V - \mu\varepsilon \frac{\partial^2}{\partial t^2} V = -\frac{\rho}{\varepsilon} \quad . \tag{4.51}$$

4.10   Subsequent treatment of the Maxwell equations

To obtain solutions for real-life field problems, a subsequent treatment of the Maxwell equations is necessary. The potentials are introduced to reduce the mathematical dimensions of the field problem. This approach results, for example for a line integral, in building a simple difference. To improve understanding of the field equations, a scheme is introduced representing the Maxwell equations. The various potentials are implemented in this scheme (Hafner [47]).

The four field equations in differential form can immediately be put into a schematic:

$$
\begin{array}{ccc}
\mathbf{E} & \xrightarrow{\;\nabla\times\;} & -\dfrac{\partial \mathbf{B}}{\partial t} \\[2ex]
\mathbf{J} & \xleftarrow{\;\nabla\times\;} & \mathbf{H} \\[2ex]
\rho \qquad \xleftarrow{\;\nabla\cdot\;} \qquad \mathbf{D} & & \\[2ex]
& \mathbf{B} & \xrightarrow{\;\nabla\cdot\;} \qquad 0 \quad .
\end{array}
$$
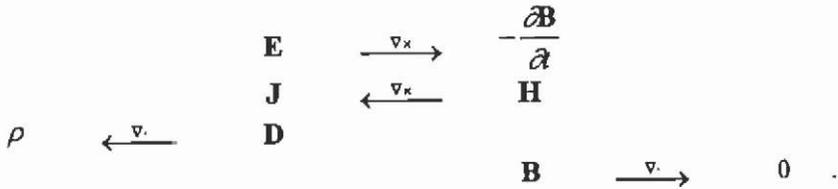
To keep magnetic and electric field quantities in the same column, the *div* and *curl* operators point to the right or left respectively. Consider the four following possible chain or arrow structures, with s a scalar and v a vector potential field.

1:	s	$\longleftarrow$	...			
2:	0	$\longleftarrow$	...	$\longleftarrow$	v	
3:	v	$\longrightarrow$	...	$\longrightarrow$	0	
4:			...	$\longrightarrow$	s	

The arrows represent the differential operators *div, curl* or *grad* respectively. A chain consists of at maximum two arrows. The second arrow always points to zero. Two arrows in one direction indicate a double derivative in the same direction in space.

Comparing this structured scheme with the structure of the Maxwell equations indicates the places where the scheme has to be completed by a potential. It is obvious that the fourth arrow chain can be completed by the vector potential $\mathbf{A}$ in the form $\mathbf{B} = \nabla \times \mathbf{A}$ because the flux density $\mathbf{B}$ satisfies the zero condition for a divergence free field eq.(iii), (3.28).

Applying this relation to the first Maxwell equation yields

$$\nabla \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}\right) = 0 \qquad . \tag{4.52}$$

Completing with the electric scalar potential of a gradient field V and in arrow notation it is:

$$
-V \quad \xrightarrow{\;\nabla\;} \quad \mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \quad \xrightarrow{\;\nabla\times\;} \quad 0 \quad .
$$
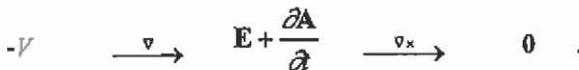
The complete scheme with the arbitrary potentials can now be written by:

$$V \xrightarrow{\;\nabla\;} -\mathbf{E} - \frac{\partial \mathbf{A}}{\partial t} \xrightarrow{\;\nabla\times\;} 0$$

$$0 \xleftarrow{\;\nabla\cdot\;} \mathbf{J} \xleftarrow{\;\nabla\times\;} \mathbf{H}$$

$$\rho \xleftarrow{\;\nabla\cdot\;} \mathbf{D}$$

$$\mathbf{A} \xrightarrow{\;\nabla\times\;} \mathbf{B} \xrightarrow{\;\nabla\cdot\;} 0 \quad .$$

To study the field equations further and to see the interdependencies between the source terms and the potentials, an improved graphical scheme can be used, the diagrams of Tonti (Bossavit[16]).

Using the same schematic as in the last section for the differential operators, and introducing an additional arrow for the time dependency of the quantities, the following arrow system is obtained (Table 4.3):

Table 4.3. Tonti's arrow system notation.

interdependency	operator	arrow
geometry	∇	vertical
	$\nabla\cdot$	
	$\nabla\times$	
material	μ	horizontal
	ε	
derivative with respect to time	$\dfrac{\partial}{\partial t}$	perpendicular to material and geometrical arrow
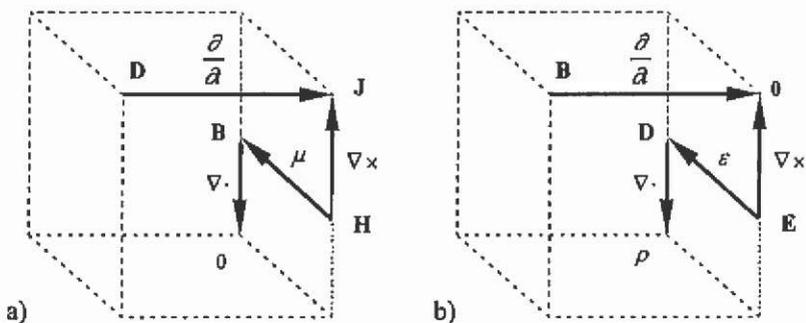
Fig. 4.6. Maxwell equations in Tonti's arrow notation for a) the magnetic field and b) the electric field.

The strong equivalence of the electric and magnetic fields is obvious using this notion (Fig. 4.6). The magnetic flux density **B** and the electric flux density **D** connect both field types. This results in the diagram of Tonti for the electro-magnetic field (Fig. 4.7) (Bossavit[16]).
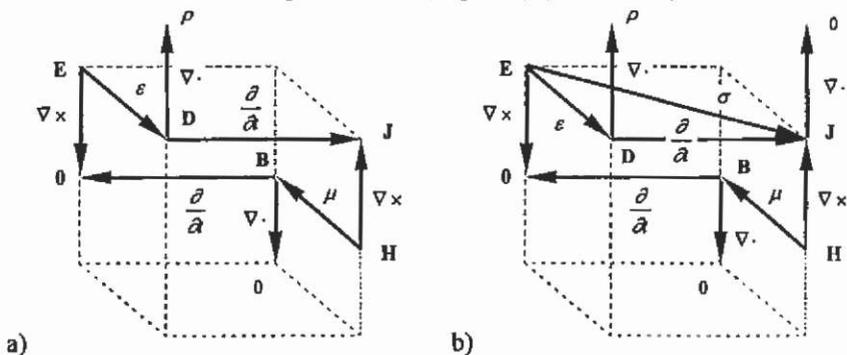
Fig. 4.7. a) Tonti's diagram for the electro magnetic field, b) by considering Ohm's law.

If conducting material is assumed, the dependency of electrical field strength and current density can be considered in the diagram of Tonti directly by Ohm's law (Fig. 4.7)

$$\mathbf{J} = \sigma\mathbf{E} \tag{4.53}$$

This equation, and respectively its arrow, represents the time independent current from eq.(ii), (3.27).

For ideal conductors we have the zero divergence condition, and therefore it can be written:

$$\nabla \cdot \mathbf{J} = 0 \tag{4.54}$$

To complete the diagrams, the defined potentials can be added to it as indicated (Fig. 4.8).
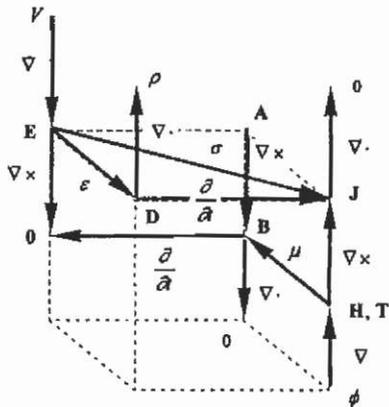
Fig. 4.8. Diagram of Tonti with potential definitions.

5 Field computation and numerical techniques

We consider that a variational principle or a boundary problem can describe a given physical-technical problem. Thus, this field problem is given by a differential equation. The problem is now to find a feasible solution of this differential equation.

Fig. 5.1 shows the various possibilities for solving general field problems. The methods applicable for use can be divided into two general classes, analytical and numerical methods. Methods which are based on simplified analytical models are called semi-numerical.
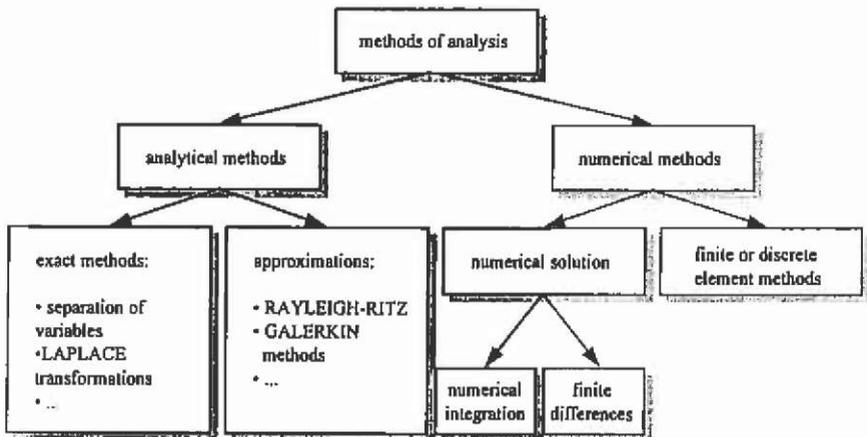
Fig. 5.1. General field analysis.

When compared to numerical techniques, analytical methods have the opportunity to deliver the exact solution of the differential equation. Those approaches, separation of the variables and Laplace transformations or other methods, can be applied to geometrically simple problem formulations only. Analytical approximations are suitable where the problem itself is very well known so that it is possible to apply appropriate simplifications. The entire above-mentioned reasons limit the

general application range for analytical methods. They are very suitable for specific problems. If an analytical approach is applicable, the solution is obtained in a rather short time; this is another great advantage of the analytical techniques when compared to the numerical methods.

Numerical integration, Runge-Kutta, Euler and other techniques can obtain direct solutions from systems of partial differential equations. The finite difference method computes the solution by applying Taylor series to approximate the field quantities in points of a mesh grid covering the domain of interest. The finite element approach belongs to the discrete methods and will be the topic of the following chapters. The discrete element methods have the ability to be used in general applications. Therefore, different types of problems can be solved by the same method. They can be solved employing the same numerical structure.

To solve a technical field problem numerically, an appropriate method has to be chosen. The most important methods are listed here:

- finite element method (FEM)
- finite difference method (FDM)
- boundary element method (BEM)
- magnetic equivalent circuit (MEC)
- point mirroring method (PMM).

Table 5.1. Numerical field computation methods.

method	principle of discretisation	geometry approximation	non-linearites	computational costs
FEM		extremely flexible	possible	high
FDM		inflexible	possible	high
BEM		extremely flexible	troublesome	high
MEC		specific geometries	possible	very low
PMM		simple geometries	by constant factors	low
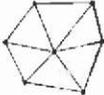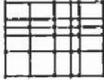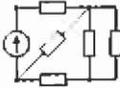
Today, the FEM is a well known method and is commonly used for electromagnetic field problems. The two most popular methods for deriving the finite element equations are the variational and the Galerkin approach, being a special case of the method of weighted residuals. For a two-dimensional analysis, the domain of interest is discretised into a number of simple triangular or rectangular elements, the finite elements, with homogenous properties. For three-dimensional problems, tetrahedra or other simple volume elements are used. The potential function is approximated in those finite elements by simple shape functions, mainly linear or quadratic. This results in a large linear system of equations. Saturation effects can be considered easily. Using triangular elements for two-dimensional field problems and tetrahedra in three dimensions, a very good approximation of the geometry is obtained. The FEM is the most flexible method when compared to all the other techniques listed in Table 5.1.

Historically the FDM is the oldest method. Here, the domain of interest is discretised by a grid with discrete points. The differential equation of the particular field problem is locally transferred into a difference equation. This leads to a linear system of equations to be solved. The solution at the grid points approximates the field.

Due to the discretisation of the domain by a grid, in xy or polar co-ordinates, this method is in some cases troublesome in accurately approximating the geometry (Table 5.1). A local grid refinement to increase the solution accuracy can not be obtained in an efficient way. Non-linearities can easily be implemented using Newton iteration schemes. A three dimensional FDM is possible under the same restrictions as mentioned before. The FDM lost its importance and is nowadays used for problems in the time domain only and is still popular in fluid dynamics.

Using a particular approach at the boundaries of a field domain including the solution of the fundamental system (Green functions) of the partial differential equation represents the basic idea of the BEM. Using this connection, only the boundaries of the region of interest are discretised. Due to this discretisation the geometry of a domain can be approximated very accurately. The coefficient matrix of the BEM is, in contrast to the FDM and FEM, completely filled, non-symmetric and not positive definite. Therefore, special solvers have to be used for the resulting system of equations. Non-linearities are very difficult to account for.

The equivalence of the steady state electrical flow field and electromagnetic field is exploited in the MEC. This method recommends

a principal knowledge of the field distribution. Here, the field domain is discretised by lumped parameter elements, representing the reluctances and sources of the problem. To obtain a potential at the nodes of this network, a rather small linear system of equations has to be solved. Therefore, the MEC is fast. The low level of discretisation, when compared to the FEM, results in an acceptable accuracy for field quantities. The computation of forces in electrical machines is troublesome, as a derivative of the reluctivity is required. Non-linearities can be implemented easily.

The PMM exploits the analogue formulations of the magnetostatic and electrostatic field problems. The PMM has its origin in analytical field calculations. Permanent magnets are considered by magnetic surface charges and are mirrored at the boundaries of the region of interest following the rules of the electrostatic field. This method is very fast, but restricted to very special geometries. Saturation can be considered by constant permeabilities only. An advantage using this method is the fact that special problems (3D high voltage transmission lines, 3D permanent magnet constructions) can be solved relatively easily and fast, whereas the FEM or other discretisation methods require huge efforts to define the problem, at high computational costs.

The advantages and disadvantages of the above-mentioned methods are collected in Table 5.1. In the following section examples are given for the MEC and PMM to demonstrate their strength in selected problem classes of the design of electromagnetic devices.

5.1 Magnetic equivalent circuit

advantages	disadvantages
fast	simple geometries only
easy to implement	flux paths must be known to build up the model
non-linearities possible	force computations are troublesome

5.1.1 Computation of field quantities of an electromagnetic actuator

For the optimisation of magnetic circuits by numerical methods, fast field computation algorithms are recommended. If the field problem is not too complicated, the MEC can be employed. Here, the computation of an actuator with permanent magnets is discussed.

Using the formal equivalence of the electric flow field, the magnetic field components of an electromagnetic device can be obtained using a magnetic equivalent circuit. With the rules of the circuit theory, this model of the electromagnetic circuit is solved. Compared to the finite

element method, this approach offers the ability to obtain accurate results with low computational costs. Inherent non-linearity due to the characteristic of the ferromagnetic parts in the magnetic circuit is implemented. Non-linearities caused by the relative displacement between moving parts are implemented as well. Results obtained by simulations are compared with measurements on a small permanent magnet-excited actuator (Fig. 5.2).
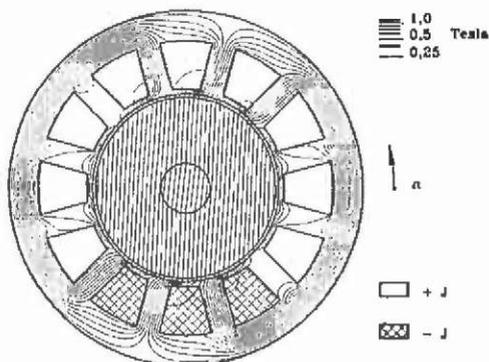
Fig. 5.2. Geometry and flux plot of the actuator computed by FEM.
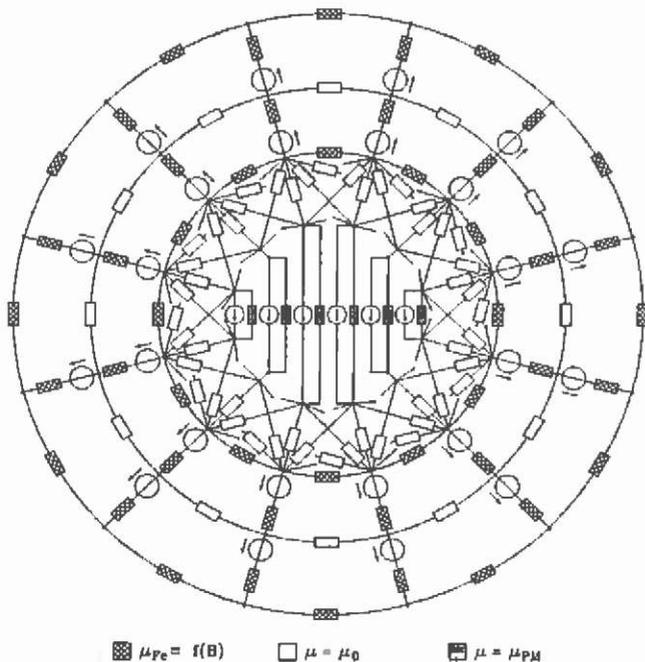
Fig. 5.3. Complete magnetic equivalent circuit.

A two-pole diametrically magnetised permanent magnet rotor ring is centred in the stator bore. The two-phase armature winding is arranged in closed stator slots.

The equivalent magnetic circuit (Fig. 5.3) consists of magnetic resistors, defined by flux tubes, and flux and/or mmf sources. The solution of this field problem is the analysis of a non-linear network.

$S(x)$ is the area perpendicular to the direction of the flux Φ at the position x where ϕ_1 and ϕ_2 are the magnetic potentials at both ends of the flux tube. The difference $\phi_1 - \phi_2$ corresponds to the magnetic voltage drop along the flux path. The magnetic resistor R_m for the equivalent magnetic circuit is

$$R_m = \frac{1}{\Lambda_m} = \int_0^l \frac{dx}{\mu(x)S(x)} \qquad . \tag{5.1}$$

Since in iron parts of the electromagnetic device the permeability $\mu(x)$ is a function of the flux density, the field problem is non-linear. Permanent magnet material with its demagnetisation characteristic is included in the equivalent magnetic circuit as well. An evaluation of Ampere's law leads to the mmf sources modelling the windings in an electromagnetic device.

To solve the network problem, a node-based method is used, enabling, when compared with branch-oriented algorithms, a more easy assembly of the node permeance matrix. The solution of three dimensional networks is possible as well. As the result of this network circuit analysis, the potentials at the nodes are obtained. From the known node potentials, the interesting field quantities can be derived. The advantages of the method used are:

- direct assembling of the system of equations
- diagonal dominant coefficient matrix
- sparse system
- no restrictions to planar graphs.
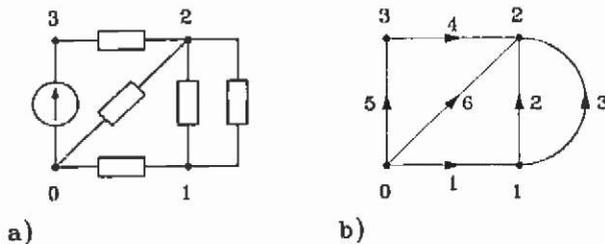
Fig. 5.4. a) Network and b) directed graphs.

The clearest and most flexible way to describe the structure of a network is by setting up a graph or mathematical topological matrices. The topology describes the properties of the network concerning its structure without considering the properties of the network elements. The structure is modelled by a directed graph (Fig. 5.4b). The incidence matrix A of the directed graph describes the topology of the network. Its columns indicate the branch number and its rows the node number. The elements of

$$
\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1z} \\ a_{21} & a_{22} & \cdots & a_{2z} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kz} \end{pmatrix}
\tag{5.2}
$$

are

$$
a_{ij} = \begin{cases} 1 : \text{if branch } j \text{ is directed away from node } i \\ -1 : \text{if branch } j \text{ is directed to node } i \\ 0 : \text{branch } j \text{ is not incident with node } i \end{cases}
\tag{5.3}
$$

The subscript k denotes the number of nodes and z the number of branches.

After introducing the vectors, Φ_0 modelling the flux and Θ_0 giving the mmf sources in the network, and with the diagonal matrix $\mathbf{D}$ representing the branch permeances, the complete system of equations is

$$
\mathbf{ADA'}\phi = \mathbf{A}(\Phi_s - \mathbf{D}\Theta_s) \quad .
\tag{5.4}
$$

In (5.4) the vector ϕ contains the required node potentials for further consideration. The system (5.4) can be solved either by direct or by iterative methods.

The network of a magnetic equivalent circuit consists mainly of non-linear elements where the permeance of a flux tube depends via the permeability μ on the flux density. Therefore, the flux density $\mathbf{B}$ as a function of field strength $\mathbf{H}$ of such elements must be given and incorporated into the solution process. The Newton algorithm obtains the iterative solution. The iteration instruction for iteration step (k+1) is

$$
\phi^{(k+1)} = \phi^{(k)} - \left\{ \mathbf{J}(\phi^{(k)}) \right\}^{-1} \mathbf{F}(\phi^{(k)}),
\tag{5.5}
$$

where $\phi^{(k)}$ represents the solution vector containing the node potentials from iteration step k, $\mathbf{J}(\phi^{(k)})$ is the Jacobi matrix and $\mathbf{F}(\phi^{(k)})$ the fundamental system. To assemble the Jacobi matrix with the term $\partial B / \partial H$ the given non-linear material characteristic B=f(H) is evaluated by cubic spline interpolations. This enables the use of the term directly

because the derivative is already included in the interpolation algorithm and thus available without additional numerical expenses.

When computing the magnetic field quantities mainly the generated forces and/or torques of an electromagnetic energy converter is of interest. The electromagnetic torque of the actuator is calculated using the energy principles according to virtual work.

Computations with increasing winding currents are performed to verify the accuracy of the magnetic equivalent circuit model at different saturation levels inside the iron parts. Fig. 5.5 shows a very good agreement between the methods. For the computed and measured torque versus position, a good agreement is found as well.

The computation time to solve the non-linear magnetic equivalent circuit with 210 elements (Fig. 5.3) is of the order of seconds.
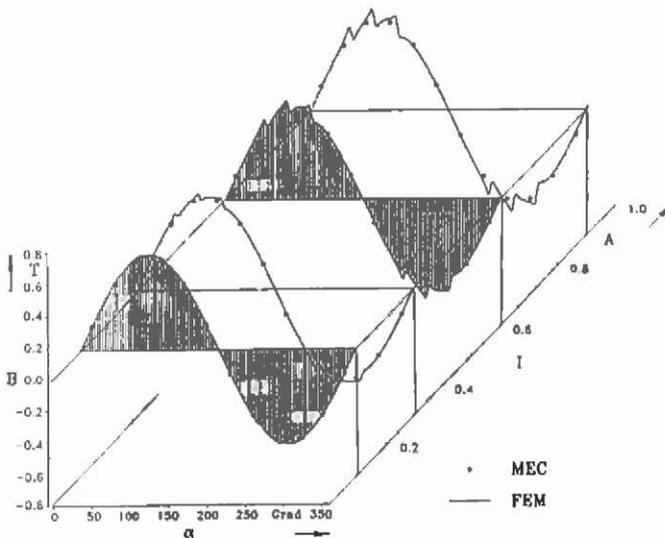
Fig. 5.5. Air gap flux density with different winding currents.

5.2 Point mirroring method

advantages	disadvantages
relatively fast	non-linearities only considered by constant factors
3D fields	special geometries only
	special boundary conditions have to be assumed

To demonstrate the strength and shortcomings of this method, on the other hand, two examples are worked out. In the first example a ferromagnetic circuit is calculated. The main limitations in this example

are seen in the assumption that the iron circuit is piecewise constant saturated. In contrast to this shortcoming, the geometry of the disc-type motor is rather complicated and requires a three-dimensional calculation of the field. The assumption of a plane mirror surface can be seen as a very strong boundary condition thus limiting the application range. Therefore, no slots are allowed in this type of model. An air gap winding is recommended in order to be able to compute the air gap field in this machine.

The second example, a high voltage transmission line, is due to the slag of the line, an inherently three-dimensional field problem as well, but in this case linear. The assumption of a plane mirror surface, the ground plane below the line, is here the strong limitation. As a consequence, the field can not be calculated in a hilly neighbourhood. Nevertheless, measurements and computations are in good agreement.

5.2.1 Computation of the field quantities of a disc-type motor

In this section the basic ideas of the method of point mirroring are given. With a relatively small amount of computation time the field quantities of very complicated geometries can be studied. The method permits three-dimensional magnet field calculations. Here, the method is used to calculate the flux distribution of a disc type motor. The rotor of the motor consists of a NdFeB permanent magnet ring. An air gap armature winding is fixed to the stator. Fig. 5.6 shows the construction of the motor.

1 distance bolt
2 guard ring
3 magnet
4 back iron
5 position bush
6 distance bush
7 shaft
8 bearing guard
9 bearing
10 bearing cap
11 yoke holder
12 armature winding
13 winding holder
14 guard ring
15 bearing
16 plate anchor
17 coil end-winding
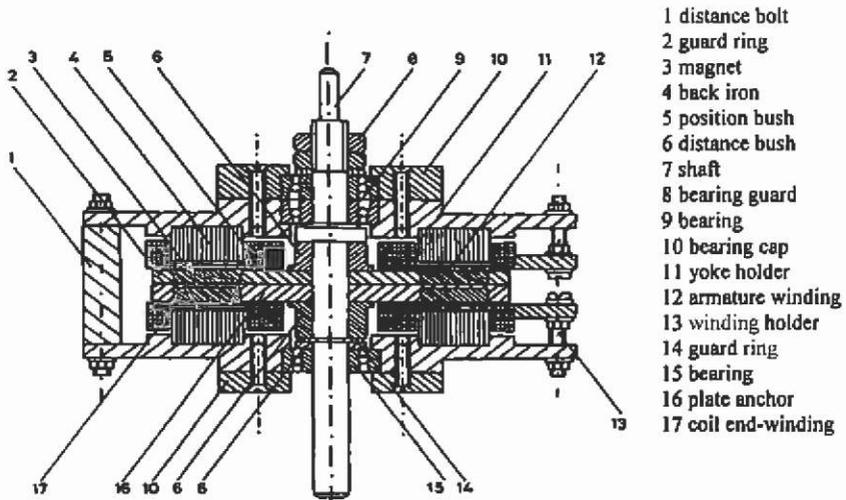
Fig. 5.6. Construction of the studied disc type motor (Hanitsch et al.[54]).

Four poles (2p=4) are axially magnetised on both rotor magnet rings. The used magnet material is NdFeB. In the case of rare earth material, the assumption of a straight-lined characteristic of the demagnetisation curve of the magnet material is realistic. In the operating range of the magnet, the magnetisation is almost independent of the demagnetising field strength.

The magnetisation of a volume is the sum of all dipole moments m_i divided by the volume v. Two different computational models, respectively representations for the magnetisation **M** are possible:

- distributed currents (Fig. 5.7),
- distributed magnetic charge (Fig. 5.8).

The magnetisation is the effect of all elemental currents inside a magnetic medium. The circulating current of one dipole cancels the current of the neighbouring dipole, if the dipoles are parallel and have the same magnitude. If, further, all dipole moments are uniformly distributed throughout the volume, all volume currents I_v vanish except the current at the surface, the surface current density I_s (Fig. 5.7). In this model, the distribution of magnetic dipole moments is equivalent to the distribution of currents at the surface of a magnetic medium and within the volume.

With the Maxwell equations the static magnetic field can be expressed as a solenoid field of flux density and the curl of the magnetic

$$\mathbf{M} = \frac{1}{v}\sum_i m_i$$

$$I_s = \mathbf{M} \times \mathbf{n}$$
$$I_v = \nabla \times \mathbf{M}$$
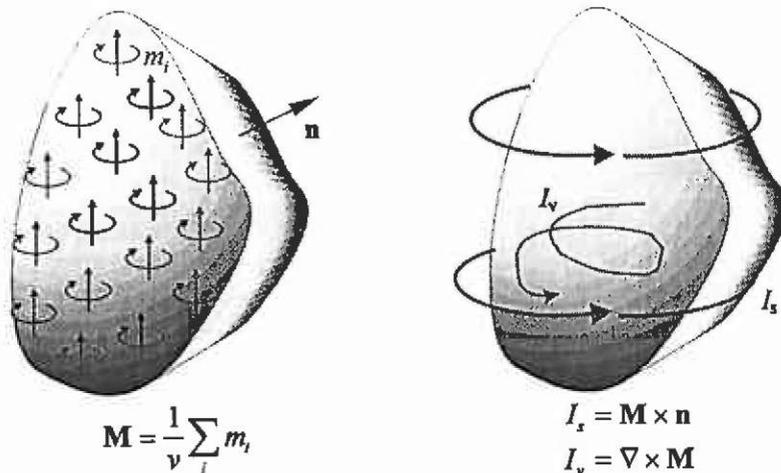
Fig. 5.7. Magnetic magnetisation **M** with current model.

field strength:

$$\nabla \cdot \mathbf{B} = 0$$
$$\nabla \times \mathbf{H} = \mathbf{J} \quad . \tag{5.6}$$

Under the assumption of no field exciting currents inside the permanent magnets, it can be written:

$$\nabla \times \mathbf{H} = 0 \quad . \tag{5.7}$$

With vanishing $\nabla \times \mathbf{H} = 0$ formally the magnetic field strength can be described as a gradient field of a scalar potential function.

$$-\nabla \varphi_m = \mathbf{H} \quad . \tag{5.8}$$

Using the demagnetisation characteristic of permanent magnet material $\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M})$ it can be formulated

$$\nabla \cdot \mathbf{B} = \mu_0(\nabla \cdot \mathbf{H} + \nabla \cdot \mathbf{M}) = 0 \tag{5.9}$$

yielding

$$\nabla \cdot \mathbf{M} = -\nabla \cdot \mathbf{H} = \nabla \cdot \nabla \varphi_m = \Delta \varphi_m \quad . \tag{5.10}$$

Formally $\nabla \cdot \mathbf{M} = \Delta \varphi_m$ is appropriate to the Poisson-differential equation of the electrostatic field. Analogous to the electric space-charge density and to the electric surface charge an auxiliary magnetic quantity can be defined.

$$\sigma_m = -\nabla \cdot \mathbf{M} \quad . \tag{5.11}$$

With known inner magnetisation $\mathbf{M}$ of the magnet, the divergence of the magnetisation can be identified as an auxiliary magnetic surface charge σ_m.
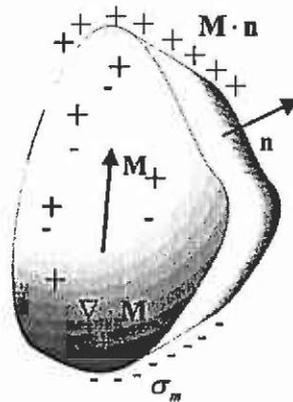
Fig. 5.8. Permanent magnet shape with auxiliary magnetic surface charge.

As shown in Fig. 5.8 for the upper pole surface, this yields

$$-\nabla \cdot \mathbf{M} = -\mathbf{n}(-\mathbf{M}) = \mathbf{M}_0 \quad , \tag{5.12}$$

and for the lower pole surface

$$-\nabla \cdot \mathbf{M} = -\mathbf{n}(\mathbf{M}) = -\mathbf{M}_0 \quad . \tag{5.13}$$

In this consideration, the laws of electrostatic field can be used to evaluate the scalar potential $\varphi_m(P)$ and magnetic field strength H(P) in a point P. Integration is performed over the surface of the north A_N and the south pole surface A_S.

$$\varphi_m(P) = \frac{\mu_0}{4\pi} \iint_{A_N} \frac{\mathbf{M}_0}{|\mathbf{r}_{pq}|} dA + \frac{\mu_0}{4\pi} \iint_{A_S} \frac{-\mathbf{M}_0}{|\mathbf{r}_{pq}|} dA \quad ,$$

$$\tag{5.14}$$

$$H(P) = \frac{-\mu_0}{4\pi} \iint_{A_N} \nabla(\frac{\mathbf{M}_0}{|\mathbf{r}_{pq}|}) dA + \frac{-\mu_0}{4\pi} \iint_{A_S} \nabla(\frac{-\mathbf{M}_0}{|\mathbf{r}_{pq}|}) dA \quad .$$

To determine the flux density of the ring-formed permanent rotor of the mentioned disc-type motor, the ring is subdivided into trapezoidal single magnet elements (Fig. 5.9).
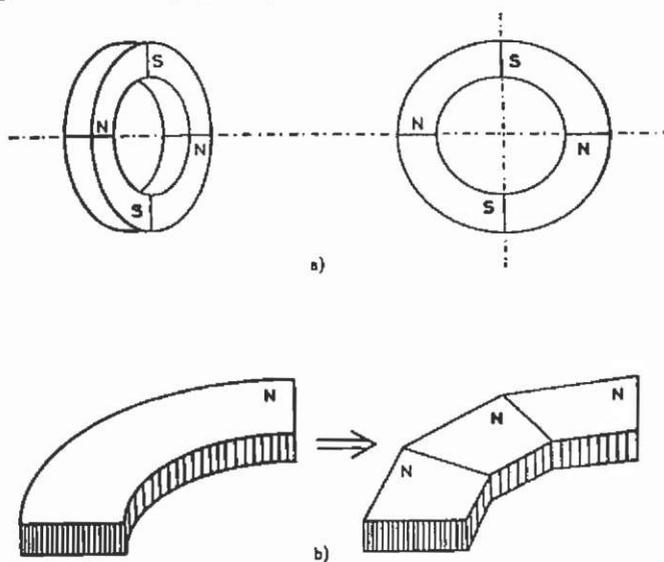
Fig. 5.9. Permanent magnet ring (Walkhoff [119]).

The superposition of elementary and simple magnet shapes is performed in order to form a complicated-shape permanent magnet. Here, the superposition of the field components of cubed and triangle magnet elements result in the required trapezoidal magnet as indicated in Fig. 5.10. The magnetic field excited by the permanent magnet ring can now be calculated at every point P outside the magnet volume.

To consider the ferromagnetic back iron inside the real machine, the magnetic surface charges have to be mirrored at the boundaries of the air gap. The laws governing the electrostatic field can be used here. A mirror interval of 4 to 5 steps is sufficient to obtain an acceptable accuracy (Walkhoff [119]).
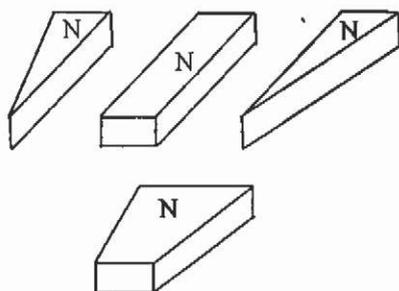
Fig. 5.10. Superposition of cube and triangle to obtain a trapezoidal shaped magnet.

With these considerations, the auxiliary configuration for the calculation of the air gap field of the disc-shaped motor, using the point mirroring method, can be constructed (Fig. 5.11).
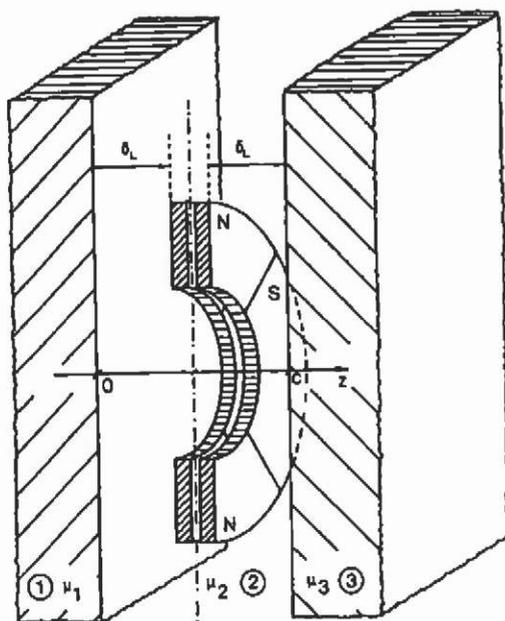
Fig. 5.11. Auxiliary configuration and co-ordinate system (Walkhoff [119]).

Mirroring the mentioned surface charges is done at the material boundaries having permeabilities μ_1 and μ_3 (Fig. 5.11). Saturation is considered by the constant factors.

The z-component of the resulting flux density distribution of a pole pitch of the mentioned disc-type motor can be taken from Fig. 5.12.
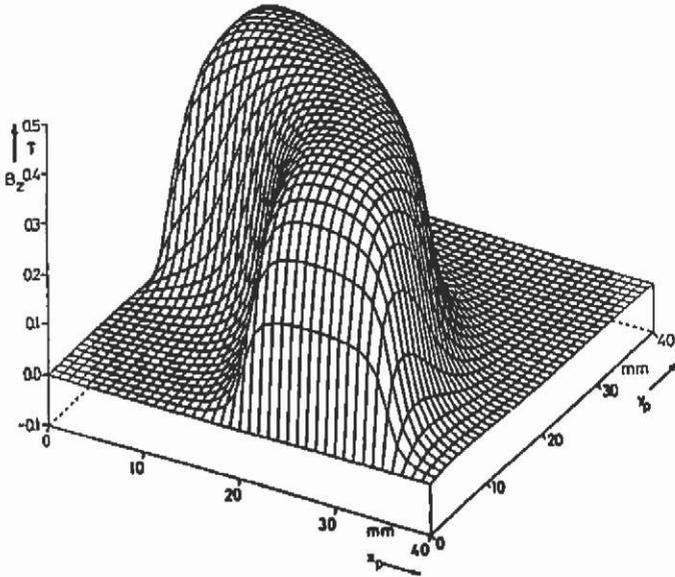
Fig. 5.12. Three-dimensional air gap flux density distribution (Walkhoff [119]).

5.2.2   Computation of the fields below AC high voltage lines

Overhead transmission lines generate in their vicinity electric and magnetic fields. The source of the magnetic field is the current in the phase conductors. The electric field is caused by the high potential at the phase conductors.

The problem specifies small diameter conductors above a large flat conducting ground plane. The phase conductors are at a time-dependent specified electrical potential and carry a time-dependent current. Due to the slag of the phase conductors, the field problem turns out to be three-dimensional. Only symmetric three-phase voltage and current systems are considered. The ground below the transmission line is a uniform plane.

The field problem may be considered as quasi static. Therefore, the solution can be determined by static techniques. With respect to the slag of the phase conductors, infinitesimally thin, segmented filaments

approximate the geometry of a single conductor. Due to the symmetry between two poles, one half of the arrangement is drawn in Fig. 5.13 only. The value of *s* indicates the slag, *l* is the distance between the two high voltage poles, the span field length.
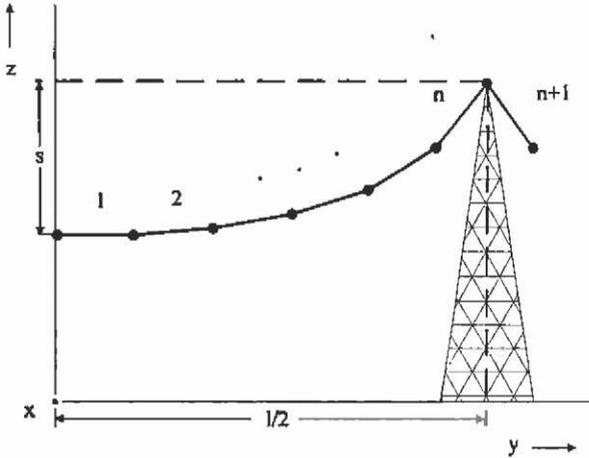
Fig. 5.13. Geometric modelling of a conductor.

5.2.2.1   Electric field The electric field is computed by mirroring single line charges at the assumed to be ideal conducting ground plane below the phase conductors. Each infinitesimally thin filament segment represents in this case a line-charge. A constant line-charge at any position in the original co-ordinate system (x, y, z) is drawn in Fig. 5.14. To evaluate the field quantities of the line-charge, this co-ordinate system has to be transformed into a system ($\widetilde{x}, \widetilde{y}, \widetilde{z}$). This transformation is performed in two steps. The first step consists of a parallel shift of the origin into the starting point of the line-charge. In a second step a rotation of this temporary co-ordinate system (x°, y°,z°) around the x°-axis is carried out in such a way that the line charge lies in the x°-y° plane. The last rotation in this step is around the z°-axis so that the line-charge lies in the x°-axis. In this co-ordinate system the potential φ of the line-charge in the point P($\widetilde{x}, \widetilde{y}, \widetilde{z}$) is given by:

$$\varphi(\widetilde{x}, \widetilde{y}, \widetilde{z}) = \frac{q}{4\pi l \varepsilon} \ln\left[\frac{l - \widetilde{x} + \sqrt{\widetilde{y}^2 + \widetilde{z}^2 + (l - \widetilde{x})^2}}{-\widetilde{x} + \sqrt{\widetilde{x}^2 + \widetilde{y}^2 + \widetilde{z}^2}}\right] \qquad . \qquad (5.15)$$
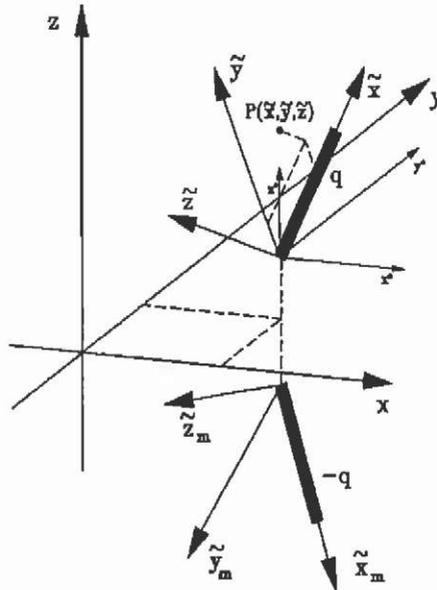
Fig. 5.14. Co-ordinate transformation of an infinitesimally thin filament segment.

To evaluate the field quantities with respect to this boundary condition, the line-charge has to be mirrored with respect to the plane x-y. Superposition of line-charge q and mirror-charge $-q$, indicated in Fig. 5.14, gives the potential φ at the point $P(x, y, z)$ inside the global co-ordinate system.

To consider the slag of the conductors, a quadratic approximation is used. Referring to Fig. 5.13, it can be written as

$$f(y) = s \cdot \left(1 - \frac{4}{l^2} y^2\right) \quad . \tag{5.16}$$

With the known complex potentials $\underline{\varphi}_i$ of the i conductors and transforming eq. (5.15) to compute the coefficient matrix $\mathbf{A}$, a linear set of equations can be formulated.

$$\mathbf{A} \cdot \underline{\mathbf{q}} = \underline{\varphi} \quad . \tag{5.17}$$

The solution determines the charge $\underline{q}_i$ of each element of the conductors. With these values the components of the electrostatic field strength in the point $P(x, y, z)$ can be computed.

$$\mathbf{E} = -\nabla \varphi = -\left(\frac{\partial \varphi}{\partial x} \mathbf{e}_x + \frac{\partial \varphi}{\partial y} \mathbf{e}_y + \frac{\partial \varphi}{\partial z} \mathbf{e}_z\right) \quad . \tag{5.18}$$
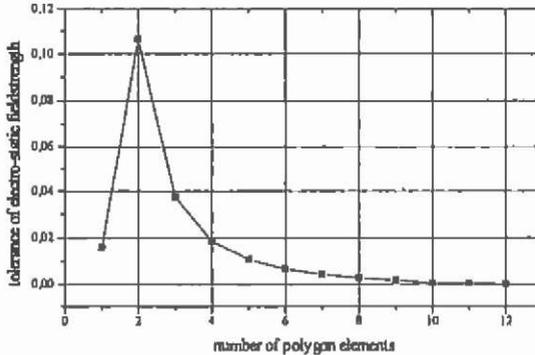
Fig. 5.15. Convergence of the simulation with respect to the slag of the high-tension line.

To illustrate the convergence behaviour of the method, attention should be paid to Fig. 5.15. With an increasing number of infinitesimally thin filament segments for one half of the span field, the electrostatic field strength converges to the correct value. Calculations with 5...7 polygon elements deliver results with a reasonable accuracy at acceptable computational costs.

5.2.2.2   Magnetic field The magnetic field problem is considered to be linear. Hence, the superposition of partial fields, calculated with the Biot-Savart law, result in the overall three-dimensional field distribution below the line.

In this case each segment of the infinitesimally-thin filament (Fig. 5.14) carries a current $i(t)$. The generated flux density of this part of the conductor is

$$|d\mathbf{B}| = \frac{\mu_0}{4\pi r^2} \cdot i(t) \cdot dl \cdot \sin\alpha \quad . \tag{5.19}$$

The point where the flux density has to be calculated has to be transformed into the co-ordinate system ($\tilde{x}, \tilde{y}, \tilde{z}$). After integrating eq. (5.19) the flux density is calculated with

$$|\mathbf{B}_j| = \frac{\mu_0}{4\pi r} \cdot i(t) \cdot \left(\frac{l - \tilde{x}}{\sqrt{(l-\tilde{x})^2 + r^2}} + \frac{\tilde{x}}{\sqrt{\tilde{x}^2 + r^2}} \right) \quad . \tag{5.20}$$

If n is the number of current carrying conductors, superposition of the individual flux densities results in the overall flux density:

$$\mathbf{B} = \sum_{j=1}^{n} \mathbf{B}_j \quad . \tag{5.21}$$

5.2.2.3   Finite   element   model   The   second   numerical   field
computation method able to compute solutions in this problem class is
the finite element method. Special boundary conditions applied to the
field problem result in an effective use of this method. The application of
open boundary conditions gives the opportunity to discretise the field
problem in regions of interest only. This results in lower computational
costs. With respect to the computational efforts, only two-dimensional
computations are performed with this numerical method.
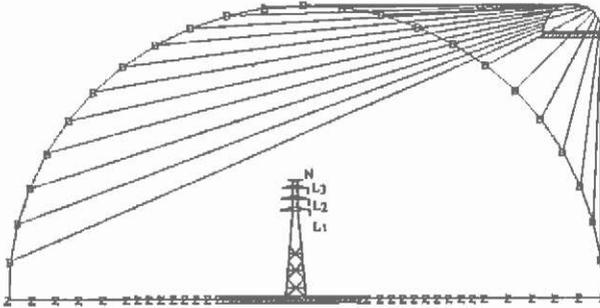
Fig. 5.16. Open boundary model to compute the electric field of a 150 kV AC
single system transmission line. (the triangulation of the domain is invisible)

A cross-section of the transmission line is made at the place where
the wires are nearest to ground level (Fig. 5.16). Here, the highest field
values   are   expected.   A   two-dimensional   finite   element   model
perpendicular to the line is built. The region of the cross-section is
subdivided in triangular finite elements (Fig. 5.17). The potential
distribution over each element is approximated by a polynomial. Instead
of solving the field equations directly, the principle of minimum potential
energy is used to obtain the potential distribution over the whole model.
The ratio of the largest size of a finite element to the smallest size in the
model of a transmission line is about 10,000. The circular boundary of
the model has a radius of about 100m, while the radius of the conductors
is a few centimetres. Therefore, special attention must be paid to obtain a
regular mesh with well-shaped elements, ensuring an accurate solution of
the field problem. Thus, a high degree of discretisation resulting in a
large system of equations must be applied. Fig. 5.17 shows a part of the
finite element model around one of the phase conductors. The change in
the size of the elements in the direction away from a conductor can be
noticed (Fig. 5.17).

When computing the electric field strength, it is assumed that the
ground plane below the transmission line is an equipotential surface.

Therefore it is not necessary to discretise the ground as indicated in Fig. 5.16. In contrast, in the magnetic field model the ground has the same magnetic properties as the surrounding air and has to be discretised as well. This results in an increased number of finite elements and thus in higher computational costs.

In contrast to the semi-numerical method, where the phase conductors are modelled taking only a few seconds to compute the field quantities, the calculation time of one transmission line on a PC-486 platform using the finite element method is about 30 minutes.
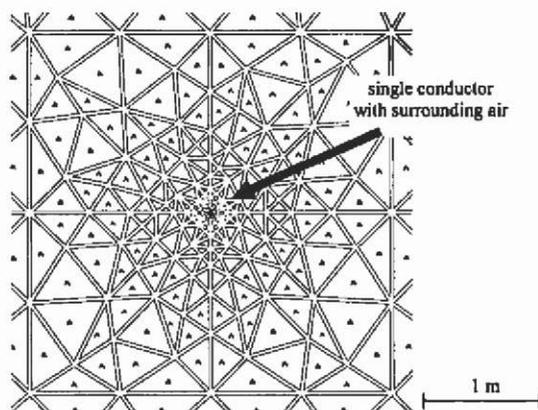
Fig. 5.17. Part of the finite element model around a single phase conductor.

5.2.2.4   Measurements The measurement of the electric field strength excited by the transmission line are based on the induced current of the charge oscillations between two halves of an isolated conductive body.

The measurements of the magnetic field strength are based on the electromotive force induced in a coil. Therefore, the probe of the field meter, Holaday Industries model HI-3604, consists both of two circular isolated parallel plates and of a circular coil. To avoid perturbations of the electric field, a fibre optic receiver and a non-conductive tripod to support the field meter are used. Only the rms value of the space component perpendicular to the plane of the probe is measured. The field quantities below the overhead transmission lines are measured at a height of 1m above ground level.

5.2.2.5   Numerical and experimental results All computations and measurements on a Belgian 150 kV AC single three-phase system transmission line are performed at the place of the maximum slag.

Because the transmission line is situated in a flat area in Belgium, the ground level is assumed to be even.

To obtain the local field values with reasonable accuracy, a third order finite element solution is necessary. The use of shape functions of third order explains the long computation time. Fig. 5.18 shows the x- and z-component of the rms value of the magnetic field. Good agreement between the measurements and computed data can be stated. The two-dimensional approach overestimates both x- and z-component of the magnetic field. The reason for this lies in the type of approximation of the geometry of the transmission line. In the two-dimensional model a phase conductor of infinite length with constant height above the ground is considered. Therefore, the two-dimensional approach represents the worst case, i.e. the highest values of field strength.
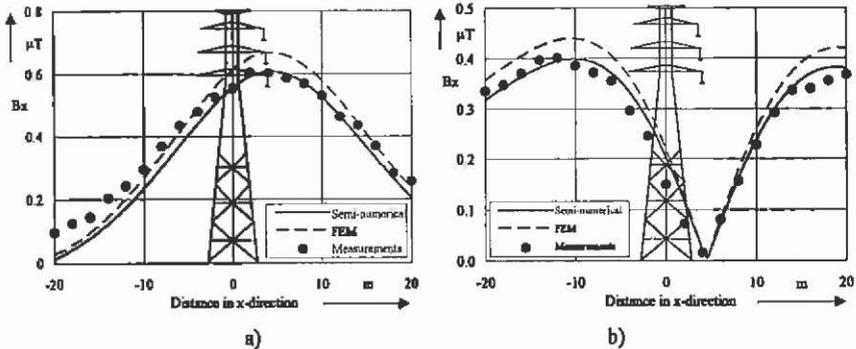
Fig. 5.18. Comparison of computed magnetic field distribution and measured data 1m above ground a) x-component and b) z-component.

Fig. 5.19 shows the effective value of the z-component of the electric field. The calculations and the measurements show good agreement.
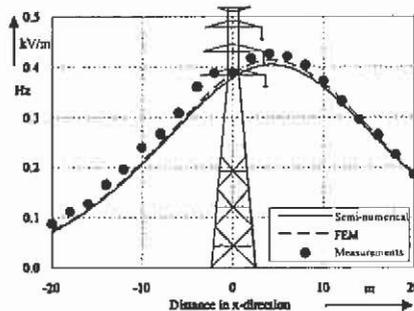
Fig. 5.19. Comparison of computed electric field distribution with measured data 1m above ground level.

In Fig. 5.20 the three-dimensional electric field distribution below the 150 kV ac transmission line obtained by the PMM method is plotted. The geometrical model of the single phase conductors consists of seven polygon elements per half of the overall span field. The high-voltage pole is located at the global co-ordinates x=0m and y=220m.

As expected, the maximum field values are found in the middle of the field span at the co-ordinates y= 0m. Here, the values of the electric field strength are in the range of 4 kV/m and thus well below the maximum allowed exposure values for the general public, given by the standards in Table 5.2.
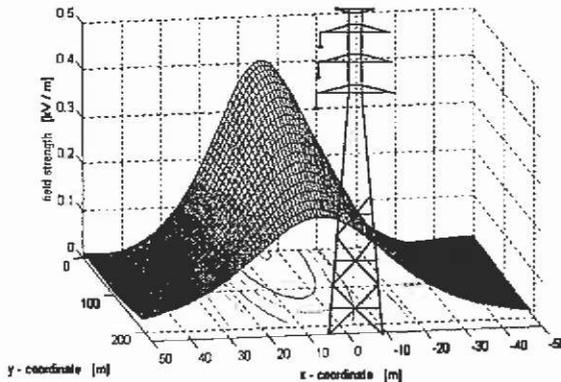
Fig. 5.20. Electric field distribution below the 150 kV AC high-voltage line computed by the point mirroring method.

Referring to Fig. 5.18, the magnetic flux density of the x- and z-component, generated by the current carrying conductors, are in the range of 0.4–0.6 μT. According to Table 5.2 those values of the magnetic flux density are far below the allowed limits as well.

Due to the linearity of the problem formulation, calculations of power lines with different types of AC-high-voltage poles carrying multiple three-phase voltage and current systems can be performed. In this case the field components generated by the single systems have to be superposed according to the relative phase angle between the systems and the considered instant of time. Fig. 5.21 shows the results, computed by the semi-numerical technique, of a high-voltage transmission line consisting of six three-phase systems with different voltage level (2x380 kV, 2x220 kV, 2x110 kV). For the magnetic flux density it is assumed that each system carries a current of 1000 A. The system with the largest transmission voltage is put at the top of the pole, while the system with the lowest voltage is located below it.
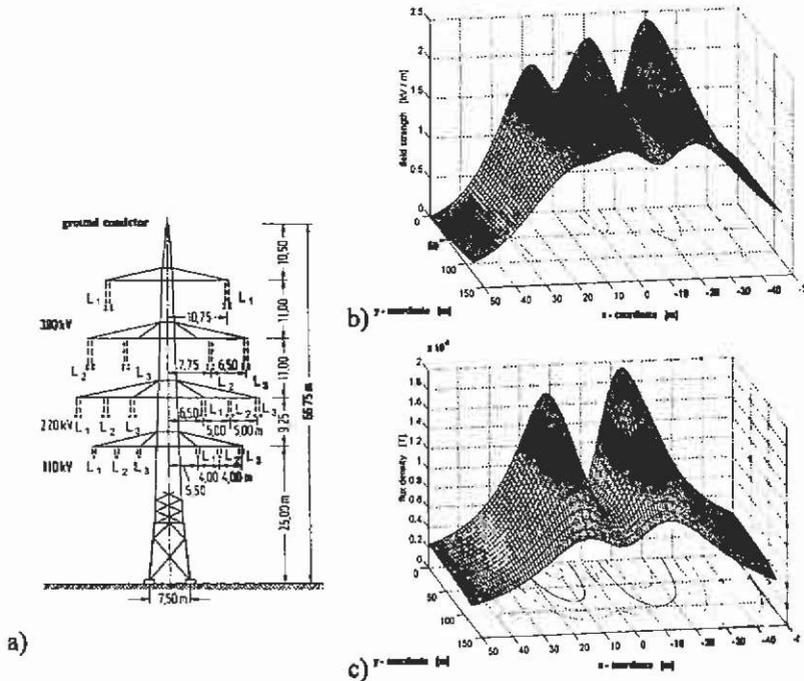
Fig. 5.21. a) High-voltage pole construction carrying 6 systems (2x380 kV, 2x220 kV, 2x110 kV) and b) the resulting electric field distribution. (All the three-phase current systems carry a rms. current of 1000 A; the pole is located at the co-ordinate x=0m, y=160m.)

5.2.2.6   Comparison of PMM and FEM Two efficient methods of computing the electric and magnetic fields below AC-high-voltage lines are demonstrated by an example; a Belgian 150 kV AC three-phase single system transmission line. Both methods, PMM and FEM, are compared with respect to accuracy and the required computational effort. To verify the results of the field simulations, measurements of a power line have been carried out, giving good agreement between computed and measured data.

In the PMM model, infinitesimally thin segmented filaments of constant charge or current are approximating the slag of the transmission line to solve the electrostatic and magnetic fields. With reasonable accuracy a three dimensional field distribution can be computed. Relatively low computation times are necessary to compute the three-dimensional field distribution below the power line with the PMM.

Using a standard PC-486/66, the calculation time lies in the range of seconds.

With the finite element method, the distribution of both electric and magnetic field quantities, is computed as well. With respect to the high computational costs when compared to the PMM, a two-dimensional approach in the middle of the span field is chosen. Due to the necessary high discretisation of the problem, the computational costs are in the range of thirty minutes using a PC-486/66. Good agreement between measured data on the Belgian 150 kV line and calculated field distributions by both methods can be stated.

The main problem employing the FEM is the small diameter of the conductors above the large flat conducting ground plane. The difference between the dimensions of a conductor and the field domain of interest is huge, the ratio lying in the range of some 10^5. This causes the generation of a large amount of finite elements and thus an enormous computation time, even for the two-dimensional problem, when compared to the efforts necessary for the PMM. A three-dimensional FEM model is difficult to build, due to this huge difference in geometrical dimensions.

This example demonstrates that for problem types, such as the high voltage line, the FEM is not very well suited. In this case, the approach using the PMM is the better choice with respect to computational time, problem dimension (2D/3D) and discretisation problems.

5.2.2.7 Effects of fields with low frequency To evaluate the influence of the transmission line, it is not sufficient to calculate the coupling impedances or capacitances of the line. It is necessary to analyse the actually generated fields in the neighbourhood of the transmission line during the planning phase and to check if given standards for maximum field values are violated.

The interaction of electromagnetic fields with living organisms can be separated into two mechanisms, thermal and non-thermal interactions. Thermal interactions mean the mechanism of the absorption of electromagnetic energy resulting in an increasing temperature. Non-thermal are these interactions where the absorbed energy is not large enough to cause a significant temperature rise. In fields at low frequency, the body does not absorb or negligibly absorbs the wave energy. This implies that biological effects caused by electric or magnetic fields of low frequency fields are non-thermal. Observed non-thermal effects on human beings can be the stimulation of nerves, upright standing skin hair, visual disturbances ... . The possible results of these effects may depend on the field characteristics which vary in intensity and frequency.

To judge the mentioned effects, existing technical standards supply the quantities of the electric, magnetic and electromagnetic fields as a function of the frequency. This is important, as the interaction of electromagnetic fields and matter strongly depends on the frequency of the considered field.

Nowadays an increasing sensitivity to ecological problems can be stated. An injurious influence to the health of human beings caused by the direct effect of low frequency electromagnetic fields (50/60 Hz) is scientifically not proven yet. For about twenty-five years research efforts to find a correlation mechanism between the field quantities and their effects on human beings have been going on, without significant success. In this situation, the electric and magnetic field quantities of high-voltage lines have to be examined in order to avoid EMC problems with the environment close to the power transmission line while planning high voltage lines.

A number of standards such as those in preparation by the European Committee for Electrotechnical Standardisation (CENELEC) are based on the known effects for short exposure times. Long term effects are not considered. However, to consider possible as yet undiscovered effects the values for technical fields are reduced by a factor. In Table 5.2 the maximum exposure values for the electric and magnetic field are summarised. The values are distinguished according to the general public and professional workers.

Table 5.2. Maximum exposure values for the electric and magnetic field (ICNIRP).

exposure	electric field strength E kV/m (rms)	magnetic flux density B mT (rms)
professionals: 8 hours general public:	10	0,5
	5	0,1

source: Health and Physics, April 1998, Vol.74, No. 4

5.3   The numerical solution of partial differential equations

Solving a differential equation analytically or by semi-numerical techniques, as demonstrated before, is only possible assuming simplifications in the differential equations that are valid in a domain that can be described in a plain mathematical way. Specific boundary conditions and various material properties in different domains make it difficult to obtain an analytical solution for problems of technical importance.

The numerical approximation methods (Table 5.1) of solving the introduced partial differential equations are not limited to such specific geometries or other rough simplifications and can handle different material characteristics with an acceptable accuracy in a single model. The two most important groups of numerical methods are the finite difference method and the projection methods.

The finite element method nowadays is the most important and most frequently used approach solving variational problems and differential equations in engineering. The most significant success of this method is founded in the possibility to develop on its base user-friendly computer programs of general application range. Due to its structured rules this is closely linked to the opportunity of the FEM to generate stable numerical schemes for considering complicated two- and three-dimensional geometries in a relatively simple way.
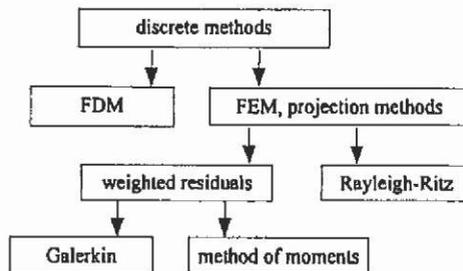
Fig. 5.22. Overview of various numerical methods.

5.4 Finite difference method

The field domain of interest is discretised by a grid, where the grid-lines are in parallel to the co-ordinate axes. This type of mesh is called an orthogonal grid and must not consist of equidistant grid-points (Fig. 5.23). The grid-distances h_E, h_S, h_W and h_N can be different.
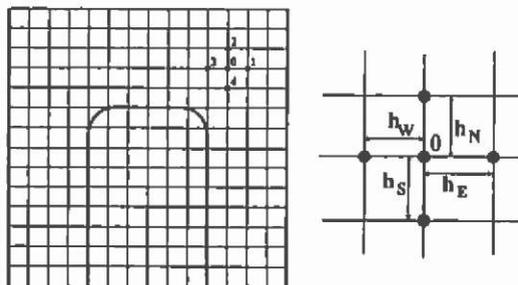
Fig. 5.23. Numerical discretisation for the FDM.
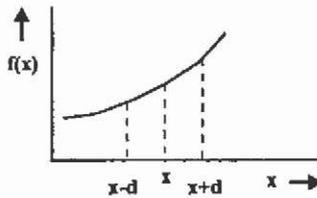
Fig. 5.24. Approximation of the potential distribution by a Taylor series.

The potential distribution over the studied domain is approximated by the first terms of a Taylor series (Fig. 5.24):

$$f(x+d)-f(x)=\sum_{m}\frac{d^{m}}{m!}f^{(m)}(x)+\eta(x) \qquad . \tag{5.22}$$

The differential equation of the particular field problem is locally transferred into a difference equation. For example, calculating the value of a potential distribution at point 0, by using the 4-point approach (Fig. 5.23), and a grid distance d with a derivative

$$\frac{\partial f}{\partial x}=\lim_{d\to 0}\frac{f(x+d)-f(x)}{d} \tag{5.23}$$

is expressed by a finite difference with a known approximation error $\eta(x)$. By considering only the first terms of the Taylor series, this error is known and dependent on the grid distance.

Forward,

$$\frac{\partial f}{\partial x}=\frac{f(x+d)-f(x)}{d}+\eta^{\cdot}(x) \qquad , \tag{5.24}$$

backward

$$\frac{\partial f}{\partial x}=\frac{f(x)-f(x-d)}{d}+\eta^{\cdot}(x) \qquad , \tag{5.25}$$

and central difference

$$\frac{\partial f}{\partial x}=\frac{f(x+d)-f(x-d)}{2d}+\eta(x) \tag{5.26}$$

are used to assemble a system of linear equations to calculate the potentials at all grid points. This leads to a large linear system of equations to be solved. The potentials at the mesh points represent the approximated field solution. To obtain an accurate field solution a fine discretisation is required.

5.5 Finite element method

In general, differential equations are hard to solve. The idea is to find a solution for the overall problem by substituting for the complicated problem a series of simpler ones. This means setting up the problem by a easily-solved linear system of equations.

Therefore, the problem has to be discretised in adequate sub-problems. The sub-problems are geometrically described by geometrically simple shaped elements such as triangles (Fig. 5.25) or rectangles for two-dimensional and mainly tetrahedrons for three-dimensional problems. Other element shapes are possible as well. When comparing the meshes from the FDM and the FEM model in Fig. 5.23, it is obvious that the FEM model approximates better the geometry of the studied domain. A necessary local mesh adaptation is possible in this model as well.
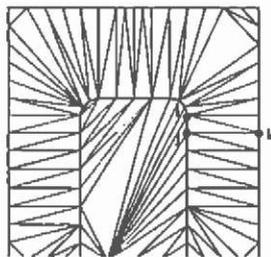
Fig. 5.25. Minimal triangular discretisation of a two-dimensional finite element model.

These elements, forming the numerical discretisation, the mesh, are called the finite elements. On this discretisation, the problem describing differential equation is locally approximated by simple basis function. The approximated overall solution is obtained by assembling all sub-problems into a system of equations and solving this. After this procedure, the approximated potential solution is known in certain points of the discretisation.

The problem is to determine the field describing potential functions for the discrete problem, the finite element equations, and to define an adequate basis or shape function to be able to assemble the overall system of equations. Different methods can be used to determine the finite element equation from the differential equation (Fig. 5.26).

If for particular field problems a variational principle is known, the discrete problem can be obtained by using the Ritz method. In this case, the generation of the discrete problem is easy to obtain. Unfortunately,

this variational principle is not known for every technical problem. In this case, for instance, the very common method of the weighted residues can be applied. The most important finite element methods are:

- various Ritz methods
- variational method
- weighted residual method (weak form of the governing equations)
- different types of Galerkin method
- approaches based on the energy-minimum functional.
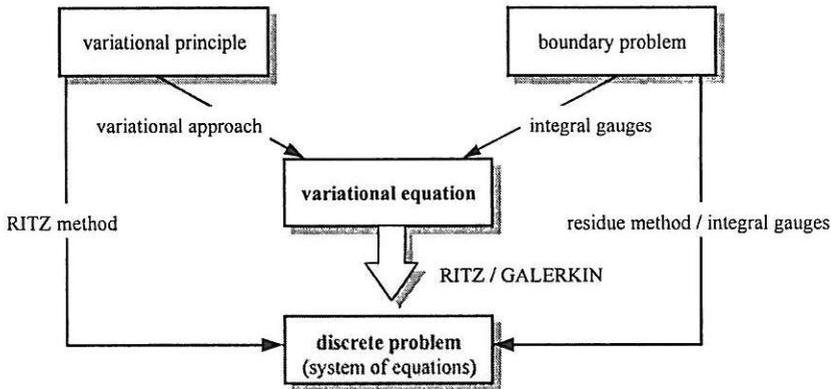
Fig. 5.26. The basic concept of the finite element.

5.5.1   Variational approach

In this section, it is assumed that a variational equation exists for the studied field problem. Therefore, the generation of the discrete problem is easy to obtain.

A problem in the form:
Find a function $u \in V$, so that for all $v \in V$
$$a(u,v) = f(v) \tag{5.27}$$
is called a variational problem.

With $a(u,v)$ is a real value depending on two functions $u, v \in V$, and V is a set of differentiable functions in the field domain Ω with $v=0$ on the boundary Γ. $f(v)$ is a linear form on V.

To determine the variational equation, a Poisson equation is chosen as an example, assuming the appropriate boundary conditions,
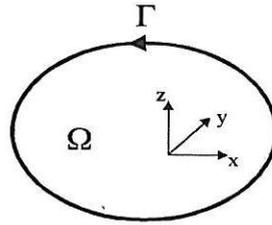
$$\nabla^2 u = \Delta u = -q \quad \text{in } \Omega$$
$$u = 0 \qquad \text{on } \Gamma \quad .$$

with the Laplacian

$$\nabla^2 = \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad , \tag{5.28}$$

and with q, a given continuous and u, a potential function.

The differential equation is multiplied with an arbitrary function $v \in V$ and integrated over Ω to obtain a linear form. It can be written:

$$-\int_\Omega (\Delta u) \cdot v d\Omega = \int_\Omega q \cdot v d\Omega \quad . \tag{5.29}$$

The Gauss-integral gauge transfers a volume integral $\int_\Omega$ into a surface integral $\int_\Gamma$:

$$\int_\Omega (\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}) d\Omega = \int_\Gamma P dy dz + Q dz dx + R dx dy \quad .$$

Γ is a closed and oriented surface, which includes the domain Ω. P, Q, R are functions of three variables defined in Ω. The partial derivative of first order from P, Q, R must exist and must be continuous. With the approach:

$$P = \frac{\partial u}{\partial x} \cdot v = u_x \cdot v \quad ,$$

$$Q = \frac{\partial u}{\partial y} \cdot v = u_y \cdot v \quad , \tag{5.30}$$

$$P = \frac{\partial u}{\partial z} \cdot v = u_z \cdot v \quad ,$$

and considering the boundary condition $v = 0$, the right hand side of the Gauss integral gauge can be written by:

$$\int_\Gamma P dy dz + Q dz dx + R dx dy = \int_\Gamma u_x \cdot \overset{0}{\cancel{v}} dy dz + u_y \cdot \overset{0}{\cancel{v}} dz dx + u_z \cdot \overset{0}{\cancel{v}} dx dy = 0 \quad . \tag{5.31}$$

To evaluate the left side of the Gauss gauge

$$-\int_\Omega (\Delta u) \cdot v d\Omega = \int_\Omega (\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}) d\Omega \quad , \tag{5.32}$$

with the derived functions

$$\frac{\partial P}{\partial x} = \frac{\partial}{\partial x}(\frac{\partial u}{\partial x} \cdot v) = \frac{\partial u}{\partial x} \cdot \frac{\partial v}{\partial x} + \frac{\partial^2 u}{\partial x^2} \cdot v \overset{0}{\nearrow} ,$$

$$\frac{\partial Q}{\partial y} = \frac{\partial}{\partial y}(\frac{\partial u}{\partial y} \cdot v) = \frac{\partial u}{\partial y} \cdot \frac{\partial v}{\partial y} ,$$

$$\frac{\partial R}{\partial z} = \frac{\partial}{\partial z}(\frac{\partial u}{\partial z} \cdot v) = \frac{\partial u}{\partial z} \cdot \frac{\partial v}{\partial z} ,$$

(5.33)

it can be written:

$$-\int_\Omega (\Delta u) \cdot v d\Omega = \int_\Omega (\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y} + \frac{\partial u}{\partial z}\frac{\partial v}{\partial z})d\Omega \qquad . \tag{5.34}$$

With this, the boundary problem described by a Poisson equation is transferred into a variational equation.

$$a(u,v) = \int_\Omega (\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y} + \frac{\partial u}{\partial z}\frac{\partial v}{\partial z})d\Omega \qquad , \tag{5.35}$$

$$f(v) = \int_\Omega q \cdot v d\Omega \qquad . \tag{5.36}$$

5.5.1.1 Discretisation of the differential equation The variational equation $a(u,v) = f(v)$ is applied to the standard example of a Poisson equation:

$$a(u,v) = \int_\Omega (\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y} + \frac{\partial u}{\partial z}\frac{\partial v}{\partial z})d\Omega$$

$$f(v) = \int_\Omega q \cdot v d\Omega \qquad .$$

$w_1, ..., w_N$ are N independent linear functions out of V, and V_h is the set of linear combinations $\sum_{i=1}^{N} c_i w_i$. V_h is called the N-dimensional partial space from V, and w_i are the basis- or global shape functions (Goering et al. [44], Zienkiewicz & Taylor [123]).

An approximated solution from $a(u,v) = f(v)$ is a function $u_s \in V_s$. The idea now is to find an approximated solution u_h, using the linear combination of the N functions

$$u_s = \sum_{i=1}^{N} u_i w_i \tag{5.37}$$

with the unknown u_i.

This yields the formulation of an equivalent problem:

$$a(u_{_h},v_{_h}) = f(v_{_h}) \qquad \text{for all } v_{_h} \in V_{_h} \quad . \tag{5.38}$$

This is a projection of the problem $a(u,v) = f(v)$ in V into a problem in V_h. If this is true for all $u_{_h} \in V_{_h}$, and for all $w_{_j} \in V_{_h}$ it can be written:

$$a(u_{_h},w_{_j}) = f(w_{_j}) \qquad j = 1(1)N \quad . \tag{5.39}$$

Equations (5.38) and (5.39) are equivalent problems and are called the discrete problems.

5.5.1.2 Practical considerations Equation (5.38) was the starting point of the theoretical considerations of the FEM. The initial position for the practical calculation of the approximation u_h is eq.(5.39). Using $u_{_h} = \sum_{i=1}^{N} u_i w_i$ and substituting in the discrete problem,

$$a(\sum_{i=1}^{N} u_i \cdot w_i, w_j) = f(w_j) \qquad j = 1(1)N \quad , \tag{5.40}$$

assuming a being a bi-linear form it can be written:

$$\sum_{i=1}^{N} a(w_i, w_j) \cdot u_i = f_j \qquad j = 1(1)N \quad . \tag{5.41}$$

This represents a system of N equations with the N unknown u_i and the coefficient matrix $\mathbf{A}_h$.

$$\mathbf{A}_{_h} = \begin{pmatrix} a_{11} & \cdots & & a_{1N} \\ & \ddots & & \\ \vdots & & a_{ij} & \vdots \\ & & & \ddots & \\ a_{N1} & \cdots & & a_{NN} \end{pmatrix}$$

$$\tag{5.42}$$

with $\quad a_{ij} = a(w_j, w_i)$

The discrete problem corresponds to a system of equations. Using the shape functions w_i to calculate the elements of the coefficient matrix $a(w_j, w_i)$ and the right hand side of $f_j = f(w_j)$ and solving the system of equations (5.41) leads to the unknown u_i. Evaluating eq.(5.37) gives the approximation of the problem $u_{_h} = \sum_{i=1}^{N} u_i w_i$. The practical realisation of the FEM depends strongly on the choice of the shape functions w_i. The basis functions must have some particular properties.

5.5.2   Shape or basis function

The definition of the shape or basis functions is not dependent on the single FEM. The same ideas used to construct the basis functions can be taken for the weighted residual methods, the Ritz methods, the different Galerkin approaches, the method using a variational equation and the energy minimum functional as well. The following ideas are of a general application.

If the approximation u_s from u must be very accurate, the number N of the basis functions must be very high. This results in a large system of equations. This is the reason why the basis functions must be chosen in such a way that the coefficient matrix A_h contains as much as possible of zero elements. The best possibility would be to choose the shape functions such that A_h is the unity-matrix. This is not practically possible.

In general a_{ij} is an integral over the domain Ω of summations of products of the w_i and w_j and their derivatives. If the shape functions are chosen such that they are only in a partial domain Ω_i of Ω non-zero and else equal to zero, the products $w_i \cdot w_j$ i, j=1(1)N are only for some combinations of i and j non-zero. This means for a FEM discretisation that for many i and j , Ω_i and Ω_j should not have common nodes. This results in the desired sparse system of equations.

The requirements to be fulfilled by the shape functions are:
- smoothness, piece-wise differentiable
- additional properties resulting from the boundary conditions must be satisfied
- the shape functions should be simple
- a good approximation u_h from u should be obtained.

5.5.2.1 Construction of the basis functions To construct a useful shape function, (Goering et al. [44]) the unit square (Fig. 5.27) is discretised into partial squares with the co-ordinates $x_v = vh$ and $y_\mu = \mu h$ with v, μ=1(1)M-1 and Mh=1 .
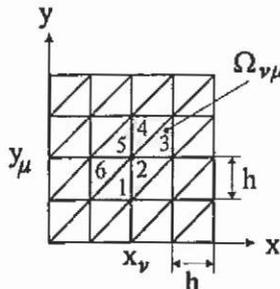
Fig. 5.27. Unit square with triangular FEM discretisation.

In this example, triangular-shaped elements are assumed, and so every partial square is represented by two triangular elements (Fig. 5.27). V_h is the $(M-1)^2 = N$-dimensional space characterised by a linear function in x and y inside each triangle.

The shape functions $w_\pi(x,y)$ in V_h are chosen in such a way that one node of the triangle has the function value 1 and the other two nodes a zero value (Fig. 5.28):

$$w_{\nu\mu}(x_k, y_l) = \begin{cases} 1 & \text{for} \quad k = \nu, \ l = \mu \\ 0 & \text{else} \end{cases} \tag{5.43}$$

Therefore, it is

$$w_\pi(x,y) \equiv 0, \quad \text{if}(x,y) \notin \Omega_\pi \quad . \tag{5.44}$$
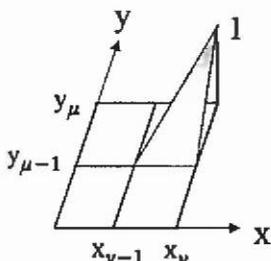
Fig. 5.28. Simple linear shape function $w_\pi(x,y)$.

$\Omega_{\nu\mu}$ is the desired partial domain in which $w_{\nu\mu}$ is non-zero (Fig. 5.28).

For example triangle 1 can be written:

$$w_{\nu\mu} = \begin{cases} 1, & x = x_\nu, & y = y_\mu \\ 0, & x = x_{\nu-1}, & y = y_{\mu-1} \\ 0, & x = x_\nu, & y = y_{\mu-1} \end{cases} \tag{5.45}$$

Now the shape functions from the triangles 1, 2, ..., 6 are considered together. If the linear function

$$w_\pi = d_0 + d_1 x + d_2 y \tag{5.46}$$

is applied, the requirements for the triangular regions (5.45) result in a system of equations.

$$1 = d_0 + d_1 x_\nu + d_2 y_\mu \quad ,$$
$$0 = d_0 + d_1 x_{\nu-1} + d_2 y_{\mu-1} \quad , \tag{5.47}$$
$$0 = d_0 + d_1 x_\nu + d_2 y_{\mu-1} \quad .$$

In terms of matrix representation:

$$\begin{pmatrix} 1 & x_\nu & y_\mu \\ 1 & x_{\nu-1} & y_{\mu-1} \\ 1 & x_\nu & y_{\mu-1} \end{pmatrix} \cdot \begin{pmatrix} d_0 \\ d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad . \tag{5.48}$$

The solution of this system of equations yields $d_1=0$, $d_2=1/h$ and $d_0=1-\mu$. The value of the shape function for all triangles can be given by:

$$w_{\nu\mu} = \begin{cases} 1+(\dfrac{y}{h}-\mu) & \text{:triangle 1} \\[2mm] 1-(\dfrac{x}{h}-\nu)+(\dfrac{y}{h}-\mu) & \text{:triangle 2} \\[2mm] 1-(\dfrac{x}{h}-\nu) & \text{:triangle 3} \\[2mm] 1-(\dfrac{y}{h}-\mu) & \text{:triangle 4} \\[2mm] 1+(\dfrac{x}{h}-\nu)-(\dfrac{y}{h}-\mu) & \text{:triangle 5} \\[2mm] 1+(\dfrac{x}{h}-\nu) & \text{:triangle 6} \quad . \end{cases} \tag{5.49}$$

Fig. 5.29 illustrates the properties of the shape function introduced. $w_{\nu\mu}$ is a simple linear shape function with the desired properties. It fulfils the conditions

$$\int_\Omega w_{\nu\mu} \cdot w_{kl} d\Omega = 0 \qquad \text{if} |\nu-k|>1 \quad \text{or} \quad |\mu-l|>1 \tag{5.50}$$

to generate a sparse overall coefficient matrix.
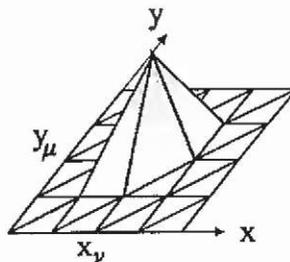
Fig. 5.29. Properties of the approximation applying linear shape functions.

Various basis functions of higher order such as quadratic, cubic etc. are possible and in common use, for example

linear:
$$w_{\nu\mu} = d_0 + d_1 x + d_2 y \qquad , \tag{5.51}$$

quadratic:
$$w_{\nu\mu} = d_0 + d_1 x + d_2 y + d_3 x^2 + d_4 y^2 + d_5 xy \qquad \text{etc.} \tag{5.52}$$

The degree of freedom (DOF) per finite element for a chosen basis function can be determined by (Kost [65]):

polynomial degree of p	DOF_{2D}	DOF_{3D}
	$\dfrac{(p+1)(p+2)}{2}$	$\dfrac{(p+1)(p+2)(p+3)}{6}$
1	3	4
2	6	10
3	10	20
4	15	35
5	21	56

Using polynomials of higher order generates a more accurate approximation of the exact solution. The DOF increases with rising order of the polynomial and this means that the computational expenses are increasing as well.

From the properties of the basis functions, general conclusions on the properties of the mesh discretising the domain Ω can be given. A mesh must consist of:

- non-overlapping elements
- nodes, corresponding to the nodes of an adjacent element (for node elements).

not corresponding node

a) b)
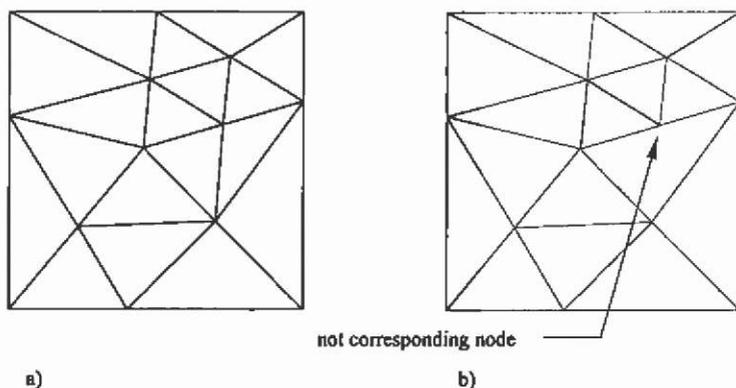
Fig. 5.30. a) Regular and b) not regular triangular element mesh.

5.5.3 Basic principle of the FEM

Three steps determine the basic principle of the FEM:

- Choose N shape functions w_i such that w_i is only in a partial region Ω_i from the domain Ω non-zero.

- Calculate the a_{ij} and f_j and solve the system of equations to obtain u_i .

$$\sum_{i=1}^{N} a_{ji} \cdot u_i = f_j \qquad\qquad j=1(1)N \quad .$$ (5.53)

- The approximated solution from $a(u,v) = f(v)$ is

$$u_h = \sum_{i=1}^{N} u_i w_i \quad .$$ (5.54)

5.5.4 Weighted residuals

With the exact solution u of a boundary problem it can be written

$$a(u,v) - f(v) = 0 \quad .$$ (5.55)

The solution of the problem obtained by an approximation $u_h = \sum_{i=1}^{N} u_i w_i$ is not exact and we must consider a residual R.

$$a(u_h,v_h) - f(v_h) = R \quad .$$ (5.56)

Assuming that the basis functions w_i satisfy the boundary conditions, in the example it was $v=0$ on Γ, the N unknown can be determined by choosing N points inside the domain Ω. For these points, the residuum is forced to be zero. Out of this, N equations to determine u_i are obtained. This method is called the point-collocation method (Binns et al. [13]).

A better approximation is obtained by averaging with an arbitrary function ω_i over the domain of interest Ω.

$$\int_{\Omega} (a(u_h,\omega_i) - f(\omega_i)) \cdot d\Omega = \int_{\Omega} R\omega_i \cdot d\Omega = 0 \quad .$$ (5.57)

Transferred to the example of a Poisson equation $\Delta u = -q$, this method yields:

$$\int_{\Omega} (\Delta u_h + q)\omega_i \cdot d\Omega = \int_{\Omega} R\omega_i \cdot d\Omega = 0 \quad .$$ (5.58)

The method is called the method of weighted residuals.

Various weighting functions are in common use. Using, for example, the Dirac-Delta function for ω_i results in the point-collocation method as a special case of the method of the weighted residuals. Usually linear weighting functions are chosen. They are easier to implement when compared to higher order functions and already deliver a sufficiently accurate approximation.

By choosing the shape functions, introduced in the last section, to be the weighting functions,

$$w_i = \omega_i \quad ,$$ (5.59)

we obtain the local Galerkin method. The combination of weighted residuals and Galerkin method is universal applicable.

5.5.4.1 Continuity Particular attention must be paid to the choice of basis function where the second derivative is present in the differential equation. This might cause difficulties with the integral at element boundaries. To avoid singular integrands, a continuous function must be chosen with continuous first derivative. This leads to the requirement of a defined continuity for the form functions.

If only the continuity of the linear form function is required, we have a C^0-continuity.

$$\frac{\partial' u(x, y, z)}{\partial n} \tag{5.60}$$

is continuous on the element's boundaries.

For a function of second order, the derivative at the element's boundary is continuous (Fig. 5.31). If in general C^s-continuity is required it can be written:

$$\frac{\partial' u(x, y, z)}{\partial n}. \tag{5.61}$$
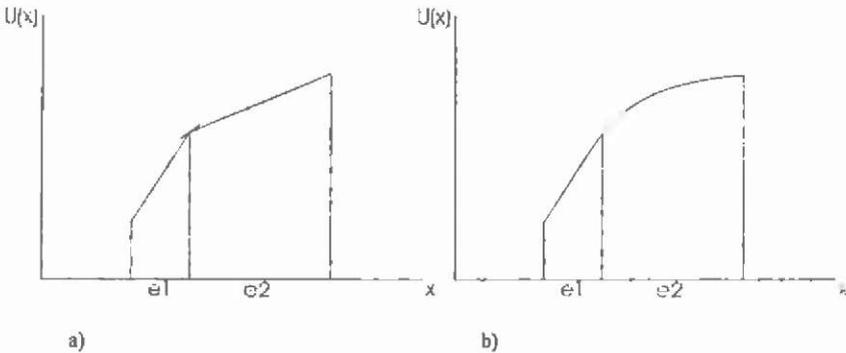
Fig. 5.31. a) C^0- and b) C^1 continuity for the solution $u(x)$ considered at the boundary from finite element e1 and e2.

5.5.4.2 Green's formula We have seen so far that we would need a C^1 continuity for the basis functions to solve the example of a Poisson equation. It is desirable to have simple linear basis functions with C^0 continuity. Applying the first Green's formula:

$$\int_\Omega P\Delta Q \cdot d\Omega = \int_\Gamma P\frac{\partial}{\partial n}Q \cdot d\Gamma - \int_\Omega \nabla P\nabla Q \cdot d\Omega \tag{5.62}$$

removes the second derivative in the equation and yields for:

$$\int_{\Omega} \omega_i \Delta u_k \cdot d\Omega = \int_{\Gamma} \omega_i \frac{\partial}{\partial n} u_k \cdot d\Gamma - \int_{\Omega} \nabla \omega_i \nabla u_k \cdot d\Omega \qquad . \tag{5.63}$$

This leads to the weak form of the Poisson equation. Weak form means that weak requirements concerning the continuity are demanded on their solution when compared to the solution of the previous differential equation. Therefore, the form of the previous differential equation is called strong.

$$\int_{\Omega} (\Delta u_k + q) \omega_i \cdot d\Omega = \int_{\Omega} \nabla \omega_i \nabla u_k \cdot d\Omega - \int_{\Omega} \omega_i q \cdot d\Omega - \int_{\Gamma} \omega_i \frac{\partial}{\partial n} u_k \cdot d\Gamma = 0 \quad . \tag{5.64}$$

As desired, a C^0 continuity of the basis function is now sufficient to solve the problem. A little disadvantage of this approach is the fact that now the weighting function ω_i must have C^0 continuity as well. Constant weighting functions are not possible with this approach. The use of the first Green's formula reduces the continuity requirements of the basis functions but increases efforts concerning the weighting functions.

It is obvious from this approach, that the solution of the strong form is always a solution of the weak form. The other way around, this does not always hold.

5.5.5   Energy-minimum functional

The principle of minimum energy requires that the potential distribution corresponds to the minimum of stored field energy. For several electrotechnical problems, this equivalent minimisation problem is known.

$$F(u_k) = \frac{1}{2} \int_{\Omega} a(u_k, v_k)' \cdot d\Omega - \int_{\Omega} f(v_k) \cdot d\Omega \rightarrow \min . \tag{5.65}$$

The quantity $F(u_k)$ is the total energy of the function $u_k(x)$, which is the sum of the internal energy $\frac{1}{2} \int_{\Omega} a(u_k, v_k)' \cdot d\Omega$ and the load potential $- \int_{\Omega} f(v_k) \cdot d\Omega$. The minimum energy functional yields the same results obtained by the local Galerkin method (Kost[65], Eriksson et al.[38]).

5.5.6   Types of elements

The domain of interest can be discretised by various types of finite elements. In this section the most commonly used element types for two- and three-dimensional FEM meshes will be introduced. It is the aim to use simple geometries for the elements. Cross-sectional elements such as triangles, quadrangles and rectangles are used for two-dimensional models and volume elements such as tetrahedrons and cuboids for the three-dimensional FEM models.

Table 5.3. Standard FEM element types.

2D	3D

The outer boundaries of the geometries are mainly approximated by a polygon. In general triangular or tetrahedral shapes can best approximate such geometries. A complicated geometry can be approximated by a large number of such simple-shaped elements.

The FEM model can be built up with element types with different properties. Nodal and edge elements can be distinguished. The most common types of elements will be briefly introduced in the next section. For further details on special types such as non-conform elements, please refer to the literature (Goering et al. [44], Eriksson et al. [38], Kost [65], Binns et al. [13]). Line elements are not considered here.

5.5.6.1 Nodal elements The triangular nodal element is the most commonly used element type for two-dimensional problem formulations. This element shape is the most adaptable to complicated geometries. Therefore, it has advantages concerning an adaptive local mesh refinement to enhance the quality of the approximated solution.
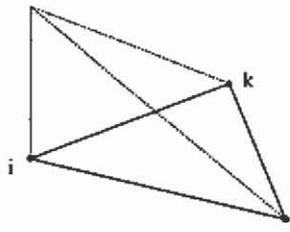
Fig. 5.32. 2D triangular nodal element with linear shape function.

The unknown values of the approximated function are defined at the nodes of element. By using first order basis functions, w_i equals 1 at one node of the domain and 0 at the other two nodes of the triangular sub-domain Ω_i (Fig. 5.32).

$$w_i^{2D} = d_0 + d_1 x + d_2 y \qquad . \qquad (5.66)$$

The equivalent element type for three-dimensional models is the nodal tetrahedron with the same advantage of being able to adapt complicated geometries very accurately.

$$w_i^{3D} = d_0 + d_1 x + d_2 y + d_3 z \qquad . \qquad (5.67)$$

5.5.6.2 Edge elements By using edge elements, the unknowns are referred to the edges of this element type. This is advantageous in three-dimensional problem definitions. Therefore, this type of element is in common use for three-dimensional FEM problems.

The basis functions w_i^e are defined by:

$$w_i^e = w_a \nabla w_b - w_b \nabla w_a \qquad . \qquad (5.68)$$

For a three-dimensional element the course of the basis function of an edge element is plotted in Fig. 5.33.
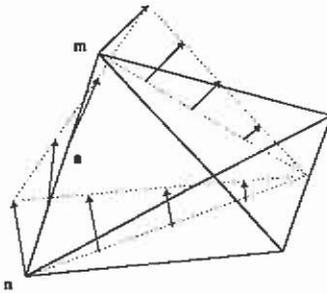
Fig. 5.33. First order three-dimensional edge element.

5.5.6.3 Facet elements The basis functions w_i^f are defined by:

$$w_i^f = 2\left(w_i \nabla w_a \times \nabla w_b + w_a \nabla w_b \times \nabla w_i + w_b \nabla w_i \times \nabla w_a\right) \qquad . \qquad (5.69)$$
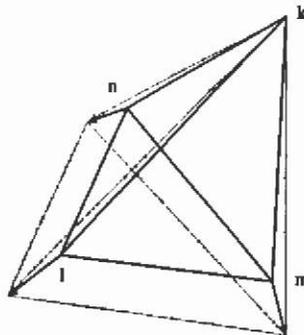
Fig. 5.34. First order three-dimensional facet element.

Fig. 5.34 shows the course of a linear basis function of a three-dimensional facet element. The function value in one node is zero. The course of the basis function value is linear between the nodes of the tetrahedron.

5.5.6.4 FEM element properties The basis functions w_i^{1D}, w_i^{1D}, w_i^e and w_i^f own special properties (Table 5.4).

Table 5.4. Properties of finite elements.g

type	element properties
w_i^{1D}, w_i^{1D}	the value of the basis function at a node is 1 at the node n and 0 at all other nodes
w_i^e	the value of the line integral over an edge is 1 at the edge e and 0 at all other edges
w_i^f	the value of the surface integral over a facet is 1 in the plane f and 0 at all other planes

5.6 Material modelling

In this section the most important material models such as permanent magnet and non-linear ferromagnetic materials are introduced.

5.6.1 Non-linear material

Most of the electromagnetic field problems are inherently non-linear. Having accurate numerical techniques able to handle non-linearities strengthens the trend to have problems out of this class because material costs and a desired minimisation of the devices force the exploitation of material to its limits. As a consequence mainly highly saturated magnetic circuits have to be analysed.
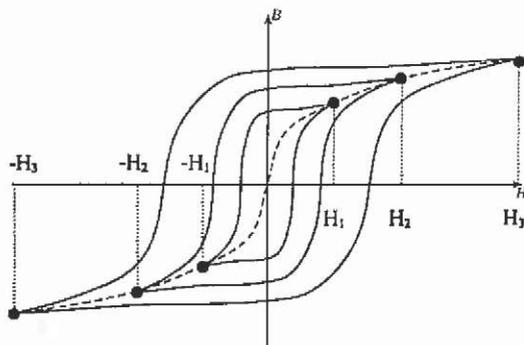
Fig. 5.35. Hysteresis loops and magnetisation characteristic (Salon [99]).

The analysis of magnetic devices requires knowledge of the physical properties of the materials used. Slowly magnetising a piece of ferromagnetic material to a value H_1 and then reducing the field until it reaches $-H_1$, and repeating the process until the characteristic remains the same, yields a hysteresis loop (Fig. 5.35). If the field is increased to a value H_2 and the process is repeated, and then to H_3, etc., a family of nested hysteresis loops is obtained as shown in (Fig. 5.35). The connected tips of the hysteresis characteristics represent the normal magnetisation characteristic. This curve is most commonly used in the finite element analysis to represent the non-linear properties of ferromagnetic materials. Using soft magnetic materials, the curve is narrow and thus an acceptable approximation of the real behaviour of the material. Here, only non-hysteretic materials are discussed. Manufacturers having different grades normally deliver the material characteristics by B-H curves (Fig. 5.36).
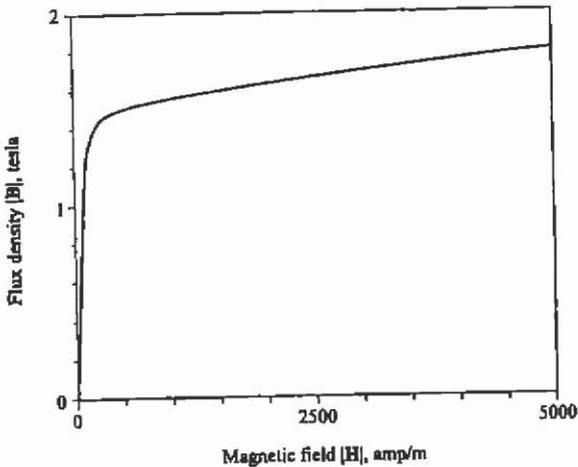
Fig. 5.36. Magnetisation characteristic in BH-form.

The basic principles of the linear finite element analysis carry over to non-linear problems almost without modification. As always, a stationary functional is constructed and discretised over finite elements. As might be expected, the equations resulting from non-linear problems are non-linear as well. They can be solved by several different methods. Simple iterative methods are not always stable and can take a long time to converge. A more common approach is to use Newton iterations.

In order to extend the linear finite element procedure to include non-linear material properties, a mathematical model describing the magnetic properties of the material is recommended. Therefore, computer readable files of material properties must be maintained ready for use. Numerous ways of modelling magnetic property curves have been tried out:

- reluctivity as a function of flux density squared, $\upsilon = \upsilon(B^2)$
- field as a function of flux density, $H = H(B)$
- permeability as a function of field squared, $\mu = \mu(H^2)$
- permeability as a function of flux density, $\mu = \mu(B)$.

The Newton method is the reason for using squared values of the independent variables B or H, rather than magnitudes. The variables are usually derived from potentials in vector component form, so that finding the magnitude involves first finding the squares of the components and then extracting the square root of their sum. Using the Newton method the Jacobian matrix can be evaluated in the form:

$$\frac{\partial^2 F}{\partial U_i \partial U_j} = \int \frac{\partial^2 W}{\partial U_i \partial U_j} d\Omega = \int (\frac{\upsilon}{2} \frac{\partial^2(B^2)}{\partial U_i \partial U_j} + \frac{1}{2} \frac{d\upsilon}{dB^2} \frac{\partial B^2}{\partial U_i} \frac{\partial B^2}{\partial U_j}) d\Omega \quad . \quad (5.70)$$

Therefore, the classical B-H characteristic is not the best choice to introduce ferromagnetic material properties in finite element software. In the most common program designs the $\upsilon = \upsilon(B^2)$ representation is employed.
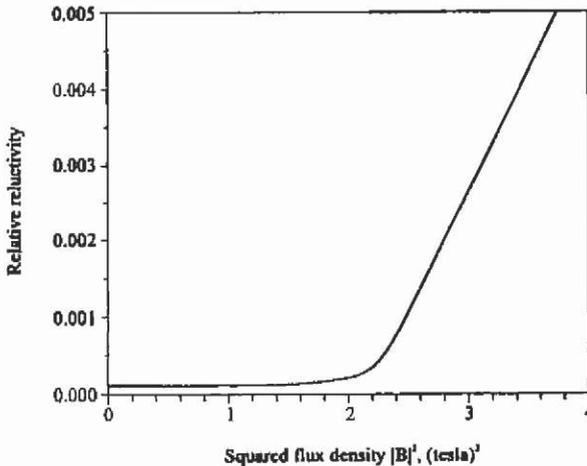
Fig. 5.37. Magnetisation characteristic in υB^2 form.

The range $0 < B^2 < B_{max}^2$ of the $\upsilon = \upsilon(B^2)$ material characteristic is subdivided into segments. The data samples υ, B^2 are tabulated in an ASCII file readable by the finite element program. For the points between the given data an interpolation scheme has to be chosen. In the simplest case a linear approximation could be chosen. The ferromagnetic characteristic would then be represented by a polygon. This approximation would be discontinuous in the points of the given data

sample. However, the Newton iteration recommends a continuous function. Therefore, the values between the single data sample are usually approximated by cubic spline-interpolating polynomials. For all points beyond the range of available data the curve can be extrapolated linearly.

5.6.1.1 Cubic spline interpolation Splines are curves to approximate functions. Cubic splines in particular have received much attention in numerical analysis in the past, replacing other polynomial and exponential approximations. The main reason for the interest in splines is that they result in a simple formulation. They interpolate exactly at the given data points and have a continuous first and second order derivative. The continuous first derivative makes the method suitable for the Newton iteration.

Assume a function $x=x(y)$ to be interpolated by cubic splines in the interval $[a, b]$.

Given a tabulated function $y_i=y(x_i)$, $i=1(1)n$ with x arranged in monotone way.

$$a = x_1 < x_2 < ... < x_n = b \quad . \tag{5.71}$$

The spline interpolating polynomial is called $S(x)$ with properties:

- $S(x)$ is two times continuous by differentiable in $[a, b]$.
- $S(x)$ is in every interval $[x_i, x_{i+1}]$ given by a cubic polynomial.
- The points of $S(x_i)$ are $y_i(x_i)$.
- As a boundary condition it is taken that $y_1'' = y_n'' = 0$ in order to obtain a so called natural spline.

Using the approach

$$S(x) = y = a_i y_i + b_i y_{i+1} + c_i y_i'' + d_i y_{i+1}'' \quad , \tag{5.72}$$

the coefficients can be calculated by applying:

$$a_i = \frac{x_{i+1} - x}{x_{i+1} - x_i} \quad ,$$

$$b_i = \frac{x - x_i}{x_{i+1} - x_i} \quad ,$$

$$c_i = \frac{1}{6}(a_i^3 - a_i)(x_{i+1} - x_i)^2 \quad , \tag{5.73}$$

$$d_i = \frac{1}{6}(b_i^3 - b_i)(x_{i+1} - x_i)^2 \quad .$$

The first derivative with respect to x of the interpolating function is:

$$\frac{dS(x)}{dx} = \frac{dy}{dx} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} - \frac{3a_i^2 - 1}{6}(x_{i+1} - x_i)y_i'' + \frac{3b_i^2 - 1}{6}(x_{i+1} - x_i)y_{i+1}'' \tag{5.74}$$

while the second is:

$$\frac{d^2 S(x)}{dx^2} = \frac{d^2 y}{dx^2} = a_i y_i'' + b_i y_{i+1}'' \quad . \tag{5.75}$$

With this, the requirement of continuity of the second derivative over the boundaries of the interval $[x_{i-1}, x_i]$ and $[x_i, x_{i+1}]$ is satisfied. Because of a required continuous derivative of first order of the interpolating spline function, the values of dS/dx at the point $x = x_i$ for $x \in [x_{i-1}, x_i]$ and $x \in [x_i, x_{i+1}]$ must be equal. Employing $x = x_i$ for both intervals yields:

$$\frac{x_i - x_{i-1}}{6} y_{i-1}'' + \frac{x_{i+1} - x_{i-1}}{3} y_i'' + \frac{x_{i+1} - x_i}{6} y_{i+1}'' = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} - \frac{y_i - y_{i-1}}{x_i - x_{i-1}} \quad . \tag{5.76}$$

This equation can be evaluated with $i=2(1)n-1$ for every interval. This results in n-2 linear independent equations for the n unknown y_i'', $i=1(1)n$ at the given data samples. Using the conditions of a natural spline $y_1'' = y_n'' = 0$ yields two additional constraints. This is a symmetric tri-diagonal system of equations and is easy to solve. With the derivatives known, the coefficients of the interpolating spline function are now determined.

A great advantage of using cubic splines is the fact that this linear system of equations has to be solved only once to obtain the values of the second derivative. Therefore, space for the solution of dimension n has to be allocated in the memory of the computer only. For the finite element method this means that this system of equation has to be solved only once independently of the number of finite elements used in the model of the magnetic circuit. For the practical use of the method a field of 20-25 data samples is sufficient to represent a non-linear ferromagnetic material characteristic.

The non-linear material characteristics are normally delivered by the manufacturer in a B=B(H) form. The data can be given in a tabular or a graphical form. This implies that the user of finite element software takes the values and rearranges them into the appropriate $\upsilon = \upsilon(B^2)$ representation. Converting the given data sample in this way may lead to numerical difficulties. To ensure stable convergence and the highest possible computational speed of the Newton iteration, particular attention has to be paid to the numerical representation of the data samples. The curve representation has to fulfil some numerical requirements. The function $\upsilon = \upsilon(B^2)$ has to be monotonic. If the characteristic is not monotonic, the derivative changes sign, eventually leading to slow or even to non-convergence of the iteration scheme. A possible way to overcome this problem lies in the optimisation of the given data samples. Here, a numerical optimisation algorithm can change the values of the

given points to ensure a monotonic behaviour of the characteristic. After optimisation a technically pure curve is constructed.

5.6.2   Permanent magnets

The development of high energy permanent magnet materials such as SmCo and NdFeB grades has led to increased interest in the use of permanent magnet material in electrical machines and actuators. The representation of these hard magnetic materials is difficult and the subject of ongoing research. As mentioned in the last section, ferromagnetic materials are characterised by a narrow hysteresis loop. In contrast, hard magnetic materials such as permanent magnets exhibit wide loops. It is often acceptable to consider the magnetic characteristic of a permanent magnet by a straight line in the second quadrant of the hysteresis loop. This is not a limitation of the finite element method. During the design of permanent magnet excited devices, particular attention must be paid to the operating temperature of the magnets.

The intersection of the hysteresis loop with the ordinate is called the residual or remanence flux density B_R. The intersection of the abscissa and the loop is called the coercive force H_C.

There are two possibilities allowing the modelling of a permanent magnet material:

- magnetisation model
- current sheet approach.

Although these two methods have a different starting point, they both result in the same set of equations. Assuming a straight line as the characteristic of the permanent magnet material (Fig. 5.38), there are only two parameters required to define the characteristic:

- the slope of the line $\mu_{_m}$ and
- the y-axis intercept B_R.
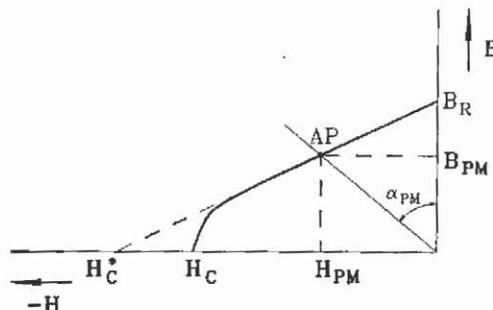
Fig. 5.38. Definition of permanent magnet material.

In Table 5.5 the most important permanent magnet properties are collected. During a design of a device excited by permanent magnets, particular attention has to be paid to the temperature dependence of the grade used.
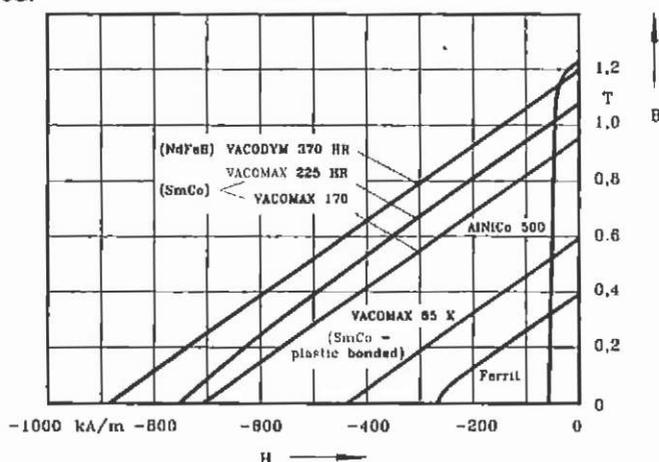
Fig. 5.39. Demagnetisation characteristics for different permanent magnet material at room temperature 20° C.

Table 5.5. Properties of different permanent magnet material at room temperature.

		SmCo VACOMAX 255 HR	NdFeB Alpha Magnet MQ1	NdFeB VACODYM 400 HR	AlNiCo KOERZIT 700	Ferrite KOEROX 400
B_R	T	1.00 ... 1.15	0.63	1.05 ... 1.15	1.35	0.4
H_c	kA/m	600 ... 900	380	770 ... 900	58	255
$(BH)_{max}$	kJ/m³	190 ... 240	75	210 ... 265	62.1	31
μ_m	-	< 1.1	1.15	< 1.1	1.5 ... 3.0	1.1
$RTC(B_R)$	%/K	-0.03 ... -0.035	-0.15	-0.12	-0.02	-0.2
$RTC(H_c)$	%/K	-0.2 ... -0.3	-0.5 ... -0.65	-0.5 ... -0.65	-0.03 ... -0.07	+0.3 ... +0.5
T_c	° C	800	310	310	790 ... 900	450
T_{max}	° C	300	180	125	500	325
H_{max}	kA/m	3500	2800	2500	280	1000
ρ	g/cm³	8.4	6.0	7.4	7.3	4.9

Typical demagnetisation characteristics of different grades are shown in Fig. 5.39.

5.6.2.1 Magnetic vector model The demagnetisation characteristic is defined by

$$B = \mu_0 \{(1 + \chi_m)\mathbf{H} + \mathbf{M}\} \tag{5.77}$$

where χ_m is the magnetic susceptibility, $\mathbf{M}$ the magnetisation vector and $\mathbf{H}$ the field strength at the operating point AP. In terms of the remanent flux density

$$\mathbf{B}_R = \mu_0 \mathbf{M} \quad . \tag{5.78}$$

The incremental permeability, the slope of the demagnetisation characteristic, is

$$\frac{\partial |\mathbf{B}|}{\partial |\mathbf{H}|} = \mu_0 (1 + \chi_m) \quad . \tag{5.79}$$

χ_m is a very small positive number so that the apparent permeability of the magnet is only slightly larger than that of the free space (Table 5.5).

The reluctivity is defined as

$$v = \frac{1}{\mu_0 (1 + \chi_m)} \tag{5.80}$$

applying to the demagnetisation characteristic, yields

$$\mathbf{H} = v(\mathbf{B} - \mu_0 \mathbf{M}) \quad . \tag{5.81}$$

Using the Maxwell equation for a magneto static problem

$$\nabla \times \mathbf{H} = \mathbf{J} \tag{5.82}$$

yields

$$\nabla \times (v\mathbf{B}) = \mathbf{J} + \nabla \times (v\mu_0 \mathbf{M}) \quad . \tag{5.83}$$

The second term, the magnetic vector, on the right-hand side represents a source term and can be identified as an equivalent magnetic current.

5.6.2.2 Current sheet approach Using an equivalent current sheet representing the permanent magnet material is an easy way to introduce the material properties in a finite element program. In its original form it is not easy to apply permanent magnets with an odd shape. However, if the model is extended, an arbitrary shaped magnet can be described. In the following section a linear demagnetisation characteristic of the material is assumed.

a)   b)
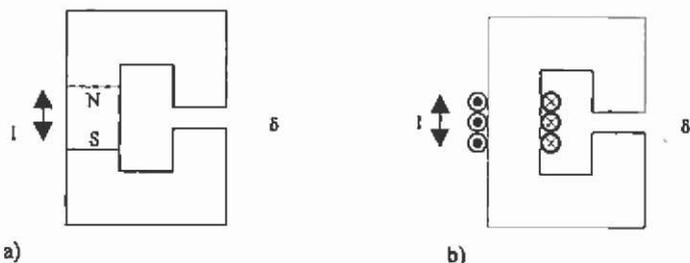
Fig. 5.40. Idealised magnetic core excited a) by a permanent magnet and b) by a current sheet.

Taking the permeability of the iron core in Fig. 5.40 to be infinite, Ampere's law yields

$$H_m l + H_s \delta = 0 \quad . \tag{5.84}$$

Now, for a uniform B, it can be written (H is negative)

$$B = -\mu_0 \frac{l}{\delta} H_m \quad . \tag{5.85}$$

The intersection between the air gap characteristic, the load line, and the demagnetisation curve represents the point AP of the magnet material used (Fig. 5.38) with

$$B = B_R + \frac{B_R}{H_C} H = B_R + \mu_0 (1 + \chi_m) H = B_R + \mu H \quad . \tag{5.86}$$

The permanent magnet in Fig. 5.38 can now be represented as a current sheet with the total ampere-turns $NI = H_c \ell$ and a material of equivalent permeability $\mu = \dfrac{B_R}{H_c}$ (Fig. 5.38).

Again assuming an infinite permeability of the iron parts of the magnetic core,

$$H_m l + H_s = H_c \ell \quad . \tag{5.87}$$

This yields

$$H_c = \frac{H_c}{B_R} B + \nu_0 \frac{\delta}{l} B \quad . \tag{5.88}$$

All magnetic quantities outside the magnet remain the same as in the case of the magnetic vector, but are shifted to the first quadrant of the magnetisation characteristic.

This method is easy to implement for rectangular magnets with a magnetisation parallel to two sides of the rectangle (Fig. 5.41).
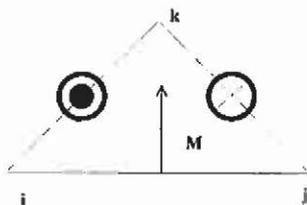
Fig. 5.41. Triangular finite element with magnetisation.

These ideas now can be transferred to permanent magnets with an arbitrary shape. Therefore, current sheets are assumed on all sides of the finite element. After some elementary trigonometrically manipulations it can be written

$$I_i = \frac{1}{2} H_c (c_i \cos\theta_n - b_i \sin\theta_n) \quad,$$

$$I_j = \frac{1}{2} H_c (c_j \cos\theta_n - b_j \sin\theta_n) \quad, \tag{5.89}$$

$$I_k = \frac{1}{2} H_c (c_k \cos\theta_n - b_k \sin\theta_n) \quad,$$

with

$$c_i = x_k - x_j \quad,$$

$$b_i = y_k - y_j \quad. \tag{5.90}$$

In a similar way the c_j, b_j, c_k and b_k are calculated. With the equivalent magnetisation vector

$$M = -v_R H_c \tag{5.91}$$

the currents are

$$I_i = \frac{1}{2} v \mu_0 (M_x c_i - M_y b_i) \quad. \tag{5.92}$$

The other edge currents can be calculated in a analogue way I_i. This procedure applied on an element by element base enables the construction of arbitrary shaped permanent magnet material.

5.7   Numerical implementation of the FEM

In this section, the theoretically derived abstract knowledge of the finite element method will be practically applied to magnetostatic field problems. Standard linear triangular finite elements will be used. The system of equations will be derived using the energy minimum functional.

Starting with the numerical solution of the Laplace equation in two dimensions, then introducing impressed currents and permanent magnet material, the difficulties programming a FEM code will be introduced and treated. Reading this section should enable the reader to obtain an understanding of the practical realization of the method in a computer program.

5.7.1   Laplace's equation

For simplicity in this example of the implementation of the FEM, a two-dimensional approach is considered. A Cartesian co-ordinate system is assumed. In a two-dimensional field the electrical field strength consists of one component in the z-direction only and the magnetic fields are in the xy-plane.

$$\mathbf{E} = (0, 0, E_z)$$
$$\mathbf{H} = (H_x, H_y, 0) \tag{5.93}$$
$$\mathbf{B} = (B_x, B_y, 0) \quad .$$

The vector potential **A** and the current density **J** have a component only in the z-direction. For reasons of simplicity the index z, to indicate the direction, will not be written in the following.

$$\mathbf{A} = (0, 0, A_z) = A$$
$$\mathbf{J}_0 = (0, 0, J_{0z}) = J_0 \quad . \tag{5.94}$$

The static magnetic field in terms of the vector potential is given by the A-formulation of Laplace's equation:

$$\nabla^2 A = 0 \quad . \tag{5.95}$$

At this moment air is considered as material. Using the energy minimum functional

$$F(A) = \frac{1}{2} \int_\Omega |\nabla A|^2 \, d\Omega \quad . \tag{5.96}$$

For a complete problem definition a magnetic flux has to be imposed by applying the appropriate Dirichlet conditions at the boundaries of the field domain of interest.

5.7.1.1   Linear basis function The problem region is discretised by triangular node elements. The vector potential is approximated by linear shape functions. By connecting all triangular elements at their nodes and forcing the node potential to be equal, the magnetic vector potential becomes continuous over the defined field region. Linear shape functions are applied by using:

$$A = a + bx + cy \quad . \tag{5.97}$$

The coefficients a, b and c are found from the values of the magnetic vector potential A_1, A_2 and A_3 at the three nodes of an element (Fig. 5.42).

$$\begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad . \tag{5.98}$$
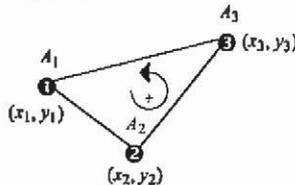
Fig. 5.42. Triangular finite element.

The magnetic vector potential within an element is:

$$A = \begin{bmatrix} 1 & x & y \end{bmatrix} \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}^{-1} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} \qquad (5.99)$$

Using the shape functions, the magnetic vector potential is approximated by:

$$A = \sum_{i=1}^{3} A_i N_i(x, y) \qquad (5.100)$$

with

$$N_1 = \frac{1}{2\Delta_e}[(x_2 y_3 - x_3 y_2) + (y_2 - y_3)x + (x_3 - x_2)y] \quad , \qquad (5.101)$$

$$N_1 = \frac{1}{2\Delta_e}(a_1 + b_1 x + c_1 y) \quad . \qquad (5.102)$$

The shape functions N_2 en N_3 are found by cyclic permutation of the indices in (5.101) and (5.102). The area Δ_e of an element is:

$$\Delta_e = \frac{1}{2}[(x_2 y_3 - x_3 y_2) + (y_2 - y_3)x_1 + (x_3 - x_2)y_1] = \frac{1}{2}\det\begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix} . \qquad (5.103)$$

It is easy to verify that the three shape functions are one at one node and zero at both other nodes.

$$N_i(x_j, y_j) = \begin{cases} 0 & \forall \quad i \neq j \\ 1 & \forall \quad i = j \end{cases} \quad . \qquad (5.104)$$

5.7.1.2   Functional within an element The gradient of the magnetic vector potential in terms of the shape functions can be calculated by:

$$\nabla A = \sum_{i=1}^{3} A_i \nabla N_i \quad . \qquad (5.105)$$

Substituting (5.105) into the functional within an element gives

$$F^{(e)} = \frac{1}{2}\iint_{\Delta_e} |\nabla A|^2 \, d\Omega \quad , \qquad (5.106)$$

$$F^{(e)} = \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3} A_i \cdot \underbrace{\int_{\Delta_e} \nabla N_i \cdot \nabla N_j \, d\Omega}_{K_{ij}^{(e)}} \cdot A_j \quad . \qquad (5.107)$$

The elements of the 3 x 3 element matrix $\mathbf{K}^{(e)}$ are

$$K_{ij}^{(e)} = \frac{1}{4\Delta_e}(b_i b_j + c_i c_j) \quad . \qquad (5.108)$$

In a matrix-vector notation, the functional within an element is:

$$F^{(e)} = \frac{1}{2} \mathbf{A}^T \underbrace{\mathbf{K}^{(e)}}_{\text{'stamp'}} \mathbf{A} \qquad . \tag{5.109}$$

5.7.1.3   Assembling all elements The overall functional is found as the sum of the functionals within the elements.   ·

$$F = \sum_{\forall e} F^{(e)} \qquad , \tag{5.110}$$

$$F = \frac{1}{2} \mathbf{A}^T \mathbf{K} \mathbf{A} \qquad . \tag{5.111}$$

Summation is practically done element by element. The element matrix $\mathbf{K}^{(e)}$ is calculated for each finite element and added to the already existing coefficient matrix K. For the two finite elements in Fig. 5.43 this yields:

$$\mathbf{K} = \begin{bmatrix} K_{11}^{(1)} & K_{12}^{(1)} & K_{13}^{(1)} & 0 \\ K_{21}^{(1)} & K_{22}^{(1)} + K_{11}^{(2)} & K_{23}^{(1)} + K_{13}^{(2)} & K_{12}^{(2)} \\ K_{31}^{(1)} & K_{32}^{(1)} + K_{31}^{(2)} & K_{33}^{(1)} + K_{33}^{(2)} & K_{32}^{(2)} \\ 0 & K_{21}^{(2)} & K_{23}^{(2)} & K_{22}^{(2)} \end{bmatrix} \qquad . \tag{5.112}$$

Two elements combined with each other have four nodes, i.e. the coefficient matrix K becomes a 4 x 4 matrix.
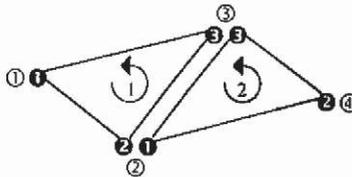
Fig. 5.43. Joining two finite elements.

5.7.1.4   The system of equation The system of linear equations is found by forcing the partial derivatives with respect to all unknowns to be zero. This is in fact an energy minimisation.

$$\frac{\partial F}{\partial A_k} = 0 \quad \forall k \qquad . \tag{5.113}$$

Applying Dirichlet conditions at boundaries is fixing the magnetic vector potential at some prescribed values. Two groups of nodes can be distinguished: nodes with free potentials that are to vary (index f) and nodes with prescribed potential values (index p). The node numbering is such that all nodes with potentials that are free to vary are numbered first.

$$F = \frac{1}{2}\begin{bmatrix} \mathbf{A}_f^T & \mathbf{A}_p^T \end{bmatrix} \begin{bmatrix} \mathbf{K}_{ff} & \mathbf{K}_{fp} \\ \mathbf{K}_{pf} & \mathbf{K}_{pp} \end{bmatrix} \begin{bmatrix} \mathbf{A}_f \\ \mathbf{A}_p \end{bmatrix} \quad , \tag{5.114}$$

$$\frac{\partial}{\partial \begin{bmatrix} \mathbf{A}_f \end{bmatrix}} \begin{bmatrix} \mathbf{A}_f^T & \mathbf{A}_p^T \end{bmatrix} \begin{bmatrix} \mathbf{K}_{ff} & \mathbf{K}_{fp} \\ \mathbf{K}_{pf} & \mathbf{K}_{pp} \end{bmatrix} \begin{bmatrix} \mathbf{A}_f \\ \mathbf{A}_p \end{bmatrix} = 0 \quad , \tag{5.115}$$

$$\begin{bmatrix} \mathbf{K}_{ff} & \mathbf{K}_{fp} \end{bmatrix} \begin{bmatrix} \mathbf{A}_f \\ \mathbf{A}_p \end{bmatrix} = 0 \quad . \tag{5.116}$$

Applying the boundary conditions, the system of linear equations becomes:

$$\begin{bmatrix} \mathbf{K}_{ff} & 0 \\ 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A}_f \\ \mathbf{A}_p \end{bmatrix} = \begin{bmatrix} -\mathbf{K}_{fp}\mathbf{A}_p \\ \mathbf{A}_p \end{bmatrix} \quad . \tag{5.117}$$

The generated coefficient matrix owns some interesting properties:

- Sparse, K_{ij} is different from zero only if node i and j are connected by an element.
- The matrix is symmetric $K_{ij} = K_{ji}$, diagonal dominant and positive definite.

The assembly of the matrix is straightforward and is done element by element (Fig. 5.44).
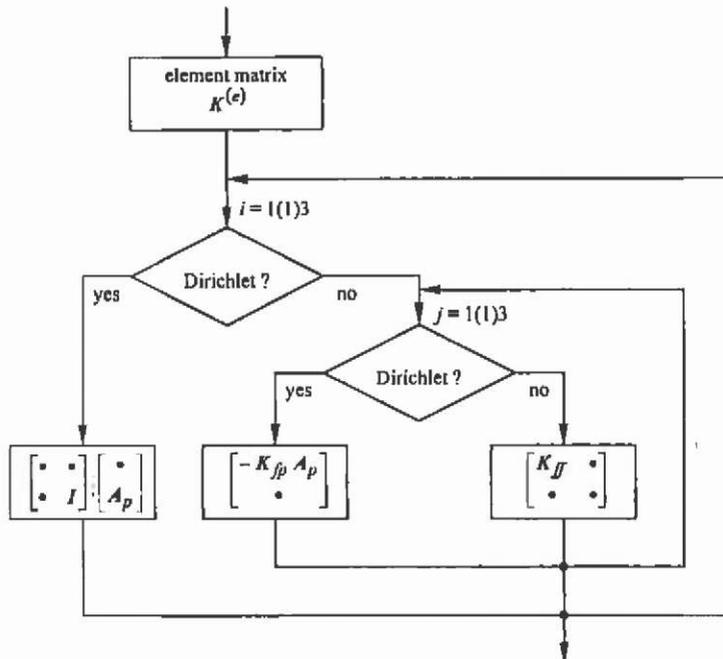
Fig. 5.44. Assembling of the coefficient matrix and the right hand side.

5.7.1.5 Material properties Laplace's equation in two dimensions is given by

$$\nabla \cdot (\nu \nabla A) = 0 \qquad . \tag{5.118}$$

With the reluctivity, respectively permeability and their relative values, index r, $\nu = \nu_r \cdot \nu_0 = \dfrac{1}{\mu} = \dfrac{1}{\mu_r \cdot \mu_0}$, characterizing the ferromagnetic properties of the material used. Due to numerical reasons the relative reluctivity ν_r is used:

$$\nabla \cdot (\nu_r \nabla A) = 0 \qquad . \tag{5.119}$$

The functional is:

$$F(A) = \frac{1}{2} \int_\Omega \nu_r |\nabla A|^2 \, d\Omega \qquad . \tag{5.120}$$

and is valid for magnetostatic problems with different materials and where only Dirichlet and Neumann conditions are applied.

5.7.1.5.1 Functional within an element The functional within an element becomes:

$$F^{(e)} = \frac{1}{2} \int_{A_e} \nu_r^{(e)} |\nabla A|^2 \, d\Omega \qquad , \tag{5.121}$$

$$F^{(e)} = \frac{1}{2} \sum_{i=1}^{3} \sum_{j=1}^{3} A_i \cdot \underbrace{\nu_r^{(e)} \int_{A_e} \nabla N_i \cdot \nabla N_j \, d\Omega}_{K_{ij}^{(e)}} \cdot A_j \qquad . \tag{5.122}$$

The elements of the element matrix $\mathbf{K}^{(e)}$ are given by

$$K_{ij}^{(e)} = \frac{\nu_r^{(e)}}{4\Delta_e} (b_i b_j + c_i c_j) \qquad . \tag{5.123}$$

5.7.2 Poisson's equation

Poisson's equation in two dimensions is given by:

$$\nabla \cdot (\nu \nabla A) = -J_0 \tag{5.124}$$

or due to numerical reasons

$$\nabla \cdot (\nu_r \nabla A) = -\frac{J_0}{\nu_0} \qquad . \tag{5.125}$$

The functional is

$$F(A) = \frac{1}{2} \int_\Omega \left(\nu_r |\nabla A|^2 - 2\frac{J_0 A}{\nu_0} \right) d\Omega \qquad , \tag{5.126}$$

and is valid for magnetostatic problems with different materials and given current densities. Dirichlet and Neumann conditions can be imposed at the boundaries.

5.7.2.1   Functional within an element The functional within an element becomes

$$F(A) = \frac{1}{2} \int_{\Delta_e} \left(v_r^{(e)} |\nabla A|^2 - 2\frac{J_0^{(e)} A}{v_0} \right) d\Omega \quad . \tag{5.127}$$

The evaluation of the second term gives:

$$\int_{\Delta_e} \frac{J_0^{(e)} A}{v_0} d\Omega = \sum_{i=1}^{3} \int_{\Delta_r} \frac{J_0^{(e)}}{v_0} A_i N_i d\Omega = \sum_{i=1}^{3} \underbrace{\frac{J_0^{(e)}}{v_0} \frac{\Delta_e}{3}}_{T_i^{(e)}} A_i \quad . \tag{5.128}$$

In matrix-vector notation, the functional within an element is

$$F^{(e)} = \frac{1}{2} \mathbf{A}^T \mathbf{K}^{(e)} \mathbf{A} - \mathbf{A}^T \mathbf{T}^{(e)} \quad . \tag{5.129}$$

The elements of the source vector $\mathbf{T}^{(e)}$ are found as

$$T_i^{(e)} = \frac{J_0^{(e)}}{v_0} \frac{\Delta_e}{3} \quad . \tag{5.130}$$

5.7.2.2   Solution of Poisson's equation The system of linear equation is found by forcing the partial derivatives with respect to all unknowns to be zero.

$$\begin{bmatrix} \mathbf{K}_{ff} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A}_f \\ \mathbf{A}_p \end{bmatrix} = \begin{bmatrix} -\mathbf{K}_{fp} \mathbf{A}_p + \mathbf{T} \\ \mathbf{A}_p \end{bmatrix} \quad . \tag{5.131}$$

5.7.3   Permanent magnet material
It is assumed that the demagnetisation characteristic in the second quadrant of the magnetisation curve of the considered magnet material is linear (Fig. 5.45). This assumption is realistic for most of the modern rare earth magnets such as the grades of NdFeB, SmCo or the Strontium or Barium Ferrites. The permanent magnet is characterised by its remanence flux density B_r and its coercive force H_c. With the known material equation $B = \mu H$ the reversible permeability $\mu_{rev} = \dfrac{1}{v_{rev}}$ can be determined. Linear permanent magnets are described by:

$$\mathbf{B} = \mu_0 \mu_{rev} \mathbf{H} + \mathbf{B}_r \quad . \tag{5.132}$$

This results in an extra exciting term in the differential equation and in the functional.
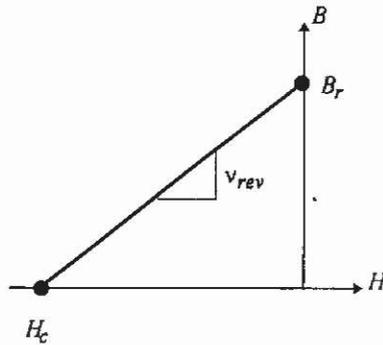
Fig. 5.45. Demagnetisation characteristic of a linear permanent magnet.

$$\nabla \cdot \left(\nu_r \nabla A\right) = \frac{-J_0}{\nu_0} + \nabla\left(\nu_{rev} \mathbf{B}_r\right) \quad , \tag{5.133}$$

$$F(A) = \frac{1}{2} \int_{\Omega} \left(\nu_r \left|\nabla A\right|^2 - 2\frac{J_0 A}{\nu_0} - 2A\nabla\left(\nu_{rev}\mathbf{B}_r\right) \right) d\Omega \quad g \quad , \tag{5.134}$$

with

$$\mathbf{B}_r = B_{rx}\mathbf{e}_x + B_{ry}\mathbf{e}_y \quad . \tag{5.135}$$

This functional is valid for magnetostatic problems where a magnetic field is imposed by currents and permanent magnets.

5.7.3.1   Functional within an element The third term in the functional within an element becomes

$$\int_{\Delta_e} A\nabla\left(\nu_{rev}^{(e)} B_r^{(e)}\right) d\Omega = \sum_{i=1}^{3} \nu_{rev}^{(e)} \int_{\Delta_e} A_i N_i \left(\frac{\partial B_{ry}^{(e)}}{\partial x} - \frac{\partial B_{rx}^{(e)}}{\partial y} \right) d\Omega \quad . \tag{5.136}$$

Applying Green's formula gives:

$$\int_{\Delta_e} A\nabla\left(\nu_{rev}^{(e)} B_r^{(e)}\right) d\Omega = \sum_{i=1}^{3} \nu_{rev}^{(e)} A_i \int_{\Delta_e} \left(B_{ry}^{(e)} \frac{\partial N_i}{\partial x} - B_{rx}^{(e)} \frac{\partial N_i}{\partial y} \right) d\Omega \quad , \tag{5.137}$$

$$\int_{\Delta_e} A\nabla\left(\nu_{rev}^{(e)} B_r^{(e)}\right) d\Omega = \sum_{i=1}^{3} \nu_{rev}^{(e)} A_i \frac{B_{ry}^{(e)} b_i - B_{rx}^{(e)} c_i}{2} \quad . \tag{5.138}$$

This results in an additional term in the source vector $\mathbf{T}^{(e)}$.

$$T_i^{(e)} = \frac{J_s^{(e)}}{\nu_0} \frac{\Delta_e}{3} + \frac{\nu_{rev}^{(e)}}{2}\left(B_{rx}^{(e)} c_i - B_{ry}^{(e)} b_i\right) \quad . \tag{5.139}$$

5.7.3.2   Radial magnetisation The direction of the magnetisation for a radial magnetised permanent magnet must be calculated for each element separately (Fig. 5.46).
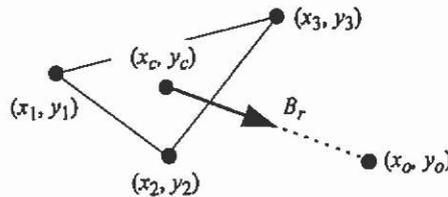
Fig. 5.46. Radial magnetisation of magnet material within an element.

$$B_{x}^{(e)} = \frac{x_{o} - x_{c}}{d} B_{r}^{(e)} \quad , \tag{5.140}$$

$$B_{ry}^{(e)} = \frac{y_{o} - y_{c}}{d} B_{r}^{(e)} \quad , \tag{5.141}$$

with

$$x_{c} = \frac{x_{1} + x_{2} + x_{3}}{3} \quad , \tag{5.142}$$

$$y_{c} = \frac{y_{1} + y_{2} + y_{3}}{3} \quad , \tag{5.143}$$

$$d = \sqrt{(x_{o} - x_{c})^{2} + (y_{o} - y_{c})^{2}} \quad . \tag{5.144}$$

5.7.4 Binary constraints

Binary constraints enforce a relation between values of the magnetic vector potential at a first boundary and values at a second boundary.

$$\mathbf{A}_{1} + k\mathbf{A}_{2} = m \quad . \tag{5.145}$$

Nodes belonging to the first boundary are indicated as binary constrained (index 1), while nodes belonging to the second boundary are seen as free nodes (index 2), yielding the system of linear equations in matrix-vector notation:

$$
\begin{bmatrix}
\mathbf{K}_{ff} & \mathbf{K}_{f1} & \mathbf{K}_{f2} & 0 \\
\mathbf{K}_{1f} & \mathbf{K}_{11} & \mathbf{K}_{12} & 0 \\
\mathbf{K}_{2f} & \mathbf{K}_{21} & \mathbf{K}_{22} & 0 \\
0 & 0 & 0 & \mathbf{I}
\end{bmatrix}
\begin{bmatrix}
\mathbf{A}_{f} \\
\mathbf{A}_{1} \\
\mathbf{A}_{2} \\
\mathbf{A}_{p}
\end{bmatrix}
=
\begin{bmatrix}
\mathbf{T}_{f} - \mathbf{K}_{fp}\mathbf{A}_{p} \\
\mathbf{T}_{1} - \mathbf{K}_{1p}\mathbf{A}_{p} \\
\mathbf{T}_{2} - \mathbf{K}_{2p}\mathbf{A}_{p} \\
\mathbf{A}_{p}
\end{bmatrix}
\quad . \tag{5.146}
$$

Elimination of A_1 using (5.145) gives:

$$
\begin{bmatrix}
\mathbf{K}_{ff} & 0 & \mathbf{K}_{f2} - k\mathbf{K}_{f1} & 0 \\
\mathbf{K}_{1f} & 0 & \mathbf{K}_{12} - k\mathbf{K}_{11} & 0 \\
\mathbf{K}_{2f} & 0 & \mathbf{K}_{22} - k\mathbf{K}_{21} & 0 \\
0 & 0 & 0 & \mathbf{I}
\end{bmatrix}
\begin{bmatrix}
\mathbf{A}_{f} \\
\mathbf{A}_{1} \\
\mathbf{A}_{2} \\
\mathbf{A}_{p}
\end{bmatrix}
=
\begin{bmatrix}
\mathbf{T}_{f} - \mathbf{K}_{fp}\mathbf{A}_{p} - m\mathbf{K}_{f1} \\
\mathbf{T}_{1} - \mathbf{K}_{1p}\mathbf{A}_{p} - m\mathbf{K}_{11} \\
\mathbf{T}_{2} - \mathbf{K}_{2p}\mathbf{A}_{p} - m\mathbf{K}_{21} \\
\mathbf{A}_{p}
\end{bmatrix}
\quad . \tag{5.147}
$$

Symmetrising the matrix (3rd row = 3rd row - k · 2nd row)

$$
\begin{bmatrix}
\mathbf{K}_{ff} & 0 & \mathbf{K}_{f2} - k\mathbf{K}_{f1} & 0 \\
0 & \mathbf{I} & 0 & 0 \\
\mathbf{K}_{2f} - k\mathbf{K}_{1f} & 0 & \mathbf{K}_{22} - k\mathbf{K}_{21} - k\mathbf{K}_{12} + k^2\mathbf{K}_{11} & 0 \\
0 & 0 & 0 & \mathbf{I}
\end{bmatrix}
\begin{bmatrix}
\mathbf{A}_f \\
\mathbf{A}_1 \\
\mathbf{A}_2 \\
\mathbf{A}_p
\end{bmatrix}
$$
$$
=
\begin{bmatrix}
\mathbf{T}_f - \mathbf{K}_{fp}\mathbf{A}_p - m\mathbf{K}_{f1} \\
0 \\
\mathbf{T}_2 - \mathbf{K}_{2p}\mathbf{A}_p - m\mathbf{K}_{21} - k\mathbf{T}_1 + k\mathbf{K}_{1p}\mathbf{A}_p + km\mathbf{K}_{11} \\
\mathbf{A}_p
\end{bmatrix}
\qquad (5.148)
$$

The values of A_1 are calculated with eq.(5.145) after solving the system of equations. Fig. 5.47 shows the matrix assembling.
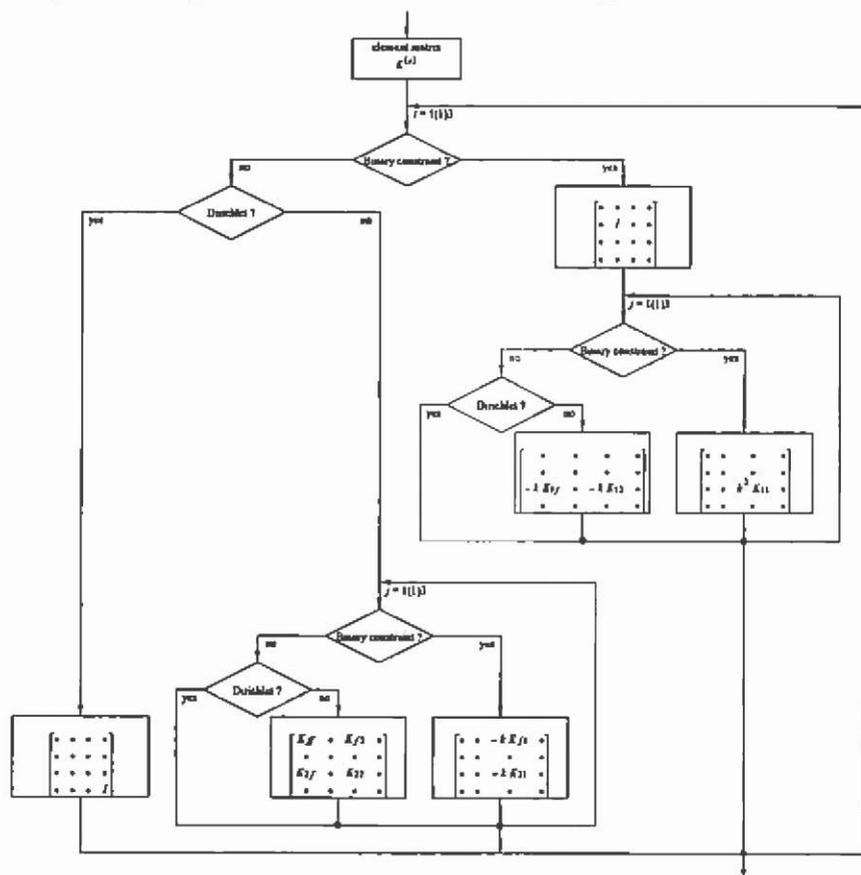
Fig. 5.47. Assembling of the coefficient matrix.

5.7.5 Non-linear materials

To implement the ferromagnetic properties of iron, the non-linear characteristics have to be considered. This results in a system of non-linear equations and this system can not be solved in a closed form. A numerical iteration scheme has to be used to obtain a field solution with the presence of ferromagnetic material. To obtain a solution of such a non-linear system, the Newton iteration can be used.

5.7.5.1   Newton iteration The Newton iteration is a fast approach. Applying appropriate start solutions, its rate of convergence is locally of quadratic order. The numerical solution of a non-linear system is transferred to a series of linear solutions, i.e. in every iteration step a linear system of equation is solved. Various modified Newton methods are in common use.

It is assumed to have the known system of equations:

$$\mathbf{K}(\mathbf{A}) \cdot \mathbf{A} - \mathbf{R} = 0 = \mathbf{F}(\mathbf{A}) \tag{5.149}$$

with $\mathbf{K}$ the coefficient matrix, $\mathbf{A}$ the vector potential and $\mathbf{R}$ the right hand side; we call this the fundamental system $\mathbf{F}(\mathbf{A})$ or residual. When compared to the set of linear equations, in the non-linear case the coefficients of the system matrix are dependent on the solution vector $\mathbf{A}$. The difference between a solution in iteration step (k) and (k+1) is:

$$\mathbf{A}^{(k+1)} = \mathbf{A}^{(k)} + \mathbf{d}^{(k+1)} = \mathbf{A}^{(k)} + \delta\mathbf{A}^{(k)} \tag{5.150}$$

with $\mathbf{d}^{(k+1)}$ as the defect vector. The defect is used as a stopping criterion for the iterations.

The fundamental system is expressed by a Taylor series neglecting the second order and higher terms:

$$0 = \mathbf{F}(\mathbf{A}^{(k+1)}) = \mathbf{F}(\mathbf{A}^{(k)} + \mathbf{d}^{(k+1)}) = \mathbf{F}(\mathbf{A}^{(k)}) + \frac{\partial \mathbf{F}(\mathbf{A}^{(k)})}{\partial \mathbf{A}^{(k)}} \cdot \mathbf{d}^{(k+1)} + \dots \tag{5.151}$$

This yields:

$$-\mathbf{F}(\mathbf{A}^{(k)}) = \mathbf{P}(\mathbf{A}^{(k)}) \cdot \mathbf{d}^{(k+1)} \tag{5.152}$$

with $\mathbf{P}(\mathbf{A}^{(k)})$ the Jacobian matrix from iteration step (k).

$$\mathbf{P}(\mathbf{A}) = \mathbf{F}'(\mathbf{A}) = \frac{\partial \mathbf{F}(\mathbf{A})}{\partial \mathbf{A}} \qquad . \tag{5.153}$$

The defect vector can now be evaluated by using:

$$\mathbf{d}^{(k+1)} = -\mathbf{P}^{-1}(\mathbf{A}^{(k)})\mathbf{F}(\mathbf{A}^{(k)}) \qquad . \tag{5.154}$$

The iteration rule to compute $\mathbf{A}^{(k+1)}$ can be given by:

$$\mathbf{A}^{(k+1)} = \mathbf{A}^{(k)} - \mathbf{P}^{-1}(\mathbf{A}^{(k)})\mathbf{F}(\mathbf{A}^{(k)}) \qquad . \tag{5.155}$$

For practical considerations the formulation of the iteration in this way is not very useful. But solving the system

$$\mathbf{P}^{-1}(\mathbf{A}^{(k)}) \cdot \mathbf{d}^{(k+1)} = -\mathbf{F}(\mathbf{A}^{(k)}) \quad , \tag{5.156}$$

and afterwards evaluating

$$\mathbf{A}^{(k+1)} = \mathbf{A}^{(k)} + \mathbf{d}^{(k+1)} = \mathbf{A}^{(k)} + \delta\mathbf{A}^{(k)} \quad . \tag{5.157}$$

results in the overall solution of the non-linear problem. If the start solution is not chosen close to the exact solution of the system, the iteration may oscillate, diverge and fail. To avoid this, a modified damped Newton method can be implemented.

5.7.5.1.1 Jacobian matrix and final system of equations The Jacobian can be assembled by applying:

$$P_{ij} = K_{ij} + 2\frac{\mathrm{d}\nu}{\mathrm{d}B^2}\Delta_e \sum_{m=1}^{3}\sum_{n=1}^{3} A_m \cdot \left(\nabla N_m \cdot \nabla N_i\right) \cdot \left(\nabla N_n \cdot \nabla N_j\right) \cdot A_n \quad . \tag{5.158}$$

The first term corresponds to the linear equations, while the second term exists only in the presence of a non-linear material. By using an arbitrary matrix $\mathbf{K}^*$ corresponding with the element matrix

$$\mathbf{K}^{(e)} = \nu_r \mathbf{K}^* \quad , \tag{5.159}$$

an arbitrary vector $\mathbf{E}$ can be defined with the elements:

$$E_k = \sum_{m=1}^{3} K_{km}^* A_m \tag{5.160}$$

to obtain a more convenient formulation for the Jacobian matrix.

$$P_{ij} = K_{ij} + \frac{2}{\Delta_e}\frac{\mathrm{d}\nu}{\mathrm{d}B^2}E_i E_j \quad . \tag{5.161}$$

The system of equations to obtain the defect can be written by:

$$
\begin{bmatrix}
\mathbf{P}_{ff} & \mathbf{P}_{f1} & \mathbf{P}_{f2} & 0 \\
\mathbf{P}_{1f} & \mathbf{P}_{11} & \mathbf{P}_{12} & 0 \\
\mathbf{P}_{2f} & \mathbf{P}_{21} & \mathbf{P}_{22} & 0 \\
0 & 0 & 0 & \mathbf{I}
\end{bmatrix}
\begin{bmatrix}
\delta\mathbf{A}_f \\
\delta\mathbf{A}_1 \\
\delta\mathbf{A}_2 \\
\delta\mathbf{A}_p
\end{bmatrix}
$$
$$
=
\begin{bmatrix}
\mathbf{T}_f - \mathbf{K}_{ff}\mathbf{A}_f - \mathbf{K}_{f1}\mathbf{A}_1 - \mathbf{K}_{f2}\mathbf{A}_2 - \mathbf{K}_{fp}\mathbf{A}_p \\
\mathbf{T}_1 - \mathbf{K}_{1f}\mathbf{A}_f - \mathbf{K}_{11}\mathbf{A}_1 - \mathbf{K}_{12}\mathbf{A}_2 - \mathbf{K}_{1p}\mathbf{A}_p \\
\mathbf{T}_2 - \mathbf{K}_{2f}\mathbf{A}_f - \mathbf{K}_{21}\mathbf{A}_1 - \mathbf{K}_{22}\mathbf{A}_2 - \mathbf{K}_{2p}\mathbf{A}_p \\
0
\end{bmatrix}
\tag{5.162}
$$

The source vector S for the calculation of the residual is introduced as:

$$\mathbf{S} = \mathbf{T} - \mathbf{K}\mathbf{A} \tag{5.163}$$

which yields:

$$
\begin{bmatrix} \mathbf{P}_{\!f\!f} & \mathbf{P}_{\!f1} & \mathbf{P}_{\!f2} & 0 \\ \mathbf{P}_{\!1f} & \mathbf{P}_{\!11} & \mathbf{P}_{\!12} & 0 \\ \mathbf{P}_{\!2f} & \mathbf{P}_{\!21} & \mathbf{P}_{\!22} & 0 \\ 0 & 0 & 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \delta\mathbf{A}_{\!f} \\ \delta\mathbf{A}_{\!1} \\ \delta\mathbf{A}_{\!2} \\ \delta\mathbf{A}_{\!p} \end{bmatrix} = \begin{bmatrix} \mathbf{S}_{\!f} \\ \mathbf{S}_{\!1} \\ \mathbf{S}_{\!2} \\ 0 \end{bmatrix} \quad . \tag{5.164}
$$

Elimination of A_1 using (5.125) gives

$$
\begin{bmatrix} \mathbf{P}_{\!f\!f} & 0 & \mathbf{P}_{\!f2}-k\mathbf{P}_{\!f1} & 0 \\ \mathbf{P}_{\!1f} & 0 & \mathbf{P}_{\!12}-k\mathbf{P}_{\!11} & 0 \\ \mathbf{P}_{\!2f} & 0 & \mathbf{P}_{\!22}-k\mathbf{P}_{\!21} & 0 \\ 0 & 0 & 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \delta\mathbf{A}_{\!f} \\ \delta\mathbf{A}_{\!1} \\ \delta\mathbf{A}_{\!2} \\ \delta\mathbf{A}_{\!p} \end{bmatrix} = \begin{bmatrix} \mathbf{S}_{\!f}-m\mathbf{P}_{\!f1} \\ \mathbf{S}_{\!1}-m\mathbf{P}_{\!11} \\ \mathbf{S}_{\!2}-m\mathbf{P}_{\!21} \\ 0 \end{bmatrix} \quad , \tag{5.165}
$$

and symmetrising the matrix by (3rd row = 3rd row - $k \cdot 2^{nd}$ row) yields

$$
\begin{bmatrix} \mathbf{P}_{\!f\!f} & 0 & \mathbf{P}_{\!f2}-k\mathbf{P}_{\!f1} & 0 \\ 0 & \mathbf{I} & 0 & 0 \\ \mathbf{P}_{\!2f}-k\mathbf{P}_{\!1f} & 0 & \mathbf{P}_{\!22}-k\mathbf{P}_{\!21}-k\mathbf{P}_{\!12}+k^2\mathbf{P}_{\!11} & 0 \\ 0 & 0 & 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \delta\mathbf{A}_{\!f} \\ \delta\mathbf{A}_{\!1} \\ \delta\mathbf{A}_{\!2} \\ \delta\mathbf{A}_{\!p} \end{bmatrix}
$$

$$
= \begin{bmatrix} \mathbf{S}_{\!f}-m\mathbf{P}_{\!f1} \\ 0 \\ \mathbf{S}_{\!2}-m\mathbf{P}_{\!21}-k\mathbf{S}_{\!1}+km\mathbf{P}_{\!11} \\ 0 \end{bmatrix} \quad . \tag{5.166}
$$

5.7.5.1.2 Damped Newton iteration

For the Newton method, an appropriate step length of the algorithm has to be chosen to save the iteration from divergence. The damped Newton iteration is a variant of the original method. The defect vector, Newton correction, is damped by a factor α.

$$
\mathbf{A}^{(k+1)} = \mathbf{A}^{(k)} - \alpha \cdot (\mathbf{P}^{-1}(\mathbf{A}^{(k)})\mathbf{F}(\mathbf{A}^{(k)})) \quad . \tag{5.167}
$$

The idea is to accept the iterated $\mathbf{A}^{(k+1)}$ only if the damping criterion

$$
\left\| \mathbf{F}(\mathbf{A}^{(k+1)}) \right\|_2 \leq \left\| \mathbf{F}(\mathbf{A}^{(k)}) \right\|_2 \tag{5.168}
$$

is satisfied to ensure convergence. If this condition is not satisfied, damping steps must be performed until

$$
\left\| \mathbf{F}(\mathbf{A}^{(k)} - \alpha\mathbf{d}^{(k+1)}) \right\|_2 \leq (1-\sigma\alpha)\left\| \mathbf{F}(\mathbf{A}^{(k)}) \right\|_2 \tag{5.169}
$$

with the parameters $\alpha = \beta^j s$ for the damping steps j=0(1)j_{max} is satisfied or a maximum number of steps is performed without success (Kosmol [64]). The parameters are chosen in the range of:

$$\beta = (0,1)$$
$$\sigma = (0,\tfrac{1}{2}) \qquad .$$
$$s > 0$$

(5.170)

Typical values are $\beta = 0.5$, $\sigma = 0.5$, $s = 1$. The maximum number of steps is typically between 5 and 10.

This scheme represents an adaptive damping factor for the Newton iteration. If further problems occur, and the solution already diverges; special algorithms such as gradient steps (Hameyer[49]) can be implemented to try to recover convergence.

5.8   Adaptive refinement for 2D triangular meshes

To obtain a high accuracy of the approximated solution using the FEM, the number of elements has to be high. Refining uniformly causes many finite elements in regions where they are not recommended. The idea of an automated adaptive mesh refinement is to refine the discretisation locally in different iteration steps, by starting with a minimum mesh and terminating this approach with a quality mesh.

The problem in this approach is to know where more elements are recommended and where fewer elements are sufficient to obtain the desired global accuracy. Here, error estimation plays an important role.

To enhance the quality of a finite element discretisation, different strategies can be followed. Errors in the solution can be identified at different stages of the field analysis:

- a priori, before the field is computed
- a posteriori using the field solution to estimate the error.

Adaptive mesh refinement is often a combination of a-priori, a-posteriori error estimation and a refinement algorithm. Error estimators indicate which elements to refine.   Due to a-priori mesh quality indicators, various steps such as node movements or edge swapping are performed. Mesh refinement based on an error estimator improves the convergence when compared to a uniform refinement of all elements. The slope of the global error $|e|$ in Fig. 5.48 is an indication for the convergence rate.

$$|e| = \left| \frac{W - W_{N_n \to \infty}}{W_{N_n \to \infty}} \right|$$

(5.171)

with W the energy stored in the model, N_n the number of nodes, $W_{N_n \to \infty}$ energy stored in a model (exact solution). In practice, $W_{N_n \to \infty}$ is calculated with a sufficiently high number of nodes.
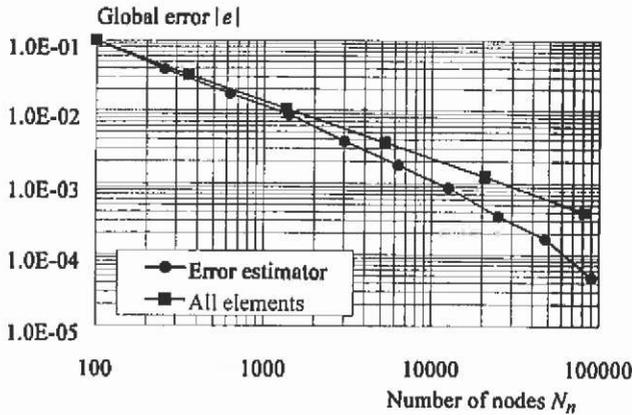
Global error |e|

Fig. 5.48. Global error versus the number of nodes.

The proper combination of the various error estimators with the appropriate element refinement algorithm ensures acceptable computational efforts and efficient use of the available memory resources. In the following sections, some notations are introduced to explain the described refinement techniques.

5.8.1 Type of element refinement

Two different types of isotropic refinements are considered here. It is assumed that the problem solution is unknown and thus no field information is available. Therefore, an anisotropic mesh refinement is not discussed here.

Red I, the first method of refinement, is element-based. Here, a new node is inserted in the centre of an element (Fig. 5.49).

Fig. 5.49. Element and edge-based refinement.

For an element-based refinement, the number of new elements per refined element and the total number of new elements per adaptation step is low because the refinement has no influence on the neighbouring

elements. The initial aspect ratio of the new elements is worse than that of the original element.

The boundaries of the studied domain must be treated in a special way using this strategy (Fig. 5.50).

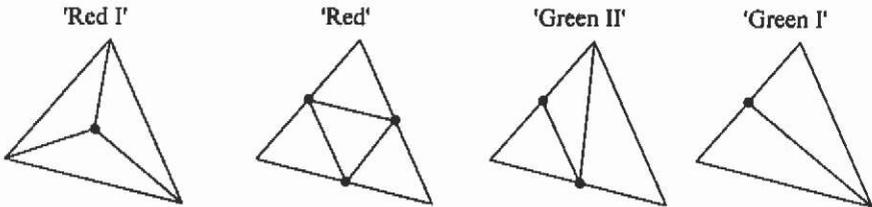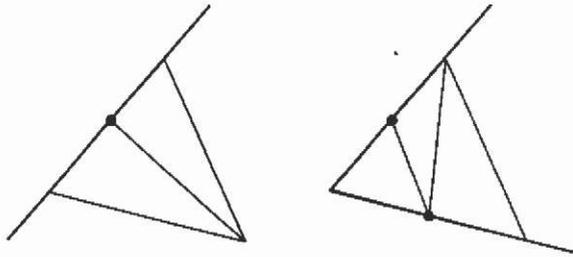Fig. 5.50. Element-based refinement at boundaries.

The second type is an edge-based refinement, i.e. a number of new nodes are inserted along the edges of an element. Using an edge-based refinement approach, the number of new elements per refined element is high and neighbouring elements are influenced. Adjacent elements must be refined with Green I and II as well. Three different types of edge-refinement can be distinguished indicated by the number of marked edges (Fig. 5.49).

No special treatment of boundaries is recommended here. The initial aspect ratio of the new elements is equal to the aspect ratio of the original element.

5.8.2 A priori error estimation

This type of error estimation is performed before a numerical computation to enhance the quality of the FEM mesh. Because no solution is known, geometrical properties, such as the shape or angles of a finite element of the mesh, are chosen to estimate the quality of the discretisation. But such a priori error estimation does not permit a general statement about the accuracy of the field solution because the overall accuracy of the field approximation depends on the precise discretisation of the geometry in regions with a large change of the field quantities as well. Therefore, an a posteriori error estimator is necessary to estimate the relative error of the field solution. Following a particular meshing strategy, minimum or quality mesh, various actions can be taken to minimise the numerical approximation error of the overall field solution before starting the field computation itself.
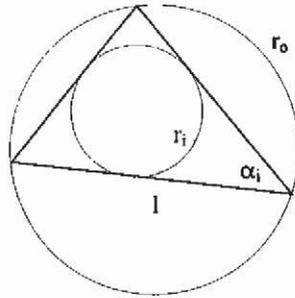
Fig. 5.51. Geometrical definitions for a standard triangular finite element.

Some general rules to generate a uniformly distributed discretisation with triangular elements can be given. Using standard triangular finite elements, the quadratic deviation from the angles of the triangle α_i from the standard angle $\pi/3$ should tend to a minimum:

$$\frac{1}{3}\sum_{i=1}^{3}(\alpha_i - \pi/3)^2 \to \min. \tag{5.172}$$

The ratio of largest and smallest angle in the mesh should tend to unity:

$$\frac{\alpha_{max}}{\alpha_{min}} \to 1 \quad . \tag{5.173}$$

The ratio of maximum to minimum side length of a triangle should tend to unity:

$$\frac{l_{max}}{l_{min}} \to 1 \quad . \tag{5.174}$$

Further details can be found in (Babuska & Aziz [6]).

5.8.2.1 Aspect ratio The aspect ratio γ of a triangular element is calculated as the ratio of radius R of the circumscribed circle to twice the radius r of the inscribed circle (Fig. 5.52). An equilateral triangle has an aspect ratio of one. Therefore, within the discretisation, the ratio of outer to twice the inner radius of a triangle should tend to unity to generate a uniform mesh:

$$\gamma = \frac{R}{2r} \to 1 \quad , \tag{5.175}$$

$$\gamma = \frac{abc}{8(s-a)(s-b)(s-c)} \quad , \tag{5.176}$$

with a, b and c the length of the three edges and s the semi-perimeter of the triangle.

$$s = \frac{a+b+c}{2} \qquad . \tag{5.177}$$
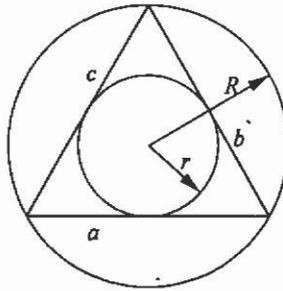
Fig. 5.52. Inscribed and circumscribed circle and edge definition of a triangle.

5.8.2.2 Delaunay triangulation The quality of a generated mesh is characterised by

- the size of the elements referred to the outer dimensions of the area of interest
- an average element aspect ratio close to one
- a low worst element aspect ratio
- the ability to restore the original geometry.

A Delaunay triangulation is a first step to a high quality mesh. Delaunay triangulation is defined in the following way:

For each pair of two adjacent triangles, the minimum of the six angles in the two triangles is larger than it would have been if the diagonal of the quadrilateral had been swapped.

This abstract definition means avoiding small angles and thus long elements with a large aspect ratio. The shape of two considered triangles must be unchanged.
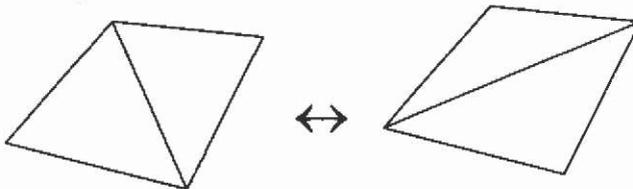
Fig. 5.53. Swapping the diagonals of a quadrilateral, formed by two adjacent triangles.

Each pair of adjacent triangles holding the same region label is tested. The circumscribed circle of one of the triangles is calculated out of the co-ordinates of the nodes. The diagonal is swapped if the fourth node

lies inside the circumscribed circle (Fig. 5.53). The circumscribed circle is in practice never calculated because this is too time consuming.
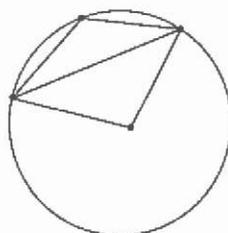
Fig. 5.54. Test for a Delaunay triangulation.

5.8.2.3 Cline and Renka test

If node b lies in the circumscribed circle of the triangle formed by nodes a, c and d, then this yields according to Fig. 5.55:

$$2\pi R - 2R\alpha < 2R\beta \quad , \tag{5.178}$$

$$\alpha + \beta > \pi \quad , \tag{5.179}$$

or

$$\sin(\alpha + \beta) < 0 \quad , \tag{5.180}$$

$$\sin(\alpha) \cdot \cos(\beta) + \cos(\alpha) \cdot \sin(\beta) < 0 \quad . \tag{5.181}$$

Fig. 5.55. Cline and Renka test.

Round-off errors may cause problems if $\sin(\alpha + \beta)$ is close to zero. This happens when:

- $\alpha + \beta$ is near π
- α and β are both near 0
- α and β are both near π.

Node *b* lies almost on the circumscribed circle in the first case and swapping the diagonal has no bad effects. Swapping the diagonal may result in a wrong triangulation in the last two cases. Therefore, additional tests are included in the algorithm.

Table 5.6. Cline and Renka algorithm.

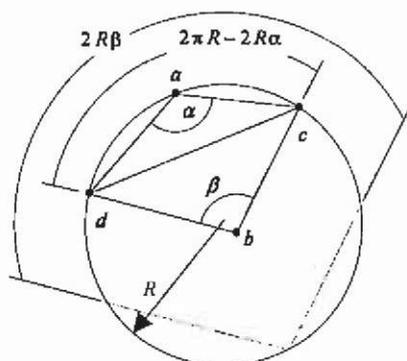|	
Step 1:	$\cos(\alpha) = (x_a - x_c)\cdot(x_a - x_d) + (y_a - y_c)\cdot(y_a - y_d)$ $\cos(\beta) = (x_b - x_c)\cdot(x_b - x_d) + (y_b - y_c)\cdot(y_b - y_d)$

Step 2:	if $(\cos(\alpha) \geq 0$ and $\cos(\beta) \geq 0)$
	   then swap = FALSE
	        exit

Step 3:	if $(\cos(\alpha) < 0$ and $\cos(\beta) < 0)$
	   then swap = TRUE
	        exit

Step 4:	$\sin(\alpha) = (x_a - x_c)\cdot(y_a - y_d) - (y_a - y_c)\cdot(x_a - x_d)$ $\sin(\beta) = (x_b - x_c)\cdot(y_b - y_d) - (y_b - y_c)\cdot(x_b - x_d)$ $\sin(\alpha + \beta) = \sin(\alpha)\cdot\cos(\beta) + \cos(\alpha)\cdot\sin(\beta)$

Step 5:	if $(\sin(\alpha + \beta) < 0)$
	   then swap = TRUE
	        exit
	   else swap = FALSE
	        exit

5.8.2.4 Lawson's Delaunay algorithm This swapping algorithm can be used for an element-based refinement strategy. Fig. 5.56 shows the initial triangulation with 11 nodes and 12 elements. When a new node is inserted, the existing triangulation is updated to a new Delaunay triangulation. This means in practice that the new nodes are inserted on a node-per-node base and each time some tests are performed.

In the Delaunay algorithm of Lawson all the triangles adjacent to the edges opposite the new node are placed on a LIFO-stack (last-in, first-out). A maximum of three elements is placed on the stack in the first step. All elements are taken one by one from the stack and a test is made as to whether the new node lies inside the circumscribed circle. If the new node lies in the circumscribed circle, the diagonal of the quadrilateral is swapped to generate two better-shaped elements. The triangles that are now adjacent to the edges opposite the new node are added to the stack.
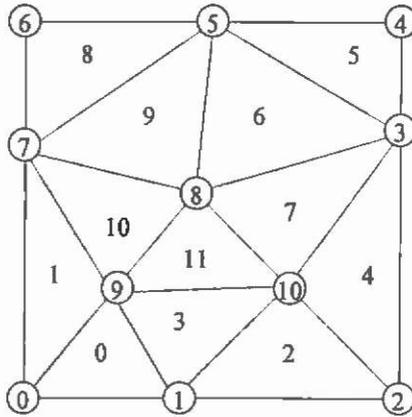
Fig. 5.56. Initial triangulation.

Fig. 5.57 illustrates Lawson's algorithm for the initial triangulation. According to an a posteriori error estimator, element number 6 is chosen to be refined. The new node receives the number 11 and is inserted in the centre of the element (Fig. 5.57a). Fig. 5.57b shows the three new elements. One of the three new elements gets the original element number and the other two get the numbers 12 and 13. For each inserted node the number of elements increases by two. The three elements adjacent to the edges opposite the new node are put on the stack (Fig. 5.57c).

The element that is put last on the stack is taken first from the stack and tested as to whether the new node lies inside the circumscribed circle of the element (Fig. 5.57d). If so, the diagonal of the quadrilateral is swapped (Fig. 5.57e). Because there are no elements that are now adjacent to the edges opposite the new node, no elements are added to the stack. The next element is taken from the stack and is tested for whether the new node lies inside the circumscribed circle (Fig. 5.57f). The diagonal of the quadrilateral is swapped (Fig. 5.57g). Two elements are now adjacent to the edges opposite the new node and are added to the stack (Fig. 5.57h). Three elements are now on the stack.

The next element is taken from the stack and a test is performed (Fig. 5.57i). No swapping is necessary and the same applies for the next element on the stack (Fig. 5.57j). The last element from the stack is tested (Fig. 5.57k). The new node does not lie inside the circumscribed circle of the last element taken from the stack. Fig. 5.57*l* shows the resulting Delaunay triangulation.
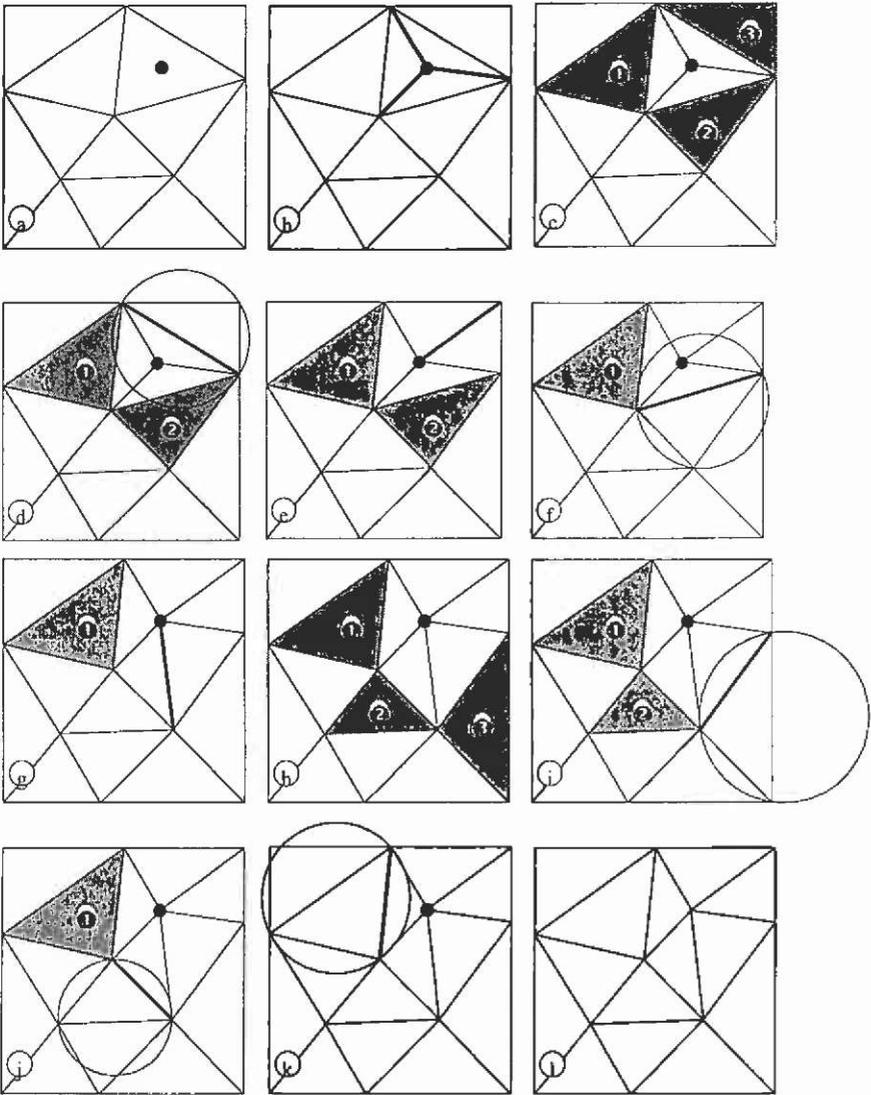
Fig. 5.57. Lawson's Delaunay algorithm.

5.8.3 Reconstruction of the original geometry

Another characteristic of a refinement algorithm to obtain high quality meshes is the ability to restore the original geometry when elements are refined. Of each edge of an element on an outline, the primitive (arc, circle or line) to which it belongs, is stored. If an edge of an arc or circle

is refined, the new node is not inserted in the middle of the edge. The new node is inserted in the middle in terms of polar co-ordinates (Fig. 5.58).

$$\rho_n = \frac{\rho_i + \rho_j}{2} \quad . \tag{5.182}$$
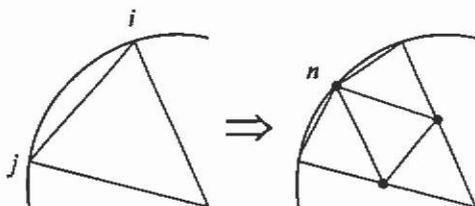
Fig. 5.58. Restoring to the geometry during refinement.

5.8.4 Moving nodes

Moving the nodes results in a lower average aspect ratio. A node is moved to the centre of the surrounding nodes (Fig. 5.59). A test is performed so that no elements with a negative area are introduced.

$$x_n = \frac{1}{n}\sum_{i=0}^{n-1} x_i \quad , \tag{5.183}$$

$$y_n = \frac{1}{n}\sum_{i=0}^{n-1} y_i \quad . \tag{5.184}$$
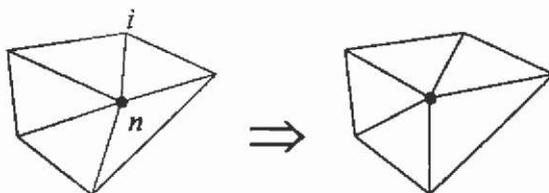
Fig. 5.59. Moving the nodes.

An edge-based refinement is easily extended with the moving node approach. A node is moved to the centre of the surrounding nodes and all pairs of two adjacent surrounding elements are tested. If swapping of a diagonal was necessary, the node is moved again to the new centre of the surrounding nodes. The loop is repeated until the number of swaps is sufficiently low. Fig. 5.61a shows the triangulation before moving the nodes, after one loop (Fig. 5.61b) and after two loops (Fig. 5.61c) of moving and swapping. No diagonals were swapped after the two loops.

```
for (i=0;i<numnode;i++)
      movenode(i);
      swapedge(i,swapped);
      if (swapped = TRUE)
            movenode(i);
```

Fig. 5.60. Extended test algorithm.
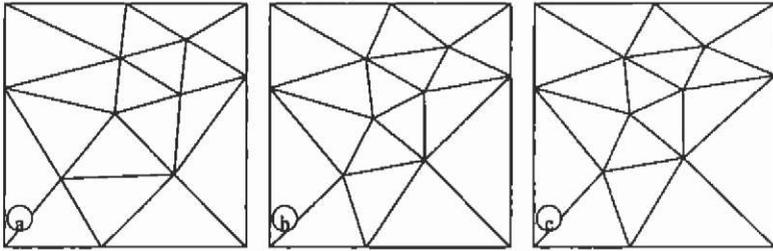
Fig. 5.61. Triangulation after moving the nodes.

5.8.5 Starting solution for the next adaptation step

An iterative solver to solve the system of linear equations needs a starting solution. The first start solution is often the zero-solution. The starting solution of the next adaptation step is determined by the interpolated solution of the previous step before movement of the nodes. When an edge is refined, the average of the solution of the two nodes forming the edge is taken as the solution for the new node. A starting solution obtained in this way can reduce the number of iterative steps by 25 %. A reduction of one or more Newton steps is possible for strongly saturated problems.

5.8.6 Additional refinement rules

Marking the edges to be refined is performed in four steps to improve the quality of the generated mesh, even in regions where no elements have to be refined.

- To improve the initial average aspect ratio of the new elements, the edges of elements with an aspect ratio > 5.0 are marked in a special way (Fig. 5.62).
- All edges of an element to be refined are marked.
- All non-outline edges that touch two outlines are marked for refinement during the first adaptation step (Fig. 5.63).
- All edges that touch an outline edge to be refined are marked (Fig. 5.64).

Fig. 5.62. Refinement of elements with an aspect ratio > 5,0.
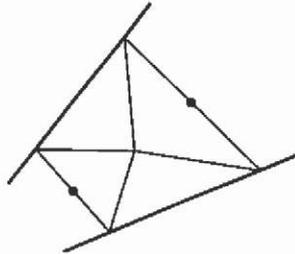
Fig. 5.63. Refinement of non-outline edges touching two outlines.
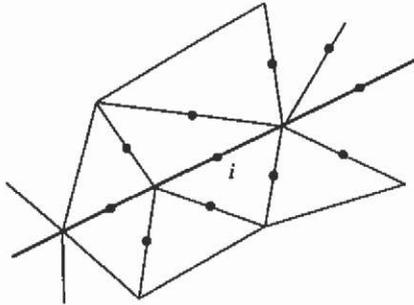
Fig. 5.64. Refinement of edges touching an outline edge i.

5.8.7 A posteriori error estimation

An a posteriori error estimator evaluates the local error e_k per element, based on a previously computed solution. This error estimation can be performed in two different ways:

- model based and
- region, label based.

To obtain a local error between 0 and 1 for the model based strategy, it is normalised, i.e. divided by the maximum error over all elements in the model.

$$e_k = \frac{e_k}{\max(e_k)} \qquad . \tag{5.185}$$

This type of estimation is called model error estimation. A maximum accepted local error ε_{local} (e.g. 10 %) can be specified and all elements with a larger error bound are considered for refinement.

$$e_k > \varepsilon_{local} \quad . \tag{5.186}$$

The elements are arranged according to their local error and the elements with the highest error are refined first. The number of refined elements is determined by the maximum percentage of new elements p specified (e.g. 100 %). The maximum number of elements N_r marked for refinement is given by:

$$N_r = p \cdot c \cdot N_e \quad . \tag{5.187}$$

with N_e the number of elements, c a constant depending on the refinement method. The constant c indicates the ratio of the number of refined elements to the number of new elements: $c = \frac{1}{4}$ for an edge-based refinement and $c = \frac{1}{3}$ for an element-based refinement.
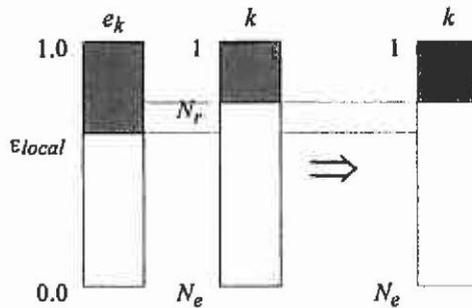
Fig. 5.65. Determining the elements to be refined.

A label-based error estimator calculates the local error within each region label separately and the error is normalised by division with the maximum error within the label. The maximum number of elements $N_{r,l}$ to be refined is determined by the number of elements $N_{e,l}$ inside the label.

$$N_{r,l} = p \cdot c \cdot N_{e,l} \quad . \tag{5.188}$$

Elements are refined all over the model, but also in regions (labels) where it is probably not necessary. This, if not desired explicitly, represents a disadvantage of a label-based refinement.

The local error can be evaluated in various ways. The error indicator has influence on the generated mesh and its proper choice is dependent on the problem definition as such, time-harmonic or static magnet field solution and on the strategy of refinement, minimum or quality meshing. Therefore, no indication regarding use of a particular estimator or estimation scheme will be given.

5.8.7.1   Energy in an element The energy W_k in an element is an indication for the local error e_k.

$$e_k = W_k \tag{5.189}$$

5.8.7.2   Flux density at a node The magnetic flux density $B_{n,i}$ at a node i is the weighted average of the values of the magnetic flux density $B_{e,k}$ of the surrounding elements k belonging to the same label (Fig. 5.66). The weighting factor is the area Δ_k of the elements.

$$B_{n,i} = \frac{\displaystyle\sum_{k=1}^{N} B_{e,k} \cdot \Delta_k}{\displaystyle\sum_{k=1}^{N} \Delta_k} \tag{5.190}$$

The local error e_k of an element is calculated as the difference between the magnetic flux density of the element and the average of the nodal values.

$$\tag{5.191}$$

$$e_k = \left| B_{e,k} - \frac{\displaystyle\sum_{i=1}^{3} B_{n,i}}{3} \right| \quad .$$
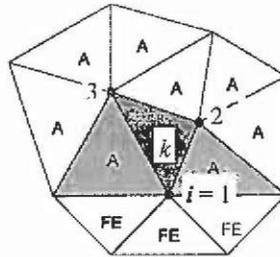
Fig. 5.66. Calculation of the magnetic flux density at a node.

5.8.7.3   Magnetic flux density in a node weighted with the energy The local error e_k based on the magnetic flux density in a node is weighted with the energy W_k in an element.

$$\tag{5.192}$$

$$e_k = W_k \left| B_{e,k} - \frac{\displaystyle\sum_{i=1}^{3} B_{n,i}}{3} \right| \quad .$$

5.8.8   Numerical implementation

Error estimation and mesh enhancements require computation time. Searching in very large data structures absorbs most of this time. To optimise such efforts with respect to the computational costs towards a

better approximation of a field solution, 'intelligent' data structures are recommended to accelerate the search routines.

Starting from the element matrix '*elem*' indicating how the elements are built out of the nodes, the node-to-element matrix '*nodeelem*' is built in a straightforward way (Fig. 5.67). This matrix gives for each node the surrounding elements. The number of surrounding elements is stored in the vector '*surrelem*'. Fig. 5.68 gives the element matrix for the initial triangulation of Fig. 5.56 and Fig. 5.69 the node-to-element matrix and the number-of-surrounding-elements vector.

```
for (k=0;k<numelem;k++)
     for (i=0;i<3;i++)
          nodeelem[elem[k,i],surrelem[elem[k,i]]]=k;
          surrelem[elem[k,i]]++;
```

Fig. 5.67. Building the node-to-element matrix.

0	0	1	9
1	0	9	7
2	1	2	10
3	1	10	9
4	2	3	10
5	3	4	5
6	3	5	8
7	3	8	10
8	5	6	7
9	5	7	8
10	7	9	8
11	8	9	10

Fig. 5.68. Element matrix.

0	0	1					2
1	0	2	3				3
2	2	4					2
3	4	5	6	7			4
4	5						1
5	5	6	8				4
6	8						1
7	1	8	9	10			4
8	6	7	9	10	11		5
9	0	1	3	10	11		5
10	0	3	4	7	11		5

Fig. 5.69. Node-to-element matrix and number-of-surrounding-elements vector.

The neighbouring-element matrix is now built. For each edge of an element the neighbouring element is determined. The two rows of the node-to-element matrix are searched for the common element numbers. The element itself is found and if another element number is found, it is that of the neighbouring element. As shown in Fig. 5.69, the rows of the node-to-element matrix are sorted and the number of surrounding elements is almost a constant. This means that very effective and fast search algorithms can be used. Fig. 5.70 shows the neighbouring-element matrix for the initial triangulation. Binary constraints are dealt with during the building of the neighbouring-element matrix. Elements connected through binary constraints are neighbouring elements.

0		3	1
1	0	10	
2		4	3
3	2	11	0
4		7	2
5			6
6	5	9	7
7	6	11	4
8			9
9	8	10	6
10	1	11	9
11	10	3	7

Fig. 5.70. Neighbouring-element matrix.

5.8.8.1 Edge-based refinement An edge-based refinement is implemented using a neighbouring-element matrix. The refinement of an element influences the neighbouring elements. An error estimator marks the elements to be refined. All the edges of the marked elements are marked for refinement. Putting the number of the new node according to the edge to be refined in the to-be-refined-edges matrix *'refedge'* does this. At the same time, the appropriate edge of the neighbouring element is marked with the same node number. When two elements are connected through binary constraints, the next node number is used. Fig. 5.71 shows the to-be-refined-edges matrix according to the refinement of Fig. 5.56. Again only element 6 is to be refined ('Red' element). Elements 5, 7 and 9 are 'Green I' elements.

0			
1			
2			
3			
4			
5			11
6	11	12	13
7	13		
8			
9			12
10			
11			

Fig. 5.71. To-be-refined-edges matrix.
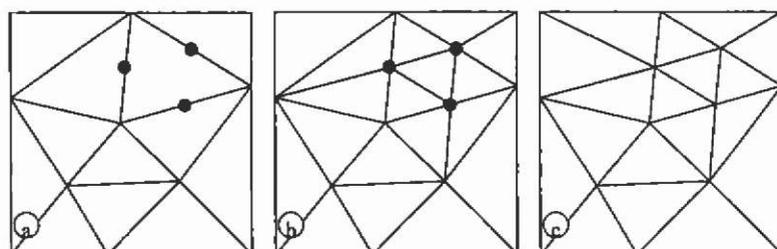

Fig. 5.72. Edge-based refinement.

After all edges are marked for refinement, each element is refined according Fig. 5.72. Testing as to whether the intermediate triangulation is a Delaunay triangulation is not done. A Delaunay triangulation is unique and therefore it is sufficient to use a global Delaunay algorithm to obtain a Delaunay triangulation with the new nodes. A possible algorithm is a loop over all nodes and testing all pairs of two adjacent surrounding elements. The number of swapped diagonals is counted and the loop is repeated until there are no swaps performed. The resulting triangulation is a Delaunay triangulation (Fig. 5.72c). In practice, the loop is repeated until the number of swaps is less than 10 % of the number of swaps in the first loop. Because the number of swaps is rather low for an edge-based refinement, the quality of the generated mesh is high, but the mesh is strictly speaking not a Delaunay triangulation.

5.9 Coupling of field and circuit equations

In the FEM models considered up to now it was assumed that current densities, permanent magnets and/or given potential distributions were imposed as field exciting sources. Many models of technical relevance

can be modelled using the mentioned sources. But electromechanical devices are operated by currents and/or voltages generated by a power supply. Thus, such real devices are fed from electrical circuits. Such external circuits consist of inductances, capacitances, resistances, current and/or voltage sources.

The two-dimensional model of an induction motor does not consider the resistor of the cage end-windings. The connection of the rotor bars is not modelled in a simple two-dimensional cross-sectional model. Therefore, additional external circuits with resistive elements have to be defined to model the short-cut rings of the rotor cage.

In the following sections such external circuit equations will be defined and coupled to the two-dimensional finite element equations. The types of stranded and solid conductor will be introduced and implemented in the FEM model. Both types can be handled simultaneously in one model using a mixed formulation.

The same notation as used in the previous sections is assumed and used.

5.9.1 Time harmonic problem

Regions with eddy currents and applied voltage gradient are called solid conductors.

The partial differential equation for a two-dimensional time-harmonic magnetic problem in solid conductors is:

$$\nabla \cdot (\nu \nabla A) - j\omega \cdot \sigma A = -J_0 \quad . \tag{5.193}$$

To stay in the same notation as in section 5.7 using the relative reluctivity ν_r and the permeability of the free space μ_0, we can write:

$$\nabla \cdot (\nu_r \nabla A) - j\omega \cdot \mu_0 \sigma A = -\mu_0 J_0 \quad . \tag{5.194}$$

Regions without eddy currents but with applied current density J_0 represent stranded conductors. They are described by Poisson's equation:

$$\nabla \cdot (\nu_r \nabla A) = -\mu_0 J_0 \quad . \tag{5.195}$$

For non-conducting regions Laplace's equation is valid:

$$\nabla \cdot (\nu_r \nabla A) = 0 \quad . \tag{5.196}$$

Different materials with respect to their conductivity σ and permeability μ can be considered element by element.

The corresponding energy-minimum functionals are
- for solid conductors:

$$F(A) = \frac{1}{2} \int_{\Omega_{solid\ conductors}} \left(\nu_r |\nabla A|^2 + j\omega \cdot \mu_0 \sigma A^2 - 2\mu_0 J_0 A \right) d\Omega \quad ; \tag{5.197}$$

- for stranded conductors:

$$F(A) = \frac{1}{2} \int_{\Omega_{stranded\ conductors}} \left(\nu_r |\nabla A|^2 - 2\mu_0 J_0 A \right) d\Omega \quad ; \tag{5.198}$$

- and for non-conducting regions:

$$F(A) = \frac{1}{2} \int_{\Omega_{non-conducting\ regions}} \nu_r |\nabla A|^2 \, d\Omega \quad . \tag{5.199}$$

5.9.1.1   Functional within an element For non-conducting regions and regions without eddy currents (stranded conductors) the 3x3 element matrix $\mathbf{K}^{(e)}$ remains the same as in the magneto static case (5.123).

$$K_{ij}^{(e)} = \nu_r^{(e)} \frac{b_i b_j + c_i c_j}{4\Delta_e} \quad . \tag{5.200}$$

For regions with eddy currents (solid conductors) the functional within an element is:

$$F^{(e)}(A) = \frac{1}{2} \int_{\Delta_e} \left(\nu_r^{(e)} |\nabla A|^2 + j\omega\mu_0 \sigma^{(e)} A^2 - 2\mu_0 J_0^{(e)} A \right) d\Omega \quad . \tag{5.201}$$

The second term is:

$$\frac{1}{2} \int_{\Delta_e} \left(j\omega\mu_0 \sigma^{(e)} A^2 \right) d\Omega = \frac{1}{2} \sum_{i=1}^{3} \sum_{j=1}^{3} A_i \underbrace{j\omega\mu_0 \sigma^{(e)} \int_{\Delta_e} N_i N_j \, d\Omega}_{L_{ij}^{(e)}} A_j \quad . \tag{5.202}$$

The terms of the 3x3 eddy current matrix $\mathbf{L}^{(e)}$ are given by:

$$L_{ii}^{(e)} = j\omega\mu_0 \sigma^{(e)} \frac{\Delta_e}{6} \quad , \tag{5.203}$$

$$L_{ij}^{(e)} = j\omega\mu_0 \sigma^{(e)} \frac{\Delta_e}{12} \quad . \tag{5.204}$$

The matrix $\mathbf{L}^{(e)}$ is symmetric. In matrix-vector notation, the functional within an element becomes:

$$F^{(e)} = \frac{1}{2} \mathbf{A}^T \left(\mathbf{K}^{(e)} + \mathbf{L}^{(e)} \right) \mathbf{A} - \mathbf{A}^T \mathbf{T}^{(e)} \quad . \tag{5.205}$$

5.9.1.2   Source vector For regions without eddy current but with an applied current density J_0 (stranded conductors) the source vector $\mathbf{T}^{(e)}$ remains the same as in the static case (5.130).

$$T_i^{(e)} = \mu_0 \frac{\Delta_e}{3} J_0^{(e)} \quad . \tag{5.206}$$

Windings of electrical machines can consist of a number of conductors connected in parallel with geometrical dimensions such that at the considered frequency eddy currents can be neglected. For this case, and stranded conductor p, the source vector is written in terms of the current per strand $I_{str,p}$.

$$T_i^{(e)} = \mu_0 \frac{\Delta_e}{3} \frac{N_{t,p}}{\Delta_{str,p}} I_{str,p} = \mu_0 P_{t,p}^{(e)} I_{str,p} \qquad . \tag{5.207}$$

$N_{t,p}$ is the number of turns of the winding or the strands of conductor p and $\Delta_{str,p}$ is the area of the stranded conductor in the FEM mesh.

In the case of a solid conductor q, the source vector can be written in terms of its voltage drop $V_{sol,q}$:

$$T_i^{(e)} = \mu_0 \frac{\Delta_e}{3} \frac{\sigma^{(e)}}{\ell} V_{sol,q} = \mu_0 Q_{t,q}^{(e)} V_{sol,q} \tag{5.208}$$

with ℓ the length of the conductor.

5.9.1.3   System of equations The system of linear equations is assembled in the same way as for the magneto static problem.

$$\begin{bmatrix} K_{ff} & 0 \\ 0 & I_{pp} \end{bmatrix} \begin{bmatrix} A_f \\ A_p \end{bmatrix} = \begin{bmatrix} -K_{fp} A_p + T \\ A_p \end{bmatrix} \qquad . \tag{5.209}$$

Dirichlet boundary conditions are considered in eq.(5.209) and binary conditions are omitted.

If the currents in the stranded conductors and the voltage drops over the solid conductors are unknown, extra circuit equations have to be added to the system. These equations can be seen as boundary conditions. The circuit conditions act as a lumped parameters model that is applied to the boundary of the differential problem.

If $I_{str,p}$ and $V_{sol,q}$ are unknowns, the system in (5.209) becomes:

$$\begin{bmatrix} K_{ff} & 0 & | & -\mu_0 P_{f,p} & -\mu_0 Q_{f,q} \\ 0 & I_{pp} & | & 0 & 0 \end{bmatrix} \begin{bmatrix} A_f \\ A_p \\ \hline I_{str,p} \\ V_{sol,q} \end{bmatrix} = \begin{bmatrix} -K_{fp} A_p \\ A_p \end{bmatrix} \qquad . \tag{5.210}$$

5.9.1.4   Stranded conductors in eddy current problems The skin depth of the current into a conducting material is given by:

$$\delta = \sqrt{\frac{2}{\omega \mu \sigma}} \tag{5.211}$$

with ω the angular frequency of the current, μ the permeability and σ the conductivity of the conductor. In general, it is assumed that the current density is uniformly distributed across the skin depth. However, when a number of conductors are connected in parallel, the induced voltage depends on both the current density and the magnetic vector potential (5.212).

The entire voltage drop over the stranded conductors is calculated by using the average voltage over the conducting area:

$$V_{str,p} = N_{t,p} \cdot V_{str,p}\Big|_{average} = \frac{N_{t,p}}{\Delta_{str,p}} \int_{\Omega_{str}} \ell\left(\frac{J_{str,p}}{\sigma f_{str,p}} + j\omega A\right) d\Omega \tag{5.212}$$

with

$$J_{str,p} = \frac{N_{t,p}}{\Delta_{str,p}} I_{str,p} \tag{5.213}$$

When a stranded conductor is considered as a region of filamentary conducting wires, the fill factor $f_{str,p}$ for the conducting material is defined as the ratio of the surface of the conducting material to the entire surface of the conductor. In this way, insulating material can be considered.

$$\begin{aligned} V_{str,p} &= \ell\frac{N_{t,p}}{\Delta_{str,p}}\left(\frac{J_{str,p}}{\sigma f_{str,p}}\Delta_{str,p} + j\omega\sum_{e=1}^{n_e^{str}}\sum_{j=1}^{3} A_j \frac{\Delta_e}{3}\right) \\ &= \ell\underbrace{\frac{N_{t,p}^2}{\sigma f_{str,p}\Delta_{str,p}} I_{str,p}}_{R_{str,p}} + \underbrace{j\omega\ell\sum_j P_{j,p}A_j}_{V_{str,p}^{ind}} \end{aligned} \tag{5.214}$$

From eq.(5.214) it can be taken that the total voltage drop consists of a resistive component $V_{str,p}^{res} = R_{str,p}I_{str\,p}$ and an inductive component $V_{str,p}^{ind}$. Therefore, in an electric network model, a stranded conductor can be represented by a series connection of an ohmic resistor and a controlled voltage source (Fig. 5.73).
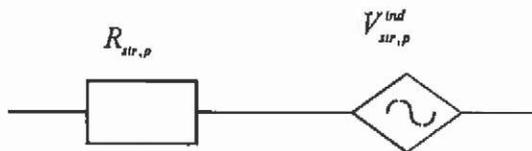

Fig. 5.73. Network elements representing a stranded conductor.

5.9.1.5   Solid conductors in eddy current problems In the case of solid conductors, the gradient of the voltage is assumed to be constant over the surface of the conductor. The eddy currents are a function of both the voltage gradient and the magnetic vector potential:

$$
\begin{aligned}
I_{sol,q} &= \int_{\Delta_{sol,q}} J_e \, d\Omega = \int_{\Delta_{sol,q}} \left(-\sigma \nabla V - j\omega\sigma A \right) d\Omega \\
&= -\frac{\sigma}{\ell} \Delta_{sol} V_{sol,q} - j\omega\sigma \sum_{e=1}^{sol} \sum_{j=1}^{3} A_j \frac{\Delta_e}{3} \\
&= \underbrace{\frac{\sigma \Delta_{sol,q}}{\ell}}_{G_{sol,q}} V_{sol,q} \underbrace{- j\omega\ell \sum_{j} Q_{j,q} A_j}_{I^{ind}_{sol,q}} \ .
\end{aligned}
\tag{5.215}
$$

From eq.(5.215) it can be taken that the current consists of an admittance current $I^{adm}_{sol,q} = G_{sol,q} V_{sol,q}$ and an eddy current $I^{ind}_{sol,q}$ component. In the circuit analysis, a solid conductor can be modelled by admittance and a controlled current source connected in parallel (Fig. 5.74).
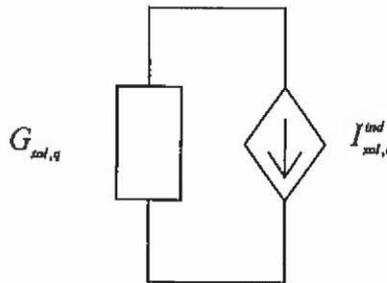

Fig. 5.74. Network elements representing a solid conductor.

5.9.2   Coupled field–circuit equations

To solve the time-harmonic magnetic field, the unknown magnetic vector potentials can be calculated by evaluating the:

- known potentials at the boundary (Dirichlet)
- known current densities in stranded conductors
- known voltage gradients in solid conductors.

In reality, it is not possible to know these quantities from the system before solving the magnetic field because they are dependent on each other. The electric and magnetic parameters, and thus the fields of the system interact very strongly.

It is recommended, therefore, that the currents of stranded conductors and the voltage drops in solid conductors are considered as unknowns of the system. The equations (5.215) and (5.214), describing an external electric circuit (Fig. 5.73 and Fig. 5.74), have to be added to the set of field equations to obtain a coupled system.

The coupling of the magnetic field with the electric circuit equations can be obtained numerically strong or by a weak coupling.

Using some initial values, a first computation of the magnetic field is performed. Out of this field solution, the induced voltage drops over the stranded conductors and the eddy currents in solid conductors are evaluated. By using these results, the electrical quantities of the electric circuit network are calculated, supplying new values for the currents in the stranded conductors and the voltage drops over the solid conductors, to be used for a new calculation of the magnetic field.

To solve the entire coupled field-circuit problem, an iterative procedure can be applied. Both partial problems are solved in successive steps. This approach is a numerically weak coupling of the two systems.

To obtain a numerically strong coupling, it is possible to assemble all unknowns in one vector and to combine all equations describing the system in one matrix. All equations describing the system are solved simultaneously. This approach is called a numerically strong coupling of the systems. If all stranded conductors are voltage driven and if all solid conductors are current driven, the coupled matrix is given by:

$$
\begin{bmatrix}
\mathbf{K}_{ff} & 0 & -\mu_0\mathbf{P}_{i,p} & -\mu_0\mathbf{Q}_{i,q} \\
0 & \mathbf{I}_{pp} & 0 & 0 \\
\hline
j\omega\ell\mathbf{P}_{j,p} & 0 & \mathbf{R}_{str,p} & 0 \\
-j\omega\ell\mathbf{Q}_{j,q} & 0 & 0 & \mathbf{G}_{sol,q}
\end{bmatrix}
\begin{bmatrix}
\mathbf{A}_f \\
\mathbf{A}_p \\
\hline
\mathbf{I}_{str,p} \\
\mathbf{V}_{sol,q}
\end{bmatrix}
=
\begin{bmatrix}
-\mathbf{K}_{fp}\mathbf{A}_p \\
\mathbf{A}_p \\
\hline
\mathbf{V}_{str,p} \\
\mathbf{I}_{sol,q}
\end{bmatrix}
. \quad (5.216)
$$

Multiplying (5.215) with a factor $\chi = \dfrac{\mu_0}{j\omega\ell}$ and (5.214) with $-\chi$

results in a symmetrical matrix.

5.9.2.1   Mixed stranded and solid conductors Mixed stranded and solid conductors in a connected network cause problems describing the circuit. The matrices obtained by a separate analysis of stranded and solid conductors can not be arranged together.

The circuit theory indicates a problem while enumerating tree and co-tree branches. In Fig. 5.75a one of the three solid conductors has to be considered as a link. In Fig. 5.75b only one of the stranded conductors can be considered as a tree branch.

Replacing the magnetic branches, as indicated in Fig. 5.73 and Fig. 5.74, fails. In this case cut-sets containing a stranded conductor branch

include solid conductor branches. Loops holding a solid conductor branch include stranded conductor branches.

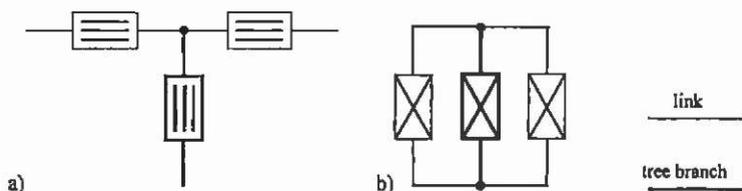To solve this problem, the connected network has to be represented in another form.

a)   b)   link

                                                                        tree branch

Fig. 5.75. a) Star-connected stranded conductors and b) solid conductors connected in parallel.

5.9.3   Network topology

With a network topology, lumped parameter networks obeying three basic laws:

- Kirchhoff voltage law (KVL),
- Kirchhoff current law (KCL) and
- Branch current-voltage relations (BCVR)

can be studied. A complete description of the network delivers information on:

- the connection of branches,
- the reference directions for branch currents and voltages and
- the branch characteristics.

Items 1 and 2 can be represented by a directed graph (Fig. 5.76).

By defining loops, cut-sets and a tree, the network description can be performed systematically. Loops are the sub-graphs to which KVL is applied. Cut-sets are the sub-graphs to which a generalised KCL is applied. The concept of a tree is a tool for a systematic formulation of independent KCL and KVL equations.
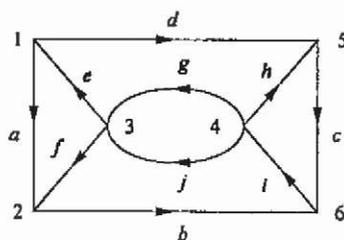

Fig. 5.76. Directed graph.

5.9.3.1 Definitions

Path

A set of branches is called a path between two nodes p and q, if the branches can be labelled in such a way that:

- consecutive branches always have a common endpoint,
- no node is the endpoint of more than two branches in the set, and
- p is the endpoint of exactly one branch in the set, and so is q.

Thus, A path is just a route between two nodes. In Fig. 5.76 branches (*dhib*) form a path between nodes 1 and 2.

Connected graph

An undirected graph is a connected graph if there exists a path between any two nodes of the graph. A network is connected if the associated graph is connected. The graph in Fig. 5.76 is connected.

Loop

A sub-graph of a graph is called a loop if

- the sub-graph is connected, and
- every node of the sub-graph has exactly two branches incident at it.

For example, in Fig. 5.76 the branches (*abcd*) form a loop.

Tree, co-tree, branches and links

A sub-graph of a connected graph is called a tree if

- the sub-graph is connected,
- the sub-graph contains all nodes of the graph, and
- the sub-graph contains no loops.

For example, in Fig. 5.76, the branches (*aedgi*) form a tree. The branches belonging to a tree are called tree branches, and those that do not belong to a tree are called links. All the links of a given tree T form what is called a co-tree with respect to the tree T. It can be shown that for a connected graph with n nodes, any tree T has exactly $n-1$ tree branches. Furthermore, each set of $n-1$ branches without loops constitutes a tree.

Cut-set

A set of branches of a connected graph is said to be a cut-set if

- the removal of the set of branches (but not their endpoints) results in a graph that is not connected and

- after the removal of the set of branches, the restoration of any branch from the set will result in a connected graph again.

For example in Fig. 5.76, the branches (*aed*) or (*dgjb*) form a cut-set.

Incidence matrix

An incidence matrix of a directed graph with n nodes and b branches is a $n \times b$ matrix $\mathbf{A_a} = \left[a_{ij} \right]$ where:

- $a_{ij} = 1$ if branch j is incident to node i, and the arrow is pointing away from node i,
- $a_{ij} = -1$ if branch j is incident to node i, and the arrow is pointing towards node i, and
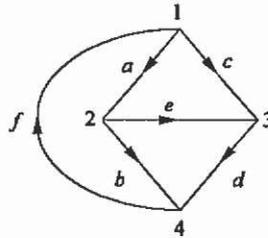- $a_{ij} = 0$ if branch j is not incident to node i.

Fig. 5.77. Directed graph.

For example, for the directed graph from Fig. 5.77:

$$A_a = \begin{array}{c} node \\ \\ \\ \\ \\ \end{array} \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{ccccccc} a & b & c & d & e & f & incident\ branches \\ \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & -1 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & -1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \end{bmatrix} \end{array}$$

(5.217)

One of the rows of $\mathbf{A_a}$ is linearly dependent on the other rows. The matrix $\mathbf{A}$ obtained from $\mathbf{A_a}$ by omitting a row is called a reduced incidence matrix. $\mathbf{A_a}$ is called the complete incidence matrix.

The KCL for the network can be written in matrix-vector notation as

$$\mathbf{A_a i} = 0 \tag{5.218}$$

The maximum set of independent KCL equations, obtained from the nodes of a connected network, can be expressed as

$$\mathbf{A i} = 0 \tag{5.219}$$

A can be partitioned as $\left[\mathbf{A}_{\mathrm{T}} \mid \mathbf{A}_{\mathrm{L}}\right]$ where the columns of $\mathbf{A}_{\mathrm{T}}$ correspond to the tree branches of a chosen tree T, and the columns of $\mathbf{A}_{\mathrm{L}}$ correspond to the links.

Loop matrix

For a directed graph with b branches and n_l oriented loops, the loop matrix is a $n_l \times b$ matrix $\mathbf{B}_\mathbf{a} = \left\lfloor b_{ij} \right\rfloor$ where:

- $b_{ij} = 1$ if branch j is in loop i, and their directions agree,
- $b_{ij} = -1$ if branch j is in loop i, and their directions oppose, and
- $b_{ij} = 0$ if branch j is not in loop i.

For example in Fig. 5.77 there are seven loops. The loop matrix is:

$$
\begin{array}{c}
\ \ \ a\ \ \ \ b\ \ \ \ c\ \ \ \ d\ \ \ \ e\ \ \ \ f\quad branches\ in\ the\ loop \\
\begin{array}{cc}
loop & 1 \\
& 2 \\
& 3 \\
B_a = & 4 \\
& 5 \\
& 6 \\
& 7
\end{array}
\left[
\begin{array}{cccccc}
1 & 0 & -1 & 0 & 1 & 0 \\
0 & 1 & 0 & -1 & -1 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 \\
1 & 1 & -1 & -1 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 1 \\
1 & 0 & 0 & 1 & 1 & 1 \\
0 & 1 & 1 & 0 & -1 & 1
\end{array}
\right]
\begin{array}{l}
aec \\
bde \\
abf \\
abdc \\
cdf \\
aedf \\
bfce
\end{array}
\end{array}
$$

$$\text{(5.220)}$$

The KVL for all loops can be expressed in matrix-vector notation as
$$\mathbf{B}_\mathbf{a} v = 0 \qquad \text{(5.221)}$$

Any sub-matrix $\mathbf{B}_\mathbf{b}$ consisting of the maximum number of independent rows of $\mathbf{B}_\mathbf{a}$ is called a basic loop matrix. Thus, the $b-n+1$ independent KVL equations may be expressed as
$$\mathbf{B}_\mathbf{b} v = 0 \qquad \text{(5.222)}$$

A systematic method of constructing a basic loop matrix is through the aid of a tree T. Each link of the associated co-tree, together with the unique path through T, forms a loop, called the fundamental loop for that link. A sub-matrix of $\mathbf{B}_\mathbf{a}$ constructed by the use of $b-n+1$ links is called a fundamental loop matrix $\mathbf{B}_f$.

For example, in Fig. 5.77 a tree T consists of the branches (abc). The corresponding fundamental loop matrix is

$$\begin{array}{c} \\ d \\ \mathbf{B}_f = e \\ f \end{array} \begin{array}{cccccc} a & b & c & d & e & f \\ \left[\begin{array}{ccc|ccc} -1 & -1 & 1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right] & \begin{array}{c} dbac \\ eca \\ fab \end{array} \end{array}$$

$$(5.223)$$

It is obvious that any fundamental loop matrix can be partitioned as:

$$\mathbf{B}_f = \begin{bmatrix} \mathbf{B}_T & \vdots & \mathbf{1} \end{bmatrix} \qquad . \tag{5.224}$$

Cut-set matrix

For a directed graph with b branches and n_C oriented cut-sets, the cut-set matrix is a $n_c \times b$ matrix $\mathbf{D}_a = \begin{bmatrix} d_{ij} \end{bmatrix}$ where:

- $d_{ij} = 1$ if branch j is in cut-set i, and their directions agree,
- $d_{ij} = -1$ if branch j is in cut-set i, and their directions oppose, and
- $d_{ij} = 0$ if branch j is not in cut-set i.

For example, there are seven cut-sets in the graph of Fig. 5.77. The cut-set matrix is

$$\begin{array}{c} cutset \\ \\ \mathbf{D}_a = \end{array} \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} \begin{array}{cccccc} a & b & c & d & e & f \\ \left[\begin{array}{cccccc} 1 & 0 & 1 & 0 & 0 & -1 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & -1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & -1 \\ 1 & 0 & 0 & 1 & -1 & -1 \\ -1 & 1 & -1 & 1 & 0 & 0 \end{array}\right] & \begin{array}{c} \textit{branches in the cutset} \\ acf \\ abe \\ cde \\ bdf \\ bcef \\ adef \\ abcd \end{array} \end{array}$$

$$(5.225)$$

The general form of the KCL in matrix-vector notation is:

$$\mathbf{D}_a \mathbf{i} = 0 \qquad . \tag{5.226}$$

Any sub-matrix $\mathbf{D}_b$ of $\mathbf{D}_a$ consisting of the maximum number of independent rows of $\mathbf{D}_a$ is called a basic cut-set matrix. A systematic method for constructing a basic cut-set matrix is through the aid of a tree T. Each tree branch of T, together with some (possibly no) links in the associated co-tree forms a cut-set, called a fundamental cut-set for that tree branch. A sub-matrix $\mathbf{D}_f$ constructed with the $n-1$ fundamental cut-sets is called a fundamental cut-set matrix.

For example, for the graph of Fig. 5.77 with tree branches (abc), the fundamental cut-set matrix is

$$
D_f = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array}
\begin{array}{cccccc} a & b & c & d & e & f \end{array}
\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & -1 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 & 1 & 0 \end{array}\right]
\begin{array}{c} adef \\ bdf \\ cde \end{array}
\tag{5.227}
$$

It is obvious that any fundamental cut-set matrix can be partitioned as:

$$
\mathbf{D}_r = \begin{bmatrix} \mathbf{1} & \vdots & \mathbf{D}_\mathbf{L} \end{bmatrix} \tag{5.228}
$$

Relationship between branch variables

When the reduced incidence matrix $\mathbf{A}$, the fundamental loop matrix $\mathbf{B}$ and the fundamental cut-set matrix $\mathbf{D}$ are built corresponding to the same tree T, the matrices can be partitioned by:

$$
\mathbf{A} = \begin{bmatrix} \mathbf{A}_\mathbf{T} & \vdots & \mathbf{A}_\mathbf{L} \end{bmatrix} , \tag{5.229}
$$

$$
\mathbf{B} = \begin{bmatrix} \mathbf{B}_\mathbf{T} & \vdots & \mathbf{1} \end{bmatrix} , \tag{5.230}
$$

$$
\mathbf{D} = \begin{bmatrix} \mathbf{1} & \vdots & \mathbf{D}_\mathbf{L} \end{bmatrix} , \tag{5.231}
$$

$$
\mathbf{v} = \begin{bmatrix} \mathbf{v}_\mathbf{T} \\ \hline \mathbf{v}_\mathbf{L} \end{bmatrix} , \tag{5.232}
$$

$$
\mathbf{i} = \begin{bmatrix} \mathbf{i}_\mathbf{T} \\ \hline \mathbf{i}_\mathbf{L} \end{bmatrix} , \tag{5.233}
$$

and the following relationships can be given:

$$
\mathbf{v}_\mathbf{L} = -\mathbf{B}_\mathbf{T}\mathbf{v}_\mathbf{T} \quad \text{(loop transformation)}, \tag{5.234}
$$

$$
\mathbf{i}_\mathbf{T} = -\mathbf{D}_\mathbf{L}\mathbf{i}_\mathbf{L} \quad \text{(cut-set transformation)}, \tag{5.235}
$$

$$
\mathbf{v} = \mathbf{A}^T\mathbf{v}_{\text{nodal}} \quad \text{(node transformation)}, \tag{5.236}
$$

$$
\mathbf{D}_\mathbf{L} = -\mathbf{B}_\mathbf{T}^T \tag{5.237}
$$

5.9.4 Circuit analysis

Linear and time invariant networks with lumped parameters can be described by various techniques. The equations are the KVL, the KCL and the branch current-voltage relationships (BCVR). The known quantities are the currents of independent current sources, the voltages of independent voltage sources and the impedances of the branches.

5.9.4.1   Tableau analysis The unknowns in a tableau analysis are the branch currents and the branch voltages. The system of equations consists

of the KCL for each node of the circuit, the KVL for each loop of the circuit and the BCVR for each branch of the circuit.
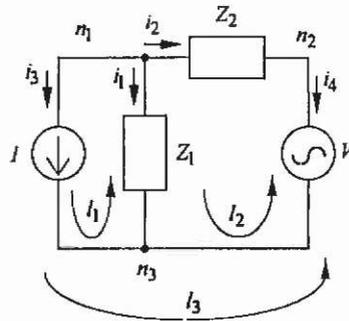

Fig. 5.78. Example circuit.

For example, consider the circuit of Fig. 5.78. The system of equations is given by:

$$
\begin{array}{llll}
\text{KCL}(n_1) & \left\{\begin{array}{l}\end{array}\right. & i_1 & +i_2 & +i_3 & & & & & & = 0 \\
\text{KCL}(n_2) & & & -i_2 & & +i_4 & & & & & = 0 \\
\text{KCL}(n_3) & & -i_1 & & -i_3 & -i_4 & & & & & = 0 \\
\hline
\text{KVL}(l_1) & \left\}\begin{array}{l}\end{array}\right. & & & & & v_1 & & -v_3 & & = 0 \\
\text{KVL}(l_2) & & & & & & -v_1 & +v_2 & & +v_4 & = 0 \\
\text{KVL}(l_3) & & & & & & & v_2 & -v_3 & +v_4 & = 0 \\
\hline
\text{BCVR}(1) & & -Z_1 . i_1 & & & & +v_1 & & & & = 0 \\
\text{BCVR}(2) & & & -Z_2 . i_2 & & & & +v_2 & & & = 0 \\
\text{BCVR}(3) & & & & i_3 & & & & & & = I \\
\text{BCVR}(4) & & & & & & & & & v_4 & = V
\end{array}
$$

$$(5.238)$$

It is obvious that the number of equations is larger than the number of unknowns. There are linear dependent equations in the system. The matrix is sparse, singular and not symmetric.

5.9.4.2   Modified nodal analysis The unknowns in a modified nodal analysis (MNA) are the nodal voltages and the currents through the independent voltage sources. For each node per connected circuit component, except the chosen reference node, the KCL is written.

Applying the BCVR to each branch immediately eliminates the currents. The branch voltages are expressed with respect to the nodal voltages. For each independent voltage source, an extra unknown current is added to the set of unknowns. An extra equation, representing the

difference between the two nodal voltages of the voltage source to the known voltage drop, is added.

For the circuit of Fig. 5.78, this approach results in:

$$\begin{bmatrix} Y_1 + Y_2 & -Y_2 & 0 \\ -Y_2 & Y_2 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} v_{n_1} \\ v_{n_2} \\ i_4 \end{bmatrix} = \begin{bmatrix} -I \\ 0 \\ V \end{bmatrix} \qquad (5.239)$$

The system of equations is sparse and smaller when compared to the tableau analysis. There are no superfluous equations. The unknowns are the nodal voltages and branch currents.

5.9.4.3 Compacted modified nodal analysis In the compacted form of the MNA, the unknown currents through the independent voltage sources are eliminated by substitution of the extra equations concerning the independent voltage sources.

$$[Y_1 + Y_2] \cdot [v_n] = [-I + Y_2 V] \qquad (5.240)$$

The system of equations is again reduced, moreover, there are no zero diagonal elements left in the matrix. The matrix is symmetric.

5.9.5 Topological methods
A topological method for circuit analysis is a technique deriving parameters describing the circuit behaviour from the structure of a graph, associated with the network. Some topological methods are the Signal Flow Graph method and the tree-enumeration method.

5.9.5.1 Signal flow graph A signal flow graph (SFG) is a weighted directed graph representing a system of linear equations. The nodes represent the variables; the branch weights represent the coefficients in the relations between the variables. A node variable and the sum of the incoming branch weights, multiplied by the node variable from which the incoming branch originates, represent an equation. Consider for example the SFG of Fig. 5.79 and the corresponding system of equations:

$$x_1 = ax_0 + dx_2 \quad ,$$
$$x_2 = bx_0 + cx_1 + ex_2 \quad , \qquad (5.241)$$
$$x_3 = fx_1 + gx_2 \quad .$$

In a SFG, a node with outgoing branches only is called a source node. A node with some incoming branches is called a dependent node. A dependent node with incoming branches only is called a sink node. In Fig. 5.79, x_0 is a source node, x_1, x_2 and x_3 are dependent nodes and x_3 is a sink node.
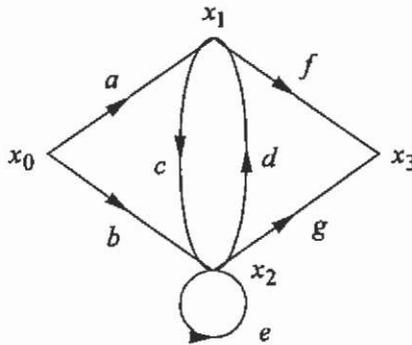
Fig. 5.79. Signal flow graph.

5.9.5.2   Formulation of the signal flow graph It is assumed that the voltage sources do not contain loops, and that the current sources do not contain cut-sets. It is possible to select a tree T such that all voltage sources v_{Σ} are tree branches, and all current sources i_J are links. The impedance tree branches are characterised by the impedance matrix Z_T and the immittance co-tree branches are characterised by the admittance matrix Y_L. The following steps can construct a SFG:

1. Apply KVL to express each element of v_L in terms of elements of v_{Σ} and v_T.

2. Apply KCL to express each element of i_T in terms of elements of i_J and i_L.

3. For impedance tree branches, each voltage is expressed in terms of the current through the branch: $v_T = Z_T i_T$.

4. For immittance co-tree branches, each current is expressed in terms of the voltage across the branch: $i_L = Y_L v_L$.

A SFG formed in this way displays the KCL, KVL and BCRV relations in their most primitive way. As a consequence, it is called a primitive signal-flow graph. The primitive SFG of the circuit from Fig. 5.79 is represented in Fig. 5.80.

The number of nodes in the SFG can very easily be largely reduced by the use of a compacted signal flow graph, which is obtained from the primitive signal flow graph by eliminating all sink nodes and the variables i_T and v_L. The compacted SFG of the circuit of Fig. 5.78 is represented in Fig. 5.81.
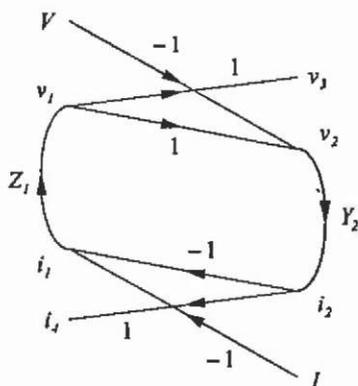
Fig. 5.80. Primitive signal flow graph.
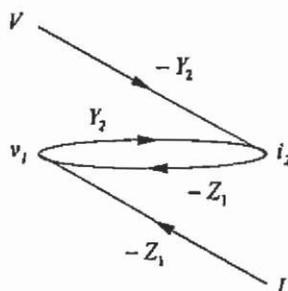
Fig. 5.81. Compacted signal flow graph.

5.9.6 Circuit analysis for the coupled field-external circuit system

The solid conductors are considered as an admittance branch and a controlled current source connected in parallel:

$$\mathbf{I}_{sol} = \mathbf{G}_{sol}\mathbf{V}_{sol} - j\omega\ell\mathbf{Q}_{sol}^{T}\mathbf{A} \quad . \tag{5.242}$$

A is the vector of the magnetic vector potentials and $j\omega\ell\mathbf{Q}_{sol}^{T}$ is the matrix with the coupling terms of (5.215). The stranded conductors are considered an impedance branch and a dependent voltage source connected in series:

$$\mathbf{V}_{str} = \mathbf{R}_{str}\mathbf{I}_{str} + j\omega\ell\mathbf{P}_{str}^{T}\mathbf{A} \tag{5.243}$$

where $j\omega\ell\mathbf{P}_{str}^{T}$ is the matrix with the coupling terms obtained by the discrete integration of (5.214).

5.9.6.1 Modified and compacted modified nodal analysis The tableau analysis is not suited for numerical calculation because the dependent equations result in a singular system matrix.

In an MNA, for each connected sub-circuit a reference node is chosen. On all other nodes an unknown voltage is defined. For each independent voltage source and for each stranded conductor, an unknown current is defined. For each node, the KCL is written. The currents are immediately written in terms of the nodal voltages by means of the BCVR. In the case of a stranded conductor branch or an independent voltage source branch, the current is unknown. For each of these branches, an extra KVL equation describes the nodal voltages of the endpoints in terms of the unknown currents and the magnetic vector potentials.
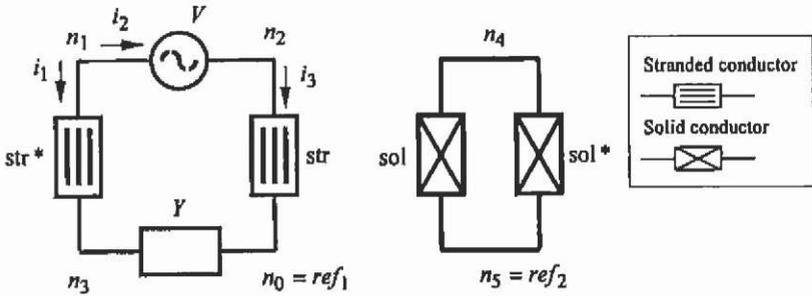
Fig. 5.82. Electric circuit with stranded and solid conductors.

This method has been applied to the circuit in Fig. 5.82.

$$
\begin{array}{ll}
\text{KCL}(n_1) & \quad i_1 + i_2 = 0 \\
\text{KCL}(n_2) & \quad -i_2 + i_3 = 0 \\
\text{KCL}(n_3) & \quad Y v_{n_3} \quad -i_1 = 0 \\
\text{KCL}(n_4) & \quad (Y_{sol} + Y_{sol*}) v_{n_4} \quad -j\omega\ell\left(Q_{sol}^T + Q_{sol*}^T\right) A = 0 \\
\hline
\text{KVL}(str*) & \quad v_{n_1} \quad -v_{n_3} \quad -Z_{str*} i_1 \quad -j\omega\ell P_{str*}^T A = 0 \\
\text{KVL}(V) & \quad v_{n_1} -v_{n_2} = V \\
\text{KVL}(str) & \quad v_{n_2} \quad -Z_{str} i_3 \quad -j\omega\ell P_{str}^T A = 0
\end{array}
$$

$$(5.244)$$

The factor χ can symmetrise the coupled magnetic-electric system of equations for a 2D time-harmonic problem.

The presents of zero elements in the diagonal of the matrix makes the choice of an appropriate method for solving the system of equations more difficult.

As with the CMNA, the extra KVL equations can be substituted in both the KCL equations and the field equations, tending to a smaller and symmetric matrix with a fully occupied diagonal. However, the

implementation of the dense circuit equations into the sparse system of field equations makes the solution procedure difficult.

The major problem of the nodal approach in the circuit analysis is that the description of the circuit is based on the voltage only. A more suitable description would be based on a hybrid use of both unknown voltages and currents. Here, the difficulty is the systematic determination of the circuit components described by currents, voltages and the interface parts in the circuit.

5.9.6.2 Signal flow graph for coupled magnetic-electric problems

The circuit in Fig. 5.83 is considered. Branches are assembled in the tree with preference:

- voltage sources, solid conductors, impedances and stranded conductors.

The preferred order for links is:

- current sources, stranded conductors, admittances and solid conductors (Fig. 5.83).

The fundamental cut-set matrix and the fundamental loop matrix are partitioned in components. They are associated with the stranded conductors being links (*str*), tree branches (*str**), solid conductors being tree branches (*sol*), links (*sol**), independent sources (*i* and *v*) and the immittance tree branches (*T*) and links (*L*).

Applying the Kirchhoff current law (KCL) for each fundamental cut-set and the Kirchhoff voltage law (KVL) for each fundamental loop arranges a SFG. The unknowns of the system are the link currents and the tree branch voltages. The SFG of a stranded and a solid conductor is shown in Fig. 5.84. The Signal Flow sub-graphs of the circuit of Fig. 5.83 are represented in Fig. 5.85. Table 5.7 shows the equivalencies between the SFG and the matrix calculus.

Table 5.7. Equivalence between circuit theory, SFG and matrix calculus.

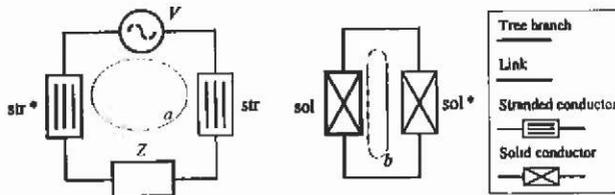| circuit analysis	signal flow graph	matrix notation
Kirchhoff current law	current nodes	$\mathbf{DI} = 0$
Kirchhoff voltage law	voltage nodes	$\mathbf{BV} = 0$
branch relations	vertical connections	$\mathbf{V} = \mathbf{ZI}$; $\mathbf{I} = \mathbf{YV}$
cut-set transformation	eliminate $\mathbf{V}_{str^*}, \mathbf{I}_{str^*}$	$\mathbf{I}_{str} = -\mathbf{D}_{str,l}\mathbf{I}_l - \mathbf{D}_{str,w}\mathbf{I}_w$
loop transformation	eliminate $\mathbf{V}_{sol^*}, \mathbf{I}_{sol^*}$	$\mathbf{V}_{sol^*} = -\mathbf{B}_{sol,v}\mathbf{V}_v - \mathbf{B}_{sol,sol}\mathbf{V}_{sol}$
compacting	eliminate $\mathbf{I}_T, \mathbf{V}_L$	$\mathbf{V}_L = \mathbf{ZI}_L$; $\mathbf{I}_T = \mathbf{YV}_T$

Fig. 5.83. Electric circuit with a) stranded and b) solid conductors.
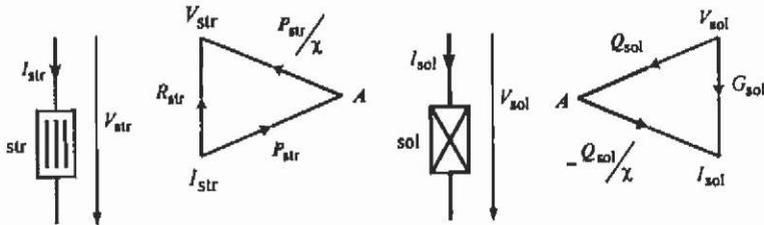

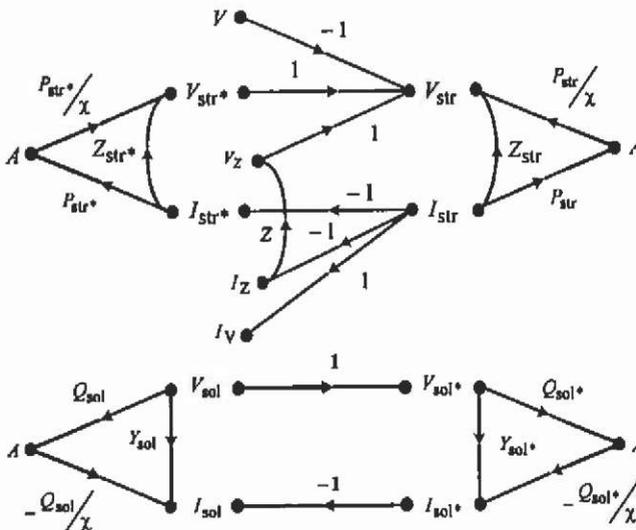Fig. 5.84. SFG of a a) stranded and b) solid conductor.

Fig. 5.85. Non-coupled signal flow graph.

Joining the graphs does not change the graph nodes so long as the dependent nodes of the first graph correspond to source nodes of the other graph and vice versa. Therefore, branch current-voltage relations (BCVR) are added either as impedances or admittances (Fig. 5.85). Two dependent nodes are joined together to one zero node by changing the sign of all incoming branch weights of one of the sub-graphs. The former

unknown of that node disappears. This happens for the stranded conductor links and the solid conductor tree branches (Fig. 5.86).
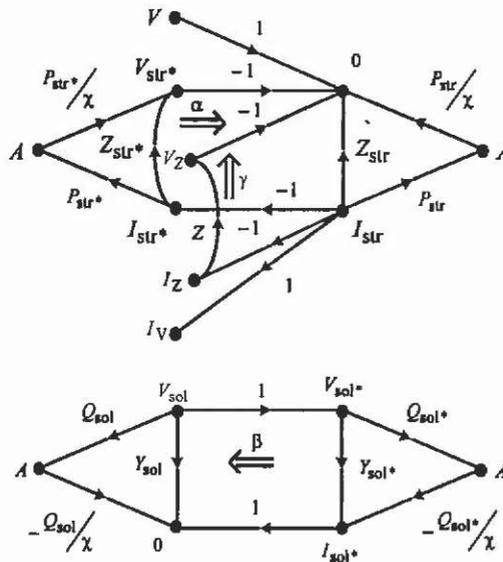
Fig. 5.86. Coupled signal flow graph.

Combining magnetically coupled branches causes three difficulties:

- Stranded conductor tree branches are described by a dependent current (Fig. 5.85a).
- Solid conductor links are described by a dependent voltage (Fig. 5.85b).
- The coupling terms are not symmetric.

Three operations can solve the above mentioned problems:

1. Partial cut-set transformation

The preferences while choosing tree branches result in a fundamental cut-set associated with a stranded conductor tree branch which contains only current sources and stranded conductors. A partial cut-set transformation

$$\mathbf{I}_{str} = -\mathbf{D}_{str,i}\mathbf{I} - \mathbf{D}_{str,str}\mathbf{I}_{str} \tag{5.245}$$

contracts the graph in direction α (Fig. 5.86). The current of the stranded conductor tree branch is expressed as a combination of independent currents and other stranded conductor currents (Table 5.7).

2. Partial loop transformation

A fundamental loop associated with a solid conductor link only exists of voltage sources and solid conductors. A partial loop transformation

$$\mathbf{V}_{sol^*} = -\mathbf{B}_{sol^*,v} \cdot \mathbf{V} - \mathbf{B}_{sol^*,sol} \cdot \mathbf{V}_{sol} \tag{5.246}$$

contracts the graph in direction β (Fig. 5.86). The voltage of the solid conductor link is expressed as a combination of independent voltages and other solid conductor voltages.

3. Symmetrising the system

A contraction of the BCVR (direction γ in Fig. 5.86, Table 5.7) leads to a compact signal flow graph (CSFG) (Fig. 5.87).
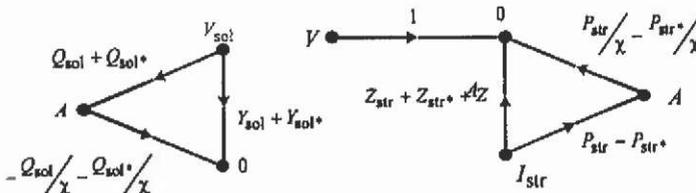
Fig. 5.87. Compact signal flow graph.

From the CSFG, the coupling matrices are extracted in a simple way. The unknown graph nodes become system unknowns. The dependent graph nodes represent matrix equations. The coupling terms are kept symmetric. Compared to tableau analysis, MNA and CMNA, a reduction of additional circuit equations is obtained. Multiplying the circuit loop equations with χ and the circuit cut-set equations with $-\chi$ leads to the coupled field-circuit matrix

$$\begin{bmatrix} \mathbf{K} & -\mathbf{F} \\ -\mathbf{F}^{\mathrm{T}} & \chi\mathbf{S} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} = \begin{bmatrix} \mathbf{H} \\ \chi\mathbf{W} \end{bmatrix} \tag{5.247}$$

where

$$\mathbf{S} = \begin{bmatrix} -\mathbf{R}^*_{str} & 0 & -\mathbf{B}_{str,sol} & -\mathbf{B}_{str,T} \\ 0 & -\mathbf{Z}_L & -\mathbf{B}_{L,sol} & -\mathbf{B}_{LT} \\ \mathbf{D}_{sol,str} & \mathbf{D}_{sol,L} & \mathbf{G}^*_{sol} & 0 \\ \mathbf{D}_{T,str} & \mathbf{D}_{T,L} & 0 & \mathbf{Y}_T \end{bmatrix} \tag{5.248}$$

with

$$\mathbf{R}^*_{str} = \mathbf{R}_{str} - \mathbf{B}_{str,str^*} \mathbf{R}_{str^*} \mathbf{D}_{str^*,str} \tag{5.249}$$

$$\mathbf{G}^*_{sol} = \mathbf{G}_{sol} - \mathbf{D}_{sol,sol^*} \mathbf{G}_{sol^*} \mathbf{B}_{sol^*,sol} \tag{5.250}$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{I}_{str} & \mathbf{I}_L & \mathbf{V}_{sol} & \mathbf{V}_T \end{bmatrix}^{\mathrm{T}} \tag{5.251}$$

$$\mathbf{H} = -\mathbf{P}_{sr} \cdot \mathbf{D}_{sr^*,i} \mathbf{I}_i - \mathbf{Q}_{sol} \cdot \mathbf{B}_{sol^*,v} \mathbf{V}_v$$

$$\mathbf{F} = \begin{bmatrix} \mathbf{P}_{sr} - \mathbf{P}_{sr} \cdot \mathbf{D}_{sr^*,sr} & 0 & \mathbf{Q}_{sol} - \mathbf{Q}_{sol} \cdot \mathbf{B}_{sol^*,sol} & 0 \end{bmatrix} \tag{5.252}$$

and

$$\mathbf{W} = \begin{bmatrix} \mathbf{B}_{sr,v} \mathbf{V}_v - \mathbf{B}_{sr,sr^*} \cdot \mathbf{R}_{sr^*} \cdot \mathbf{D}_{sr^*,i} \mathbf{I}_i \\ \mathbf{B}_{L,v} \mathbf{V}_v \\ -\mathbf{D}_{sol,i} \mathbf{I}_i + \mathbf{D}_{sol,sol^*} \mathbf{G}_{sol^*} \cdot \mathbf{B}_{sol^*,v} \mathbf{V}_v \\ -\mathbf{D}_{T,i} \mathbf{I}_i \end{bmatrix} \tag{5.253}$$

In the case of a quasi-static problem, $\mathbf{K}$ is complex symmetric. $\mathbf{S}$ is symmetric because $\mathbf{B}_{x,y} = -\mathbf{D}_{y,x}^T$ and that $\mathbf{R}_{sr}$, $\mathbf{R}_{sr^*}$, $\mathbf{G}_{sol}$, $\mathbf{G}_{sol^*}$, $\mathbf{Z}_L$ and $\mathbf{Y}_r$ are diagonal matrices.

5.9.7 Solution of the system of coupled equations

Particular attention must be paid to the solution procedure of the coupled system of FEM and circuit equations. The matrix obtained for a 2D time-harmonic solution coupled with an electric circuit described with the proposed method is complex, symmetric, has no zero diagonal elements but is not hermitian.

The FEM block is positive definite and the circuit coupling block is negative definite. Therefore, the conjugate gradient (CG) method can not be used. Other suggestions are the bi-conjugate gradient (BiCG) method, the conjugate gradient method on the normal equations (CGN), other orthogonal Krylov-subspace methods and block elimination schemes (BES).

5.9.7.1 Conjugate gradient on the normalised equations Instead of solving the system

$$\mathbf{AX} = \mathbf{B} \quad , \tag{5.254}$$

the system

$$\mathbf{A}^* \mathbf{AX} = \mathbf{A}^* \mathbf{B} \tag{5.255}$$

is solved. The matrix is hermitian.

5.9.7.2 Block elimination schemes Due to the fact that the different parts of the matrix have different properties, it is advantageous to split up the matrix into several blocks. By matrix calculus, some block elimination schemes can be derived. The kind of elimination is dependent on the criterion to divide the matrix.

The system of equations can be partitioned as:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{B}^T & \mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathbf{R} \\ \mathbf{S} \end{bmatrix} \tag{5.256}$$

where $\mathbf{A}$ is a sparse matrix with the same dimension as the mesh, $\mathbf{C}$ is a full matrix with all circuit equations and $\mathbf{B}$ is a full matrix containing the coupling terms between the FEM magnetic field description and the circuit equations.

The system can be written as:

$$\left(\mathbf{C} - \mathbf{B}^T \mathbf{A}^{-1} \mathbf{B}\right)\mathbf{Y} = \mathbf{S} - \mathbf{B}^T \mathbf{A}^{-1} \mathbf{R} \tag{5.257}$$

$$\left(\mathbf{C} - \mathbf{B}^T \mathbf{P}\right)\mathbf{Y} = \mathbf{S} - \mathbf{B}^T \mathbf{Q} \tag{5.258}$$

The vector $\mathbf{Q}$ and each column of $\mathbf{P}$ can be calculated by

$$\mathbf{A}\mathbf{Q} = \mathbf{R} \tag{5.259}$$

$$\mathbf{A}\mathbf{P}_i = \mathbf{B}_i \tag{5.260}$$

using an iterative equation solver for sparse systems. The number of systems that has to be solved in this manner equals the amount of network unknowns plus one. This is an important reason for describing the electric circuit with as few unknowns as possible.

Another method separates the complex equations (these are the equations in the solid conductor regions and the electric circuit equations) from $\mathbf{A}$. This results in dividing the system matrix in a real and a complex part.

$$\begin{bmatrix} \mathbf{A} & | & \mathbf{B} \\ \hline \mathbf{B}^T & | & \mathbf{C} + j\mathbf{D} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{X} + j\mathbf{M} \\ \mathbf{Y} + j\mathbf{N} \end{bmatrix} = \begin{bmatrix} \mathbf{R} + j\mathbf{V} \\ \mathbf{S} + j\mathbf{W} \end{bmatrix} \tag{5.261}$$

The system of equations can be written as:

$$\mathbf{A}\mathbf{X} = \mathbf{R} - \mathbf{B}\mathbf{Y}$$

$$\mathbf{A}\mathbf{M} = \mathbf{V} - \mathbf{B}\mathbf{N} \tag{5.262}$$

$$\left(\mathbf{C} + j\mathbf{D}\right)\cdot\left(\mathbf{Y} + j\mathbf{N}\right) = \mathbf{S} + j\mathbf{W} - \mathbf{B}^T\left(\mathbf{X} + j\mathbf{M}\right) \tag{5.263}$$

The system can be solved with the unknowns $\mathbf{X}, \mathbf{M}, \mathbf{Y}$ and $\mathbf{N}$ by iteration.

5.9.8   Two examples

In the first example, eddy currents are induced in a conducting plane, passing between two symmetric inductors (Fig. 5.88). The external circuit is shown in Fig. 5.82. In the second example an induction machine at start-up is analysed (Fig. 5.89). Stator windings and rotor bars are connected as shown in Fig. 5.90. The numbers of additional circuit equations for the different methods are shown in Table 5.8. The structure of the matrix is shown in (Fig. 5.90).

Table 5.8. Number of circuit equations.

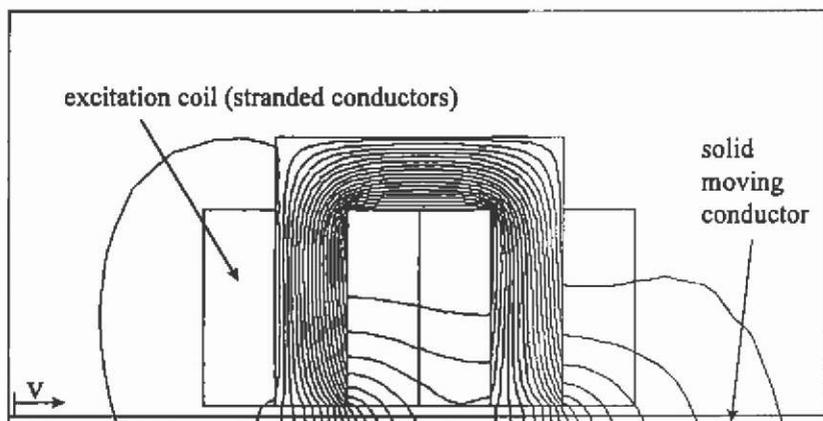| circuit analysis	tableau	MNA	CMNA	SFG
inductor	13	5	4	2
induction motor	166	46	43	31

Fig. 5.88. Plot of the equipotential lines of an inductor and a conducting plane moving at 10 m/s to the right side.
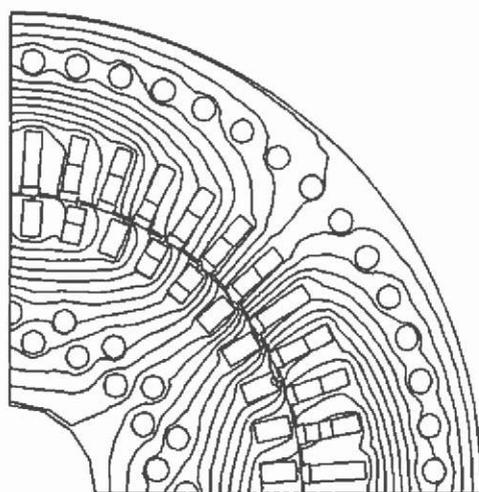

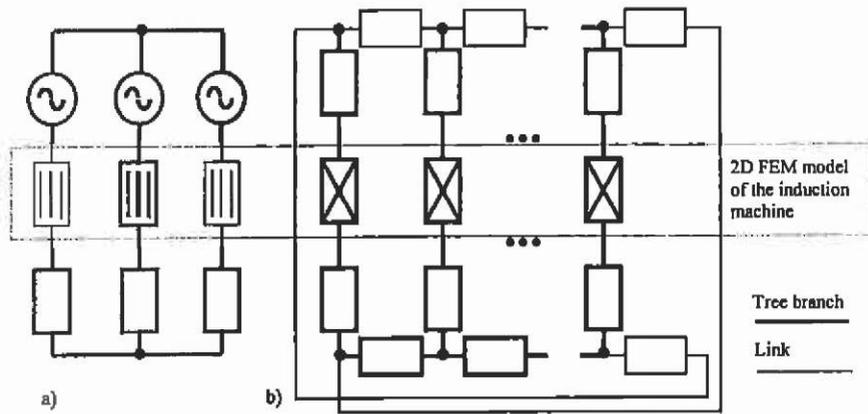Fig. 5.89. Equipotential plot of a 50 Hz time-harmonic solution for an induction motor.

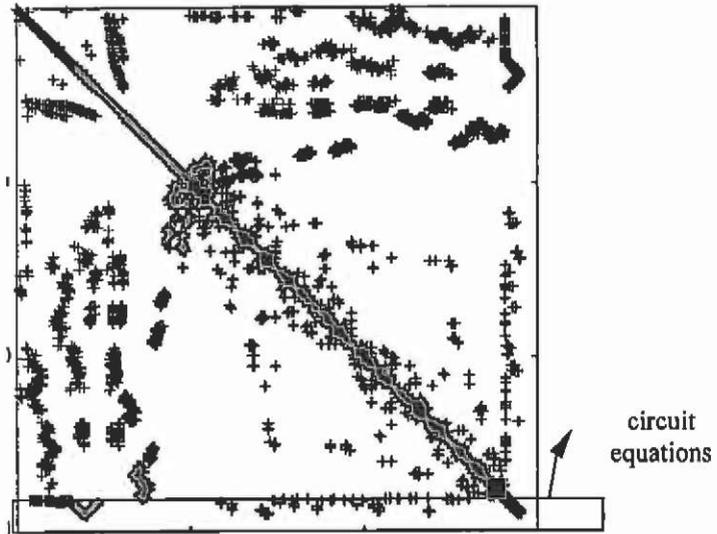Fig. 5.90. Electric circuit of a) the stator and b) the rotor of an induction motor.

Fig. 5.91. Matrix structure of a 50 Hz time-harmonic solution of the induction motor example.

5.9.9 Transient electromagnetic problem

The partial differential equation for a two-dimensional transient magnetic problem is:

$$\nabla \cdot (\nu \nabla A) - \sigma \frac{\partial}{\partial t} A = -J_0 \tag{5.264}$$

To stay in the same notation as in section 5.7 using the relative reluctivity ν, we can write:

$$\nabla \cdot (\nu_r \nabla A) - \frac{\sigma}{\nu_0} \frac{\partial}{\partial t} A = -\frac{J_0}{\nu_0} \tag{5.265}$$

and the functional is:

$$F(A) = \frac{1}{2} \int_\Omega \left(\nu_r |\nabla A|^2 - 2 \frac{J_0 A}{\nu_0} + 2 \frac{\sigma}{\nu_0} A \frac{\partial A}{\partial t} \right) d\Omega. \tag{5.266}$$

5.9.9.1 Functional within an element

The functional within an element becomes

$$F(A) = \frac{1}{2} \int_{\Delta_e} \left(\nu_r^{(e)} |\nabla A|^2 - 2 \frac{J_0^{(e)} A}{\nu_0} + 2 \frac{\sigma^{(e)}}{\nu_0} A \frac{\partial A}{\partial t} \right) d\Omega. \tag{5.267}$$

Evaluating the third term yields:

$$\int_{\Delta_e} \frac{\sigma^{(e)}}{\nu_0} A \frac{\partial A}{\partial t} d\Omega = \sum_{i=1}^{3} \sum_{j=1}^{3} A_i \cdot \underbrace{\frac{\sigma^{(e)}}{\nu_0} \int_{\Delta_e} N_i N_j \, d\Omega}_{R_{ij}^{(e)}} \cdot \frac{\partial A_j}{\partial t} \tag{5.268}$$

with the matrix entries $\mathbf{R}^{(e)}$:

$$R_{ij}^{(e)} = \begin{cases} \dfrac{\sigma}{\nu_0} \dfrac{\Delta_e}{6} & \forall \quad i = j \\[2mm] \dfrac{\sigma}{\nu_0} \dfrac{\Delta_e}{12} & \forall \quad i \neq j \end{cases} \tag{5.269}$$

5.9.9.2 Time stepping

The system of equations in matrix-vector notation can be written by:

$$\mathbf{K}\mathbf{A} + \mathbf{R} \frac{\partial \mathbf{A}}{\partial t} - \mathbf{T} = 0. \tag{5.270}$$

The time discretisation can be applied to the Galerkin approach in the time domain. The solution is only computed at discrete points in time, spaced in finite intervals Δt, the time-steps. First order shape functions are chosen for A and T as functions in time.

$$\begin{aligned} \mathbf{A}(t) &= \tau \mathbf{A}_k + (1-\tau)\mathbf{A}_{k-1} \\ \mathbf{T}(t) &= \tau \mathbf{T}_k + (1-\tau)\mathbf{T}_{k-1} \end{aligned} \tag{5.271}$$

with

$$\tau = \frac{t - t_{k-1}}{t_k - t_{k-1}} = \frac{t - t_{k-1}}{\Delta t} \tag{5.272}$$

and

$$\frac{\partial \mathbf{A}}{\partial t} = \frac{\partial \mathbf{A}}{\partial \tau}\frac{\partial \tau}{\partial t} = \frac{\mathbf{A}_k - \mathbf{A}_{k-1}}{\Delta t} \tag{5.273}$$

If the Galerkin approach is applied with τ as the weighting function then

$$\int_0^1 \tau\left\{ K\left(\tau\mathbf{A}_k + (1-\tau)\mathbf{A}_{k-1}\right) + \mathbf{R}\frac{\mathbf{A}_k - \mathbf{A}_{k-1}}{\Delta t} - \tau\mathbf{T}_k + (1-\tau)\mathbf{T}_{k-1} \right\} d\tau = 0, \tag{5.274}$$

$$\left(\frac{2K}{3} + \frac{\mathbf{R}}{\Delta t}\right)\mathbf{A}_k + \left(\frac{K}{3} - \frac{\mathbf{R}}{\Delta t}\right)\mathbf{A}_{k-1} - \left(\frac{2\mathbf{T}_k}{3} + \frac{\mathbf{T}_{k-1}}{3}\right) = 0. \tag{5.275}$$

This corresponds to a standard central difference formula. By introducing a more general set of weighting functions other difference schemes can be obtained. (5.275) can be generalised by using the parameter α.

$$\begin{aligned}\mathbf{A}(t) &= \alpha\tau\mathbf{A}_k + (1 - \alpha\tau)\mathbf{A}_{k-1} \\ \mathbf{T}(t) &= \alpha\tau\mathbf{T}_k + (1 - \alpha\tau)\mathbf{T}_{k-1}\end{aligned} \tag{5.276}$$

Equation (5.275) can be written in a general form:

$$\left(\alpha K + \frac{\mathbf{R}}{\Delta t}\right)\mathbf{A}_k + \left((1-\alpha)K - \frac{\mathbf{R}}{\Delta t}\right)\mathbf{A}_{k-1} - \left(\alpha\mathbf{T}_k + (1-\alpha)\mathbf{T}_{k-1}\right) = 0 \ . \tag{5.277}$$

With $\alpha = 0$ the forward difference Euler method is obtained. This approach is an explicit method because the term $\mathbf{KA}$ is evaluated at the beginning of time interval Δt.

$$\frac{\mathbf{R}}{\Delta t}\mathbf{A}_k = \left(\alpha\mathbf{T}_k + (1-\alpha)\mathbf{T}_{k-1}\right) - \left(K - \frac{\mathbf{R}}{\Delta t}\right)\mathbf{A}_{k-1} \tag{5.278}$$

$\alpha = 1$ gives the backward difference fully implicit method since the term $\mathbf{KA}$ is evaluated at the end of the time interval Δt.

$$\left(K + \frac{\mathbf{R}}{\Delta t}\right)\mathbf{A}_k = \left(\alpha\mathbf{T}_k + (1-\alpha)\mathbf{T}_{k-1}\right) + \frac{\mathbf{R}}{\Delta t}\mathbf{A}_{k-1} \tag{5.279}$$

The Crank-Nicolson scheme is obtained for $\alpha = 1/2$

$$\left(\frac{1}{2}K + \frac{\mathbf{R}}{\Delta t}\right)\mathbf{A}_k = \left(\alpha\mathbf{T}_k + (1-\alpha)\mathbf{T}_{k-1}\right) - \left(\frac{1}{2}K - \frac{\mathbf{R}}{\Delta t}\right)\mathbf{A}_{k-1} \tag{5.280}$$

and for $\alpha = 2/3$ the Galerkin scheme of (5.275) is obtained.

Particular attention must be paid to the choice of the value of α because it influences the stability and the accuracy of the numerical results.

5.9.9.3 **Stability and accuracy** Stability is important when time stepping is used for the solution of partial differential equations. By using the time-stepping formula eq.(5.277), the·scheme is unconditionally stable for $\alpha \geq 1/2$. The stability does not prevent oscillations, but guarantees that oscillations do not grow out of control. An oscillation-free scheme is the fully implicit method with $\alpha = 1$ while for all values $1/2 \leq \alpha < 1$ the implicit scheme oscillates if the time-step Δt is too large. The stability limit for $\alpha < 1/2$ has the form:

$$\Delta t \leq \gamma \frac{\rho c}{k} (\Delta L)^2 \tag{5.281}$$

with

$\frac{\rho c}{k}$ the diffusivity, ΔL the characteristic length of the elements and γ a numerical constant which depends on the elements used and on the choice of α.

The error $e_{\Delta t}$ in the approximation of the time derivative is for $\alpha = 1$ and $\alpha = 0$, i.e. for the fully implicit and the Euler explicit methods, of the order

$$e_{\Delta t} = O(\Delta t) \tag{5.282}$$

for $\alpha = 1/2$; i.e. for the Crank-Nicolson method it is of the order:

$$e_{\Delta t} = O((\Delta t)^2) \tag{5.283}$$

Therefore, the choice $\alpha = 1/2$ is advantageous, since it corresponds to an implicit method which is unconditionally stable and gives thus second order accuracy for the time integration. On the other hand, the value $\alpha = 1$ gives only first order accuracy for the time integration but completely avoids numerical oscillations even with large time steps. Thus a value of $\alpha = 1/2$ may be used, with a relatively small time step, in order to obtain an accurate solution. A value of $\alpha = 1$ can be used with a large time-step to obtain a less accurate solution to estimate the transient behaviour of the problem in principle.

5.9.9.4 **Slow motion** We will discuss motion problems with a static magnetic field and with a uniform moving Cartesian geometry at relatively low speed v:

$$\nabla \cdot (v_r \nabla A) + \mu_0 \sigma(\mathbf{v} \nabla A) = -\mu_0 \sigma \nabla V \tag{5.284}$$

Problems considering high speed with the solution of this A-formulation will be discussed.

The first term of the differential equation is called the diffusion term. The second term is called the convection term. The right-hand-side is called the load.

Unlike the previous formulations, the variational principle corresponding to the solution of the differential equation is not known. Therefore, a weighted residual method has been applied.

Galerkin approach

The multiplication of the differential equation with the weighting functions w_i yields:

$$-\int_\Omega \nabla(v_r \nabla A)w_i\, d\Omega - \int_\Omega \frac{\sigma}{V_0}\mathbf{v}\cdot\nabla A w_i\, d\Omega - \int_\Omega \frac{\sigma}{V_0}\nabla V w_i\, d\Omega = 0 \qquad . (5.285)$$

The magnetic vector potential A is written in terms of the basis functions N_j :

$$A = \sum_j N_j A_j \tag{5.286}$$

In a Galerkin approach the same functions are chosen as weighting functions. The first term is then written as:

$$-\int_\Omega \nabla(v_r \nabla A)w_i\, d\Omega = \sum_e \sum_{i=1}^{3}\sum_{j=1}^{3} v_r^{(e)}\underbrace{\int_{A_e}\nabla N_i\cdot\nabla N_j\, d\Omega A_j}_{\kappa_{ij}^{(e)}} \tag{5.287}$$

The third term can be evaluated by:

$$-\int_\Omega \frac{\sigma}{V_0}\nabla V w_i\, d\Omega = -\sum_e\sum_{j=1}^{3}\frac{\sigma}{V_0}\nabla V^{(e)}\underbrace{\int_{A_e} N_i\, d\Omega}_{T_i^{(e)}} \tag{5.288}$$

The second term is:

$$-\int_\Omega \frac{\sigma}{V_0}\mathbf{v}\cdot\nabla A w_i\, d\Omega = -\sum_e\sum_{i=1}^{3}\sum_{j=1}^{3}\frac{\sigma}{V_0}\underbrace{\int_{A_e}\mathbf{v}\cdot\nabla N_j N_i\, d\Omega A_j}_{M_{ij}^{(e)}} \tag{5.289}$$

The Galerkin approach and the variational technique give the same result for the diffusion term and the load term. In the case of linear basis functions, the integral in the convection term becomes:

$$M_{ij}^{(e)} = \frac{\sigma}{V_0}\frac{v_x b_j + v_y c_j}{6} \tag{5.290}$$

System of equations

The system of equations is given by:
$$(\mathbf{K} + \mathbf{M})\mathbf{A} = \mathbf{T} \tag{5.291}$$

This system is not symmetric because:
$$M_\mu^{(e)} \neq M_\nu^{(e)} \tag{5.292}$$

Appropriate iterative methods to solve the system of equations are minimal residual methods such as BiCG and GMRes.

GMRes has the problem that the memory requirement increases with the number of iteration steps. Therefore, restarted versions can be used. Petrov-Galerkin schemes such as BiCG are fast iterative equation solvers but have the disadvantage of breakdowns while solving the system.

External circuits

The current density in a moving conductor is calculated as
$$J = -\sigma \nabla V - \sigma \mathbf{v} \cdot \nabla A \tag{5.293}$$

In the case of a stranded conductor, the current density is assumed to be constant. The voltage drop over a stranded conductor with N_t turns moving at the speed $\mathbf{v}$ is

$$V_{str} = \frac{N_t \ell}{\sigma \Delta_{str}} I_{str} + \frac{N_t \ell}{\Delta_{str}} \int_\Omega \sigma \mathbf{v} \cdot \nabla A \, d\Omega \tag{5.294}$$

In the case of a solid conductor, the voltage is assumed to be constant. The current through the solid conductor is

$$I_{sol} = \frac{\sigma \Delta_{sol}}{\ell} V_{sol} - \int_\Omega \sigma \mathbf{v} \cdot \nabla A \, d\Omega \tag{5.295}$$

If the currents through one or more stranded conductors or the voltages through one or more solid conductors are unknown, extra circuit equations are needed to describe the full behaviour of the model. The choice of the circuit unknowns, the construction of the extra equations and the coupling terms is done by the signal flow graph methods already described for the case of time-harmonic magnetic fields. The non-moving conductors, however, have to be treated in a slightly different way. No induced effects appear in these conductors so long as the geometry does not change while moving.

Example

A conductive plane moves between two inductors. If the speed of the plane is zero, the solution corresponds to the static solution (Fig. 5.92). If motion is considered, the flux lines are pushed away in the direction of the motion (Fig. 5.93). If speed increases, the flux lines have less space to pass through the conducting plane. For very high speeds,

numerical problems result in unstable solutions (Fig. 5.94). In the figure, a separated closed flux line occurs. Up-winding schemes promise to surmount this difficulty. For further details please refer to the literature.
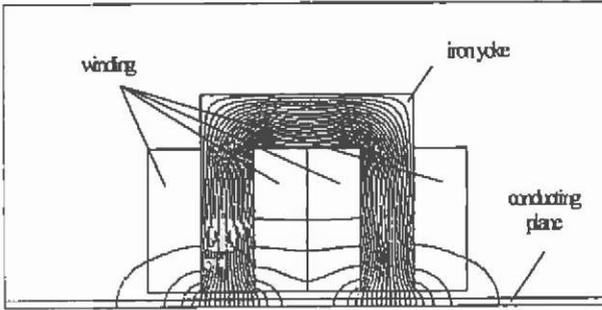
Fig. 5.92. Static solution of a non-moving conductive plane.
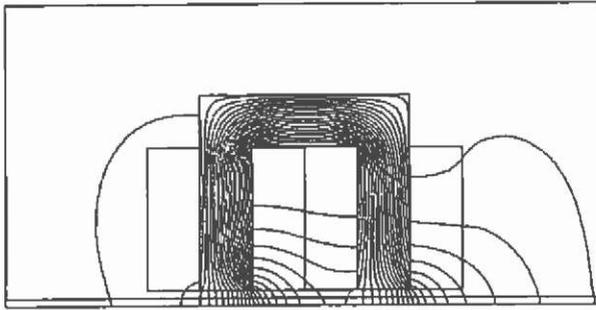
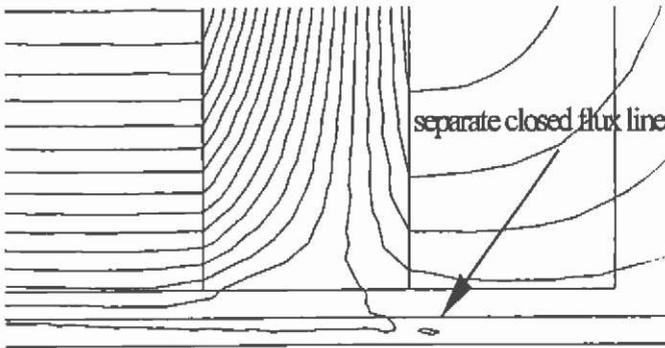Fig. 5.93. Static solution of a conductive plane moving at 10 m/s to the right side.

Fig. 5.94. Static solution of a conductive plane moving at high speed to the right side.

5.10 Post-processing

Using arbitrary potentials instead of physical quantities and the associated functionals in the formulation of the equations, raises the need for a closer look at the post-processing. The user of a FEM system desires to analyse a physical system in terms of field strength, energies, forces, densities etc. The potential itself does not necessarily have a physical meaning. In some cases, such as in the electrostatic and in the thermal analysis, the potential represents the electric potential and the temperature respectively (Table 5.9). Therefore, most of the interesting quantities in the post-process are numerically derived quantities. The type and order of the shape function of the potential over an element (linear, quadratic, etc.) and the element type (nodal, edge, etc.) determine the achievable relative accuracy of numerically derived values. The accuracy of the results is influenced by the discretisation and, related to it, the choice of the error estimator for an adaptive mesh refinement, if applied. Another difficulty arises in the calculation of lumped parameters (inductances, reactances, etc.), used in non-FEM analysis procedures, such as circuit analysis. Several different definitions of these quantities may exist, as for the inductance calculation of linear and non-linear energy transducers.

The aim of this chapter is to provide an overview of possible derived quantities, the necessary formulations and ways of influencing the accuracy of the results.

5.10.1   Potentials
As shown in the previous chapters, the chosen potentials for the different types of problems do not necessarily directly represent a physical quantity. The formulations for defining these potentials are chosen such that their application might impose simplifications in the formulation of the functionals or the choice of the gauges. A selection of problem types and the physical meaning of their potentials are collected in Table 5.9.

Table 5.9. Physical meaning of selected potentials.

Type of analysis	differential equation	type of potential	physical meaning
electrostatic	$\varepsilon \nabla^2 V = -\rho$	scalar	electric potential
magnetostatic	$\nu \nabla^2 A = -J_0$	vector	none, $\mathbf{B} = \nabla \times \mathbf{A}$
thermostatic	$\lambda \nabla^2 T = Q$	scalar	temperature
time-harmonic magnetic	$\nu \nabla^2 A - j\omega\sigma A = -J_0$	vector	induced currents related to A

The achievable accuracy of all derived values cannot be better than the accuracy of the computed potentials. The latter is determined by the choice of the element type, the shape function and the functionals used.

5.10.2   Energies

To calculate particular quantities for the elements of lumped parameter models such as inductivities, reactances, resistors etc, or to compute local forces and/or torques acting on bodies present in a magnetic field, energies are used to determine such quantities. There are various definitions.

5.10.2.1 Stored energy Energies are global quantities. It has already been discussed that in the finite element method, using the variational technique, an energy term, the functional, is minimised. This energy term does not necessarily have the meaning of a physical energy, for example a stored energy. Potentials are chosen in such a way that the minimisations of the related functional approximate the solution of the partial differential equation. In the case of a Laplace equation, the functional is:

$$F(A) = \frac{1}{2} \int_\Omega v_r |\nabla A|^2 \, d\Omega \qquad (5.296)$$

For a Cartesian, non-linear magnetostatic problem the stored magnetic energy can be calculated by:

$$W = \frac{\ell}{2\mu_0} \int_\Omega \left(\int_0^B v_r B \, dB \right) d\Omega \qquad (5.297)$$

or for linear materials, with the material reluctivity v_r:

$$W = \frac{\ell}{2\mu_0} \int_\Omega v_r B^2 \, d\Omega . \qquad (5.298)$$

Equations (5.296) and (5.297), (5.298) are similar. The total stored energy in the overall system, as an integration value is more accurate than any locally derived quantity.

Functionals for other differential equations can have additional terms in the integrand, but the similarity with an energy formulation is still maintained. The functional for the Poisson equation includes such an additional term:

$$F(A) = \frac{1}{2} \int_\Omega \left(v_r |\nabla A|^2 - 2 \frac{J_0 A}{v_0} \right) d\Omega \qquad (5.299)$$

The second term of the integrand can be thought of as being related to the energy input from the supply. In linear systems the energy input from the supply is calculated from the integral over the coil area by:

$$W = \frac{\ell}{2}\int_\Omega J_o A d\Omega \quad . \tag{5.300}$$

For linear systems, the stored energy equals the energy input from the supply. This only holds if all potentials of Dirichlet boundaries are set to zero, i.e. no additional flux is forced into the system at the boundaries. This last expression is useful even for non-linear problems for the calculation of flux linkages and inductances in special cases, when the behaviour with changing current is important. The functional used for problems with linear permanent magnets is:

$$F(A) = \frac{1}{2}\int_\Omega \left(v_r |\nabla A|^2 - 2\frac{J_o A}{v_0} - 2A\nabla(v_{rr}M) \right) d\Omega \quad . \tag{5.301}$$

Therefore, the third term in the integrand is related to the energy output from the permanent magnet to the system.

Most electrostatic problems are linear. The stored energy in an electrostatic model is:

$$W = \frac{\ell}{2}\int_\Omega \varepsilon_0 \varepsilon_r E^2 d\Omega \quad . \tag{5.302}$$

5.10.2.2 Co-energy Associated with the energy is the concept of co-energy (Cartesian geometry, magnetic problem):

$$W_{co} = \frac{\ell}{\mu_0}\int_\Omega v_r B^2 d\Omega - \frac{\ell}{2\mu_0}\int_\Omega \left(\int_0^B v_r B dB \right) d\Omega \quad . \tag{5.303}$$

This integral is effectively the "surface under the BH-curve". The co-energy is useful for force calculation. In problems with linear materials and no permanent magnets, the value of the co-energy and the energy are equal.

5.10.3 Local field quantities
As shown in Table 5.9, some local field quantities are directly represented by the particular potentials. In this case, their accuracy is determined by:

- the simplifications made to the applied differential equation
- the choice of the gauges
- the choice of the element type
- the choice of the shape function
- the accuracy of the equation solver
- the quality of the discretisation.

Therefore, the error convergence of these quantities is of the same order as that of the related potentials. An example to illustrate this:

Using the 2D magnetic vector potential, the normal component of the flux density through the edge of an element is always continuous. The flux through the edge equals the difference of the potentials at the adjacent nodes (the unit of the vector potential is Wb/m). This allows calculation of the flux through a line span between two points just by calculating the difference of the potential value at the end-points. Practically, this could be applied to the calculation of flux linkages.

5.10.3.1 Numerically derived local field quantities Most local field quantities, as well as other derived quantities such as force, require numerical derivatives of the potentials. Using nodal elements, the potentials are known at each node as a result of the approximate solution of the partial differential equation. The change of the potential inside one element is determined by the choice of the shape function:

$$A = a + bx + cy \quad .$$
(5.304)

Knowing the potentials at the nodes of the elements, the coefficients a, b and c can be calculated using this basis function. The definition of the potential now determines the required mathematical operations yielding the required local field value. In two-dimensional magnetostatic problems, the vector potential $\mathbf{A}$ is defined by:

$$\mathbf{B} = \nabla \times \mathbf{A} \quad .$$
(5.305)

Using such linear shape functions to approximate the vector potential, the x- and y-components of the flux density inside a finite element are calculated as follows:

$$B_x = \frac{\partial A}{\partial y} = \frac{1}{2\Delta_e} \sum_i^3 c_i A_i = \text{const.}$$

$$B_y = -\frac{\partial A}{\partial x} = -\frac{1}{2\Delta_e} \sum_i^3 b_i A_i = \text{const.}$$
(5.306)

The flux density B inside an FEM model is piecewise constant (5.306) if a continuous distribution of the vector potential is assumed. Accounting for this and assuming a small value of h as the maximum characteristic diameter of a finite element, the FEM is convergent towards the exact solution of order $q+1$. The constant q describes the polynomial order of the elements used. With ε as the global error, the order of convergence for the potential solution is

$$\|\varepsilon\| \leq C \cdot h^{q+1} .$$
(5.307)

The factor C is independent of the size h of the elements and depends only on the

- type of discretisation
- choice of shape function

- smoothness of the exact solution.

Equation (5.307) identifies the convergence problem transferred into the approximation problem. Using first order linear shape functions the rate of convergence is of order $O(h^2)$. Deriving the field quantities from the potential formulation numerically results in a rate of convergence $O(h)$ for those quantities, i.e. a loss in accuracy of one order compared to the potential solution. Using these field quantities this inherent inaccuracy influences the results of force calculations. This fact identifies the difficulty in obtaining accurate field quantities as a problem of the order of convergence of the numerical method used. To illustrate this fact, consider a domain containing a single linear material. By applying Dirichlet boundary conditions of different values to the left and the right domain border, a constant flux is imposed.
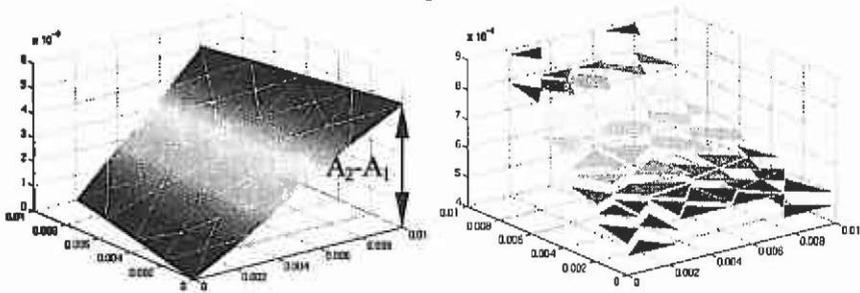
Fig. 5.95. a) Continuous vector potential; b) piece wise constant flux density.

The loss of one order of accuracy due to the numerical differentiation is inherent and effects all quantities based on such values.

There are three possibilities for possibly increasing the accuracy of local field quantities for the end-user of an FEM program package:

- Compute the model with higher order (shape function) elements.
- Increase the quality of the discretisation (adaptive mesh refinement).
- Lower the error bound where the equation solver stops.

The latter point is listed for practical reasons. Especially for eddy current and non-linear problems, a low error bound is absolutely essential. Therefore, it is a good choice to set a stopping margin close to the machine accuracy. If a high accuracy of the local field quantity is required, the relative error of the desired quantity should be monitored. The actions listed in the first two points of the above list can help to achieve this.

- If higher order elements are available in the code, rerun the computation by increasing the order of the elements in each step. Plot the desired quantity versus the number of steps. The relative accuracy can be judged from the convergence of the value towards a stable value.
- Monitor the convergence of the desired quantity over several steps of adaptive mesh refinement. Particular attention has to be paid to the choice of the error estimator. Some error estimators might have advantages for global quantities, but may not be appropriate for a local field value. The error estimator has to have effect in the region of interest.
  - The points listed above can be combined.

Another possible way to increase the accuracy of local field quantities is discussed in the next section: re-calculation of the field distribution in parts of the domain by a local post-process.

Two further points concerning the accuracy of local field values must be mentioned.

- Field values in the vicinity of singularities have a large error.

Adaptive mesh refinement minimises the effect of these regions with respect to the global solution, but the problem does not vanish.

- Smoothing techniques must be applied very carefully. They are popular as they seem to establish a principle in nature: field distributions are smooth. The danger lies in the fact that almost all of the smoothing techniques are based on geometric algorithms rather then on the underlying field equations. Smoothing may lead to just the opposite of what is intended: a loss of local field information.

5.10.4   Forces and torques

Analysing electromagnetic actuators such as electrical machines, the aim is often to find next to the field quantities the electromagnetic forces generated by the studied device. Various methods are in common use. Different methods, their application and limits are discussed.

5.10.4.1 Lorentz force A frequently encountered problem is that of a current-carrying conductor in an external magnetic field. The differential force equation may be written:

$$d\mathbf{F} = I\left(d\vec{\ell} \times \mathbf{B}\right)$$ (5.308)

where $\vec{d\ell}$ is the elementary length in the direction of the current I. Equation (5.308) is derived from the fundamental force relationship between two moving charges. It represents the magnetic part of the Lorentz force. If the conductor is straight and the field is constant along its length, the differential force may be integrated. In a two-dimensional magnetostatic finite element model, the field components are located in the plane, whereas the current is oriented perpendicularly to it. In this case (5.308) can be simplified to the following expression for the conductor of length ℓ:

$$F = BI\ell \quad . \tag{5.309}$$

Those force equations are theoretically valid only for a conductor in a magnetic field. However, in practice it might be used even for the force calculation in electrical machines with many slots containing current, provided that B is the average value in the air gap. This simplification already indicates a loss of accuracy, as local information about the field is not taken into account. This approach combines analytical and numerical field analysis at a rather simplified level.

5.10.4.2 Virtual work One of the most popular methods for the calculation of forces is based on the spatial rate of change of the stored co-energy in the model. The component of the force F_s in the direction of the displacement s is:

$$F_s = \frac{\partial W_{co}}{\partial s} \approx \frac{dW_{co}}{ds} \quad . \tag{5.310}$$

The accuracy of the co-energy calculation is rather high because, as mentioned before, the energy is computed very accurately. It can be expected that the force calculation based on the co-energy should be accurate as well, provided the following requirements are met.

- The method is valid for differential small displacements, which must be translated in terms of the dimensions of the model.
- It is assumed that the magnetic flux remains constant in the two FEM models necessary to compute dW_{co}.

The disadvantage of this method is the need for two finite element computations to obtain a single force value.

A corresponding expression for the torque T associated with an angular rotation θ is useful for the electrical machine analysis:

$$T = \frac{\partial W_{co}}{\partial \theta} \approx \frac{dW_{co}}{d\theta} \quad . \tag{5.311}$$

5.10.4.3 Maxwell stress tensor Probably the most common approach to determine electromagnetic forces is known as the Maxwell stress tensor method. In contrast to the virtual work method, based on the energy, the Maxwell stress tensor method describes the forces directly in terms of the magnetic field strength. This method is advantageous, as forces can be determined with only one FEM-solution.

The Maxwell stress approach computes the local stress at all points of a bounding surface and then sums the local stresses (using a surface integral) to find the overall force. The expression for the Maxwell stress tensor can be derived from (5.308). In three dimensions the force is a surface integral:

$$\mathbf{F} = \nu \oint \mathbf{T} dS \tag{5.312}$$

where the surface vector dS is taken as the outward normal on S. In two dimensions, this reduces to a line integral with the magnetic stress tensor $\mathbf{T}$ written by:

$$\mathbf{T} = \begin{bmatrix} \left(B_x^2 - \frac{1}{2}|B|^2 \right) & B_x B_y \\ B_y B_x & \left(B_y^2 - \frac{1}{2}|B|^2 \right) \end{bmatrix}. \tag{5.313}$$

The expression given above may be rewritten in terms of the normal and tangential components of flux density at each point on the closed contour C along which the line integral has to be evaluated. Therefore the associated components of force for an axial length ℓ is:

$$F_x = \frac{\ell}{2\mu_0} \oint \left(B_x B_n - B_y B_t \right) dl$$
$$F_y = \frac{\ell}{2\mu_0} \oint \left(B_y B_n + B_x B_t \right) dl \tag{5.314}$$

These expressions assume the following notations for the directions of B_n and B_t: for a contour parallel to the y-axis, and traversed in the direction of increasing y, $B_n = B_x$ and $B_t = B_y$. Also the component values of (5.313) have units of stress; they do not necessarily give correct local stress values. However their closed line integral has the physical meaning of the total force on the enclosed object. The contour must be entirely in air and not pass through any other material. In many cases the contour does not need to be closed. Parts of the closed contour may be skipped if their integral value is negligible. The expressions for the force computed on a single straight line are:

$$F_n = \frac{\ell}{2\mu_0} \int (B_n^2 - B_t^2) \mathrm{d}l$$

$$F_t = \frac{\ell}{\mu_0} \int B_n B_t \mathrm{d}l \qquad (5.315)$$

The torque on an arc of radius r is similarly given by:

$$T = \frac{1}{\mu_0} \int B_n B_t r \mathrm{d}l \qquad . \qquad (5.316)$$

Similar expressions may be used for electric field problems by substituting E for B and ϵ_0 for $1/\mu_0$.

The advantage of the Maxwell stress tensor method over the virtual work method recommending only one FEM-solution is lost when the accuracy of the results is compared. As the method is based on derived quantities, particular attention has to be paid. The loss of one order of accuracy compared to the potential solution can lead to large errors, especially when computing the tangential component F_t (5.315) and the torque T in (5.316). In electrical machines the reason for this is the huge difference of the field quantities in magnitude when comparing normal and tangential component. The normal components can differ some decades. This yields large truncation errors in the computed force.

The practical implementation of the algorithm introduces additional error sources. The stress values have to be evaluated at specified points along the contour, usually equidistantly distributed. If such a point is positioned exactly at the edge between two elements, the numerically derived B is double-valued (piecewise constant B for first-order shape function).

The problem of accuracy of the Maxwell stress tensor method and possible ways of improving it, have been extensively discussed in literature. Most of the proposed enhancements are based on smoothing algorithms or on different integration schemes. One of the most common methods proposes calculation of the force using different contours and averaging the result. This method can help to evaluate the margin of error, but it does not give any absolute error bounds or even an enhancement of accuracy.

5.10.5 Enhanced accuracy of finite element field quantities

It will be focused on the practical application of the static electromagnetic field solution of Laplace's equation in a local post-process to increase the accuracy of an existing solution obtained by the standard finite element method using first order elements. Advantages and drawbacks are discussed.

The Laplace equation $\nabla^2 A = 0$ in two dimensions, expressed in polar co-ordinates (r, Φ) is:

$$\frac{\partial}{\partial r}(r\frac{\partial A}{\partial r}) + \frac{1}{r}\frac{\partial^2 A}{\partial \Phi^2} = 0$$

$$r^2\frac{\partial^2 A}{\partial r^2} + r\frac{\partial A}{\partial r} + \frac{\partial^2 A}{\partial \Phi^2} = 0$$

(5.317)

Assuming linearity and uniformity, and applying a Fourier series to eq.(5.317), yields the harmonic function:

$$A(r, \Phi) = \frac{\alpha_0}{2} + \sum_{n=1}^{\infty} r^n \{\alpha_n \cos(n\Phi) + \beta_n \sin(n\Phi)\}$$

(5.318)

with its coefficients:

$$\alpha_n = \frac{1}{\pi R^n}\int_0^{2\pi} A(R, \Phi) \cdot \cos(n\Phi) \cdot d\Phi$$

$$\beta_n = \frac{1}{\pi R^n}\int_0^{2\pi} A(R, \Phi) \cdot \sin(n\Phi) \, d\Phi$$

(5.319)

The procedure for solving eq.(5.318) describes the solution of a Dirichlet problem on a circle with given boundary values at its circumference. The coefficients α_n and β_n can be calculated using known potentials $A = A(R, \Phi)$ at the circumference of a circle with radius R.

Now a finite number of N equi-angularly arranged points are applied onto the circumference of the circle.

$$A_i(R, \Phi_i) = A(R, i \cdot \tfrac{2\pi}{N}) \qquad i = 1(1)N$$

(5.320)

With N boundary potential values u_i known on the circumference and according to the properties of harmonic functions the first term in eq. (5.318) can be written by:

$$A\big|_{r=0} = \frac{\alpha_0}{2} = \frac{1}{N}\sum_{i=1}^{N} A_i$$

(5.321)

The Fourier coefficients are rewritten as follows:

$$\alpha_n = \frac{2}{N \cdot R^n}\sum_{i=1}^{N} A_i \cos(n\Phi_i) \quad ,$$

$$\beta_n = \frac{2}{N \cdot R^n}\sum_{i=1}^{N} A_i \sin(n\Phi_i) \quad .$$

(5.322)

With the Fourier series (5.318) and their coefficients eq.(5.322) the potential in the centre of a circle can be computed knowing only the boundary potential values on the circumference of the circle.

Using this approach inside a finite element solution, the value of the potential of a field point now depends on the solution in several finite elements. Thus, local numeric errors in single elements have a relatively small influence on the solution in the considered field point. Applying

(5.318) derivatives at the centre of the circle can be calculated in a closed analytical form, avoiding numerical differentiation.

$$\left.\frac{\partial A}{\partial y}\right|_{r=0} = \beta_1 = B_x = \frac{2}{N \cdot R}\sum_{i=1}^{N} A_i \sin\Phi_i$$

$$\left.\frac{\partial A}{\partial x}\right|_{r=0} = \alpha_1 = B_y = \frac{2}{N \cdot R}\sum_{i=1}^{N} A_i \cos\Phi_i$$

(5.323)

The idea is to adapt the described process, of solving a Dirichlet problem on a circular surface, to determine the vector potential in a point P_1 of a discretised finite element domain (Fig. 5.96). R_1 is the radius of the considered circular surface and the dots at the circumference indicate the points of known vector potential values computed beforehand. These points do not have to be nodes of the actual finite element mesh. This feature makes the technique very advantageous to automatic and adaptive meshing schemes in which the user can not guarantee the control of the mesh and especially its symmetry.

To obtain the potential distribution at a given contour inside a finite element domain, multiple circles have to be evaluated. Overlapping circles guarantee a continuous solution in the considered region after the post-process.

The numerical shape of (5.318), (5.321) and (5.322) enables an easy implementation of the procedure in a finite element program package. The derivatives in the centre of the circle are represented by the Fourier coefficients. Thus, no additional computational effort is necessary to compute the flux density in the centre.
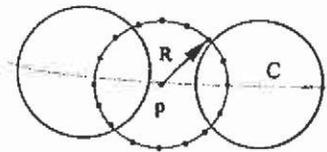
Fig. 5.96. Multiple circles to determine the vector potential on a contour.

The local solution of the Laplace equation inside an air gap of an electromagnetic device, using a Fourier series approximation for the vector potential, results in a significant increase in accuracy of the derived field quantities. To compare the results obtained by the local field evaluation to the conventionally obtained field quantities of first order elements, Fig. 5.97 shows the computed magnetic flux density derived by $B = \nabla \times A$ and B_x using the Laplace approach. For this application of the local Dirichlet problem, 24 potential boundary values on the circumference of the circle were used.

Applying such field quantities to the Maxwell stress tensor to compute the local forces acting on bodies inside an electromagnetic field yields values of higher accuracy. To obtain the torque of an electrical machine, the local force values are integrated along a contour in the air gap.

a)   b)
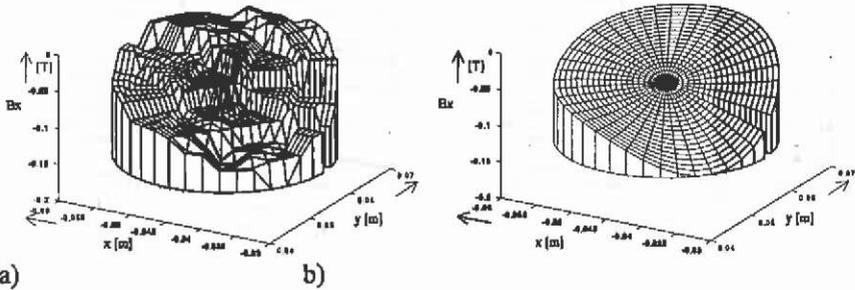
Fig. 5.97. a) Computed vector potential inside the circular FEM domain and b) the resulting flux density B_x derived by applying the local post-process.

Another approach to compute the torque more accurately, uses the values of the magnetic vector potential on two concentric circles with radii R_i and R_o as boundary conditions (Fig. 5.98). Local field values on the circular contour C with radius $R_i < r < R_o$ are calculated.
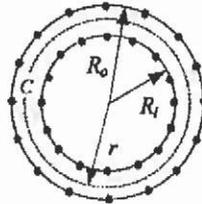
Fig. 5.98. Local Dirichlet problem for a cylindrical air gap.

If the inner radius R_i is taken as a reference, the general solution of Laplace's equation is:

$$A(r,\Phi) = \sum_{k=1}^{N}\left(a_k\left(\frac{r}{R_i}\right)^k \cos(k\Phi) + b_k\left(\frac{r}{R_i}\right)^k \sin(k\Phi)\right.$$
$$\left. + c_k\left(\frac{R_i}{r}\right)^k \cos(k\Phi) + d_k\left(\frac{R_i}{r}\right)^k \sin(k\Phi)\right)$$

(5.324)

The coefficients a_k, b_k, c_k and d_k are independently determined for each circular harmonic. A fast Fourier transformation (FFT) algorithm is used to express the magnetic vector potential at the boundaries as a series of such circular harmonics:

$$A\left(R_{i},\Phi\right)=\sum_{k=1}^{N}\left(a_{k,i}\cos(k\Phi)+b_{k,i}\sin(k\Phi)\right) \quad,$$

$$A\left(R_{o},\Phi\right)=\sum_{k=1}^{N}\left(a_{k,o}\cos(k\Phi)+b_{k,o}\sin(k\Phi)\right) \quad.$$

(5.325)

$$\left[\left(\frac{1}{\frac{R_{o}}{R_{i}}}\right)^{k}\quad\left(\frac{1}{\frac{R_{i}}{R_{o}}}\right)^{k}\right]\left[\begin{array}{c}a_{k}\\c_{k}\end{array}\right]=\left[\begin{array}{c}a_{k,i}\\a_{k,o}\end{array}\right] \quad,$$

$$\left[\left(\frac{1}{\frac{R_{o}}{R_{i}}}\right)^{k}\quad\left(\frac{1}{\frac{R_{i}}{R_{o}}}\right)^{k}\right]\left[\begin{array}{c}b_{k}\\d_{k}\end{array}\right]=\left[\begin{array}{c}b_{k,i}\\b_{k,o}\end{array}\right] \quad.$$

(5.326)

Once the magnetic vector potential at the contour C is known, the normal and tangential component of the magnetic flux density can be determined:

$$B_{n}(r,\Phi)=\sum_{k=1}^{N}\left(-ka_{k}\frac{r^{k-1}}{R_{i}^{k}}\sin(k\Phi)+kb_{k}\frac{r^{k-1}}{R_{i}^{k}}\cos(k\Phi)\right.$$
$$\left.-kc_{k}\frac{R_{i}^{k}}{r^{k+1}}\sin(k\Phi)+kd_{k}\frac{R_{i}^{k}}{r^{k+1}}\cos(k\Phi)\right) \quad,$$

$$B_{t}(r,\Phi)=\sum_{k=1}^{N}\left(-ka_{k}\frac{r^{k-1}}{R_{i}^{k}}\cos(k\Phi)-kb_{k}\frac{r^{k-1}}{R_{i}^{k}}\sin(k\Phi)\right.$$
$$\left.+kc_{k}\frac{R_{i}^{k}}{r^{k+1}}\cos(k\Phi)+kd_{k}\frac{R_{i}^{k}}{r^{k+1}}\sin(k\Phi)\right) \quad.$$

(5.327)

The tangential force component F_t results in the torque T of the device. It can be shown (Salon [99], Mertens et al. [78]) that the value of the torque is given by

$$T=\frac{2\pi}{\mu_{0}}\sum_{k=1}^{N}\left(k^{2}\left(b_{k}c_{k}-a_{k}d_{k}\right)\right) \quad,$$

(5.328)

being independent of the radius r of contour C. It is not necessary to calculate the normal and tangential component of the magnetic flux density on the contour resulting in a faster algorithm, when the overall torque is aimed at. The proposed method can easily be extended to time-harmonic problems. If all values are rms-values the torque is obtained by adding the torque calculated using the real- and the imaginary-component of the solution.

$$T=T_{real}+T_{imag} \quad.$$

(5.329)

This method has its advantage for electrical machine analysis as it is suited for small air gaps. By using eq.(5.329) the torque is evaluated directly, without the explicit calculation of the flux densities.
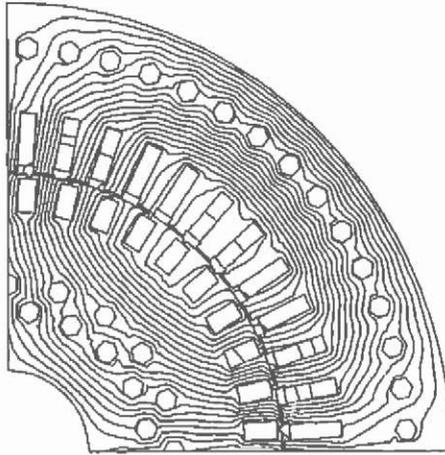
Fig. 5.99. Equipotential plot of the real component solution of a 400 kW induction motor.

The performance of this method is compared with the classical Maxwell stress tensor method using a model of a 400 kW induction machine for tests.

Table 5.10. Data of the 400 kW induction machine.

induction machine data	
voltage	0-1950 V
rated current	154 A
rated power	402 kW
cos φ	0.91
rpm	0-4350
frequency	0-140 Hz
number of pole pairs	2
number of stator slots	48
number of rotor slots	40
outer diameter	60 cm
air gap length	1.5 mm
measured torque	$T_m = 2733$ Nm

To ensure accurate results, a good trade-off between mesh refinement and using the enhanced post-processing methods is necessary. A relatively coarse discretisation in the air gap of the induction machine was chosen (Fig. 5.100).

With such a coarse discretisation, the Laplace-based method is less sensitive to the actual choice of the contour inside the air gap than the

classical method. The air gap spans a region between an inner radius of 0.186 m up to an outer radius of 0.1875 m. Fig. 5.101 shows the variation of the calculated torque. Contours with different radii are chosen. For the Laplace-based method, the inner and outer radii are varied simultaneously. Therefore, the value of the torque varies symmetrically towards the middle of the air gap.
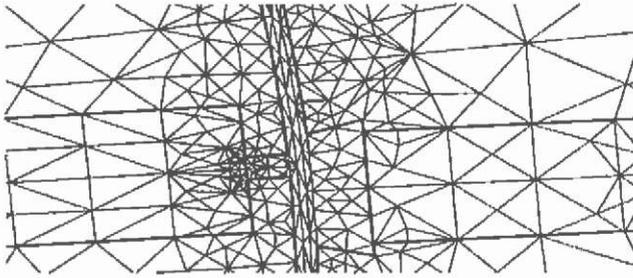
Fig. 5.100. Detail of the discretisation in the air gap.

The variation of the calculated torques using the enhanced method is much smaller when compared to the classical method. It must be stated, however, that an appropriate mesh refinement scheme would lead to better results even for the classical torque computation.
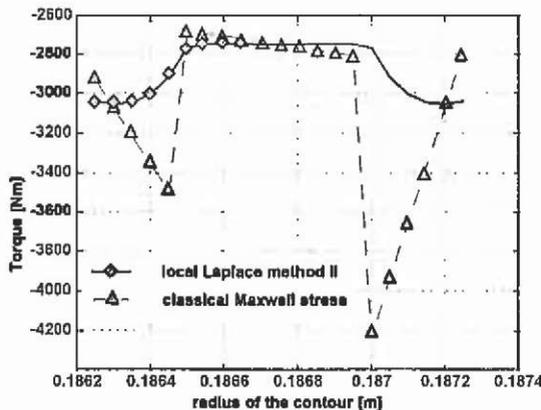
Fig. 5.101. Variation of the torque calculated along different contours inside the air gap.

The rate of convergence of the relative error indicates that the required number of sample points for the enhanced method can be chosen to be substantially fewer than for the classical method. As shown in Fig. 5.102, the smallest relative error is computed with 16384 points using the

classical Maxwell stress method, instead of 2048 points by using the Laplace-based approach.
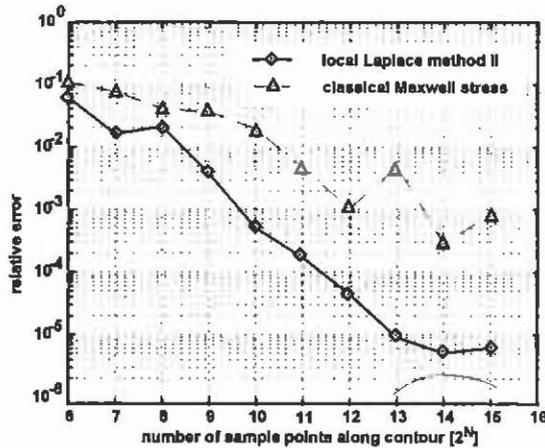
Fig. 5.102. Rate of convergence of the relative error. (Variation of the number of sample points.)

The same basic idea, as used in the 'circle' approach, yields the local solution for the three-dimensional field. The local field problem is now defined by the known potential values equally distributed along the surface of a sphere assumed to be the boundary potential values of the local field problem. According to the co-ordinate transformation:

$$x = r \sin \theta \cos \phi$$
$$y = r \sin \theta \sin \phi$$
$$z = r \cos \theta$$
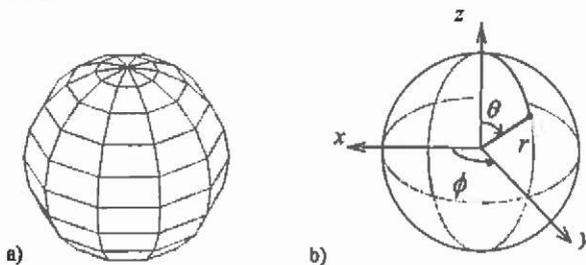
Fig. 5.103. a) Sphere with b) co-ordinate system.

a spherical co-ordinate system is applied (Fig. 5.103). Using the Laplace equation with the co-ordinate transformation yields:

$$\frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial A}{\partial r}\right)+\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial A}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2 A}{\partial\phi^2}\right]=0\,. \tag{5.330}$$

Applying the theorem of the separation of the variables $A(r,\theta,\phi)=R(r)\cdot\Theta(\theta)\cdot\Phi(\phi)$ to (5.330), a general form of the functions $R(r)$, $\Theta(\theta)$ and $\Phi(\phi)$ depending on the potential A can be written. Every solution of the Laplace equation, being finite for all θ, is a solution of:

$$A_{m,n}(r,\theta,\phi)=\left(ar^n+br^{-(n+1)}\right)P_n^m(\cos\theta)\cdot\left(\alpha\cos m\phi+\beta\sin m\phi\right) \tag{5.331}$$

where $m=0(1)\infty$, $n=m(1)\infty$, a, b, α and β are constants. P_n^m is the associated Legendre polynomial of the first kind. To simplify the notations, the surface harmonics

$$\begin{aligned}c_{m,n}&=P_n^m(\cos\theta)\cdot\cos m\phi\\s_{m,n}&=P_n^m(\cos\theta)\cdot\sin m\phi\end{aligned} \tag{5.332}$$

are introduced. Assuming (5.331) to be a linear form, the potential in the origin is finite. The constants $\alpha_{m,n}(r)$ and $\beta_{m,n}(r)$ are linear combinations of r^n and $r^{-(n+1)}$. The summation

$$A=f(\theta,\phi)=\sum_{m=0}^{\infty}\sum_{n=m}^{\infty}\left[p_{m,n}\cdot c_{m,n}(\theta,\phi)+q_{m,n}\cdot s_{m,n}(\theta,\phi)\right]\cdot r^n \tag{5.333}$$

is a solution of (5.330). Here, the magnetic scalar potential A is completely determined by the constants $p_{m,n}$ and $q_{m,n}$. The aim is to calculate the magnetic flux density at a point using known scalar potential values in its vicinity. Consequently a spherical volume with known boundary potentials at its surface around this field point is chosen to determine the field. The known boundary potentials result, as in the two-dimensional case, from a previously performed FEM computation and determine all constants in (5.333). To calculate the magnetic field quantities at the centre of the spherical volume, the Laplace equation has to be solved locally and spherically around this field point with radius r=R. The boundary potential values are available only as single values at the surface of the sphere. To distribute them equally along this surface, the spherical co-ordinates ϕ are divided into J and θ into K equal angles $\Delta\phi$ and $\Delta\theta$ respectively.

$$\begin{aligned}2\pi&=\sum_{j=1}^{J}\Delta\phi\\\pi&=\sum_{k=1}^{K}\Delta\theta\end{aligned} \tag{5.334}$$

In order to satisfy (5.333) accurately, the numbers of J and K must be sufficiently large. On the other hand, large numbers increase the computational expenses rapidly. With respect to the computation time

and accuracy a good compromise has to be found. Practical values for J and K are given in the following section.

Assuming ϕ and θ to be the co-ordinates in the local system with the interesting field point at the centre of the sphere, the coefficients $p_{m,n}$ and $q_{m,n}$ can be determined by a Legendre decomposition using the boundary potential values:

$$p_{m,n}\Big|_{m=0} = \frac{\pi}{J \cdot K}\frac{1}{2R^n}\frac{(2n+1)(n-m)!}{(n+m)!}\sum_{k=1}^{K}\left[\sum_{j=1}^{J}f(\theta_k,\phi_j)\right]$$
$$\cdot P_n^m(\cos\theta_k)\cdot\sin\theta_k$$

$$p_{m,n}\Big|_{m>0} = \frac{\pi}{J \cdot K}\frac{1}{R^n}\frac{(2n+1)(n-m)!}{(n+m)!}\sum_{k=1}^{K}\left[\sum_{j=1}^{J}f(\theta_k,\phi_j)\cos m\phi_j\right] \qquad (5.335)$$
$$\cdot P_n^m(\cos\theta_k)\cdot\sin\theta_k$$

$$q_{m,n} = \frac{\pi}{J \cdot K}\frac{1}{R^n}\frac{(2n+1)(n-m)!}{(n+m)!}\sum_{k=1}^{K}\left[\sum_{j=1}^{J}f(\theta_k,\phi_j)\sin m\phi_j\right]$$
$$\cdot P_n^m(\cos\theta_k)\cdot\sin\theta_k \qquad .$$

Retaining the local co-ordinate system in (x', y', z'), and the magnetic flux density in the original global co-ordinate system, and calculating the derivatives at the origin of the local co-ordinate system (Fig. 5.103) using $\theta = \pi/2$ and $\phi = 0$ in (5.333), yields:

$$-\mu_0\frac{\partial A}{\partial x'}\Big|_{(0,0,0)} = -\mu_0\frac{\partial A}{\partial r}\Big|_{(0,0,0)} \qquad . \qquad (5.336)$$

Analogous to (5.336) the derivatives, in y' are found by taking $\theta = \phi = \pi/2$ and in z' by taking $\theta = 0$ and $\phi = \pi/2$, in (5.333). With respect to (5.336), using (5.335) and with the Legendre terms:

$$P_1^0(\cos\theta) = \cos\theta$$
$$P_1^1(\cos\theta) = \sin\theta \qquad (5.337)$$

the components of the flux density at the centre of a sphere are explicitly rewritten by:

$$B_x = -\mu_0\frac{3}{2}\frac{\pi}{J \cdot K \cdot R}\left[\sum_{k=1}^{K}\left(\sum_{j=1}^{J}f(\theta_k,\phi_j)\cos\phi_j\right)\cdot\sin^2\theta_k\right]$$

$$B_y = -\mu_0\frac{3}{2}\frac{\pi}{J \cdot K \cdot R}\left[\sum_{k=1}^{K}\left(\sum_{j=1}^{J}f(\theta_k,\phi_j)\sin\phi_j\right)\cdot\sin^2\theta_k\right] \qquad (5.338)$$

$$B_z = -\mu_0\frac{3}{2}\frac{\pi}{J \cdot K \cdot R}\left[\sum_{k=1}^{K}\left(\sum_{j=1}^{J}f(\theta_k,\phi_j)\right)\cdot\cos\theta_k\cdot\sin\theta_k\right]$$

Using this local field approach (5.338) by arranging multiple overlapping spheres at an arbitrary surface or contour (Fig. 5.104), it is possible to obtain the required local field quantities at this surface with

the same accuracy as the previously by FEM computed potential values. The tetrahedron shown in Fig. 5.104 represents a part of the three-dimensional mesh of the FEM domain.
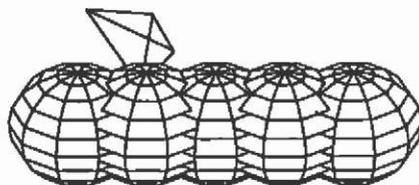
Fig. 5.104. Arrangement of multiple overlapping spheres to obtain the local field values on an arbitrary contour/surface across the centre points of the spheres inside a three-dimensional FEM domain.

a)   b)
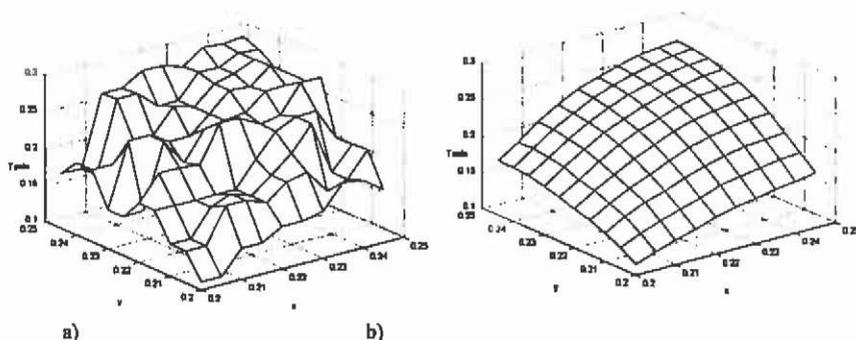
Fig. 5.105. Flux density distribution B_z on the front surface of Γ (see Fig. 5.106), a) computed by the classical direct derivation of the potential and b) using the proposed post-processor method.
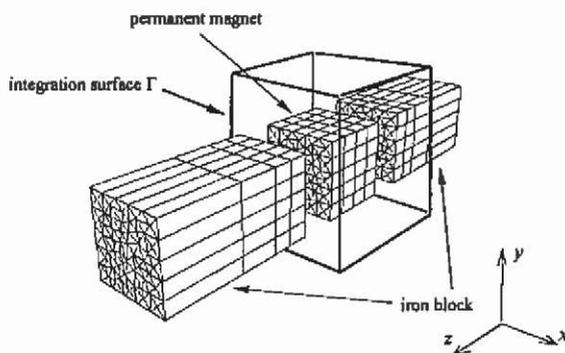
Fig. 5.106. Three-dimensional FEM model of the test example.

From Fig. 5.105 the difference between the direct evaluation of the potential and the new post-process operator is shown. Here B_z is computed for a test example (Fig. 5.106) at the front surface of Γ facing the permanent magnet cube. It is obvious that linear shape functions,

approximating the scalar potential, result in a piecewise constant flux density distribution (Fig. 5.105a). Computed forces starting from this type of solution are unreliable. The local values of B_z plotted in Fig. 5.105b show the expected continuous distribution computed using the new post-processor method.
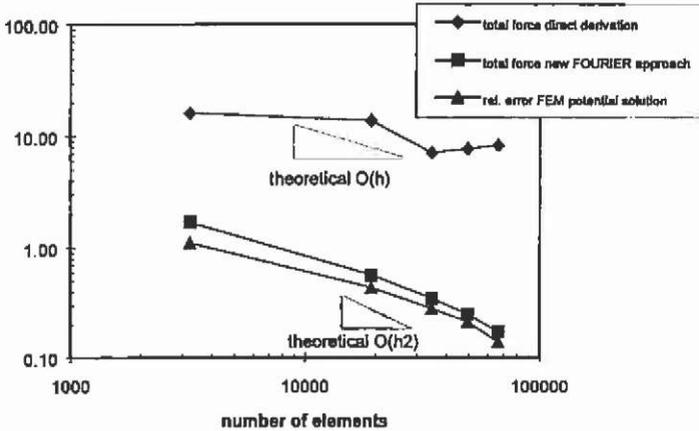
Fig. 5.107. Comparison of the convergence behaviour of the FEM potential solution with both the direct derivation and the derivative-free approach.

In Fig. 5.107 the quadratic convergence, referred to the characteristic length h of a finite element, of the FEM potential solution and the rate of convergence of the force computations using both the classical and the new post-processing approach, is plotted versus the number of tetrahedron elements. The same statement can be made for the two-dimensional approach. The triangles in Fig. 5.107 indicate the theoretical gradient of convergence eq.(5.307). The refinement of the three-dimensional discretisation is performed in such a way that the elements are of the same shape in every FEM model to obtain a regularly distributed mesh for all cases. To compute the total force, the Maxwell stress tensor is used integrating the force density calculated in points equidistantly distributed by the density D on all six sides of Γ. For the classical approach a density $D=40$ is chosen and in the case of the new method, D is set to 7. The sphere parameters are $J=K=15$. The integration surface of the force computations is located in such a way that no plane of Γ cuts through the nodes of the FEM mesh. If nodes coincide with the points of the force computation using the classical post-processor approach, this would result in a larger error due to the troublesome definition of normal and tangential field components in a node of an element. The gradient-triangles in Fig. 5.107 indicate the theoretical rate of convergence for the quadratic and the linear convergence case. It can

be seen as theoretically expected that the relative error in an energy norm of the FEM potential solution converges quadratically, referred to the specific diameter h of the elements eq.(5.307), by increasing the number of first order tetrahedron elements. Due to the analytically described potential function inside the local field volumes, the resulting overall force using this approach is of the same order of convergence. Therefore, no loss of accuracy of the derived field quantities occurs. The convergence of the total forces, computed by the classical approach, indicates the expected linear behaviour. The accuracy of the computed values is influenced by the numerically-obtained derivatives. This shows that the results obtained by the classical method are inherently inaccurate when compared to the accuracy of the potential solution.

The use of the proposed approach to enhance the accuracy of computed field quantities starting from an existing potential solution demands an additional step during the post-processing of the FEM analysis (Fig. 5.108).
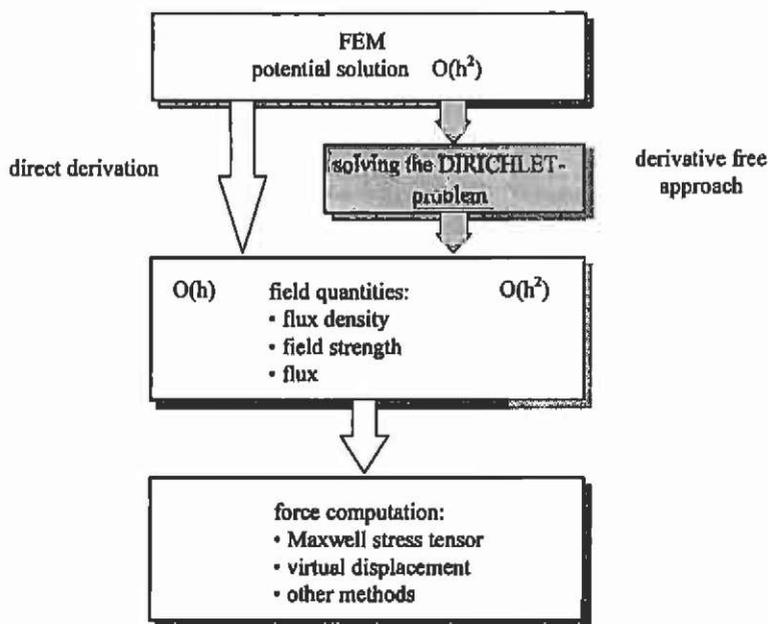
Fig. 5.108. Additional step during post-processing to enhance the accuracy of derived field quantities.

Having obtained a FEM potential solution, the user only has to define the surface of integration Γ on which the field quantities or forces have to be calculated. Defining an arbitrary contour allows the computation of field quantities or forces along it as well. For each plane

or contour the density D, the sphere parameter J, K and the radius R have to be set. The sphere parameter are problem-dependent and related to the geometry of the device, i.e. the air gap width. The planes or contours should be centred in the air gap. A suitable value for the diameter of the single spheres is about 90-95 % of the air gap width to have as many tetrahedron finite elements inside the sphere as possible. Including only one finite element in the sphere results in no enhancements in accuracy of the derived quantity. To ensure a continuous field solution, the density D should be chosen in such a way that the spheres overlap (Fig. 5.104). For the distance between two points on the surface of integration, it is suitable to choose the radius of the sphere.

$J=K=4$ $J=K=10$ $J=K=20$
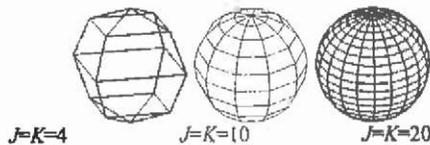
Fig. 5.109. Spheres with different numbers for the parameter J, K.

To define the number and position of boundary potential values distributed on the surface of each sphere, the parameters J and K have to be chosen. To ensure uniformly distributed boundary values J is set equal to K. In accordance with the results of Fig. 5.105 and other test calculations a number $J=K=[10 \ldots 20]$ is sufficient to meet the ratio between computational costs and accuracy. Fig. 5.109 illustrates by different $J=K$ the position and number of boundary potentials to approximate the local field inside a sphere.

5.10.6   Inductances in magnetostatic problems

Quantities, such as inductances, resitances, etc. may be determined from a numerical solution by several methods. It is just as in a laboratory, where a quantity may often be measured in many ways eventually leading to different results. Modelling implies the simplification of the complex physical phenomena determining the behaviour of a technical device. A FEM model of a device serves the purpose of predicting certain aspects of its behaviour while neglecting others. The first simplification is introduced by the choice of the formulation of the equations to solve. However, even if the basic formulation is appropriate, different results may be obtained in the post-process if inappropriate definitions of quantities are employed. This effect is illustrated by the calculation of inductances using a magneto-static analysis.

5.10.6.1 Linear models Consider a simple iron core inductor drawn in Fig. 5.110. The aim is to compute the inductance of this device to be used in an equivalent electric circuit. It is assumed that the core is loss-free without leakage flux and the winding end-effects are neglected. A two-dimensional linear, electromagnetic analysis of the device is applied.
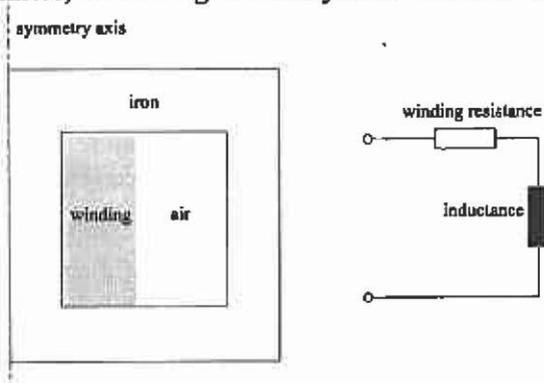
Fig. 5.110. Iron core inductor and equivalent electric circuit.

Two definitions of the inductance can be found: one based on the flux linked with the winding (5.339) and one based on the energy stored in the inductor:

$$L = \frac{N \cdot \Phi}{I} \tag{5.339}$$

$$L = 2 \cdot \frac{\int\limits_V \left(\int\limits_B H \cdot dB \right) dV}{I^2} \tag{5.340}$$

with N the number of turns, I the terminal current and V the volume of the device.

The flux linkage with the winding in (5.339) can be determined from the vector potential inside the winding area:

$$\Phi = \int\limits_S \mathbf{B} d\mathbf{S} \tag{5.341}$$

with S the surface in which the flux is penetrating. Based on the definition of the magneto static vector potential ($\mathbf{B} = \nabla \times \mathbf{A}$), the flux can be determined over the vector potential integrated along the length of the winding $d\mathbf{l}$:

$$\Phi = \oint\limits_l \mathbf{A} d\mathbf{l} \quad . \tag{5.342}$$

Equation (5.342) can be simplified when considering a two-dimensional Cartesian geometry, with no variation of the field and geometry in z-direction:

$$N \cdot \Phi = N \cdot l \cdot \left(A_{right} - A_{left} \right) \tag{5.343}$$

with l the length of the device in z-direction and A the average value of the vector potential at the right and left side of the winding. Due to the symmetry in the given device and the discretisation of the winding cross section, it can be written:

$$N \cdot \Phi = 2 \cdot N \cdot l \cdot \frac{\sum_{k=1}^{n} \Delta_k A_k}{\sum_{k=1}^{n} \Delta_k} \tag{5.344}$$

with n the number of elements over the cross section of the winding and Δ_k the surface of element k. In the general case (no symmetry), the flux linkage can be extracted from the FEM solution via an integration over the winding cross sections by applying:

$$N \cdot \Phi = l \cdot \frac{\int_{\Omega} A \cdot J d\Omega}{I} \tag{5.345}$$

with J the current density in the fractional winding cross section $d\Omega$, with Ω the surface of the winding cross section (including both sides of the winding). This last equation automatically accounts for the number of turns and the orientation of the different sides of one winding, as the sign of both the vector potential and the current density are related. The value for the inductance can be calculated from:

$$L = \frac{\Psi}{I} = \frac{N \cdot \Phi}{I} = l \cdot \frac{\int_{\Omega} A \cdot J d\Omega}{I^2} . \tag{5.346}$$

When comparing this equation with the magneto static energy functional, it can be recognised as the term for the linear energy, i.e. the energy supplied from the source:

$$F(A) = \frac{1}{2} \int_{\Omega} \left(v_r |\nabla A|^2 - 2 \frac{J_0 A}{v_0} \right) d\Omega \quad . \tag{5.347}$$

A sinusoidal variation of the source current will also result in a sinusoidal change of this energy; the system is linear.

In a linear analysis ($v_r = const.$), the calculation of the inductance via the stored magnetic energy in the model, (5.340) gives the same result. The stored magnetic energy per unit volume can be represented as the surface above the material characteristics (Fig. 5.111).

The amount of the stored energy per unit volume is thus geometrically:

$$W_\bullet = \frac{B \cdot H}{2} = \frac{\Phi \cdot N \cdot I}{2} \quad . \tag{5.348}$$

A sinusoidal variation of the source current yields a sinusoidal variation of the stored energy. Hence, the calculation of the inductances based on flux linkage and stored magnetic energy in the model give equal results.
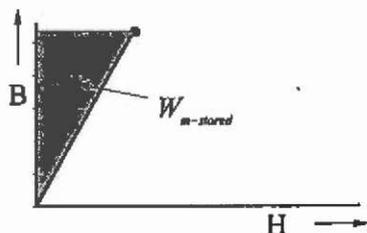
Fig. 5.111. Representation of the stored energy in the model at a defined operating point.

This "linear" definition of an inductivity is equivalent to a measurement by a ballistic flux meter.

5.10.6.2 Non-linear models For the computation of the inductance based on the flux linkage, (5.346) can be applied without changes. However, this value differs from the value of the stored energy computed by (5.340). The material is non-linear (Fig. 5.112). A sinusoidal current excitation does not yield a sinusoidal change of the stored magnetic energy due to the saturation effect. Using eq.(5.349) is not appropriate for this sinusoidal operation because it does not consider the effect of higher harmonics. Signals modulated at non-linear characteristic do contain higher harmonics. Ferromagnetic saturation usually generates harmonics of threefold fundamental frequency.

$$W_{non-linear} = \int_V \left(\int_B H \cdot dB \right) dV \tag{5.349}$$
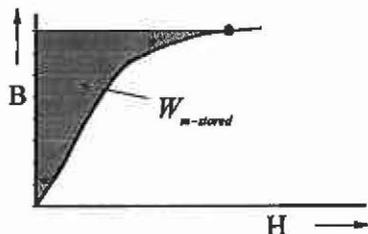
Fig. 5.112. Stored energy (5.349) in a non-linear model at a defined operating point.

The calculation of the stored energy is equivalent to an impedance measurement, which includes the measurement of the terminal voltages and currents. If sinusoidal current is applied, there will be odd harmonics

in the voltage waveform, which cannot be determined by a magneto static analysis.

5.10.6.3 Example: Inductor for a fluorescent lamp To illustrate the problem, consider a ballast inductor from a circuit of a fluorescent lamp (Fig. 5.113). The inductor is about seven times longer than it is wide. It can be treated as a two-dimensional Cartesian problem neglecting end-effects.
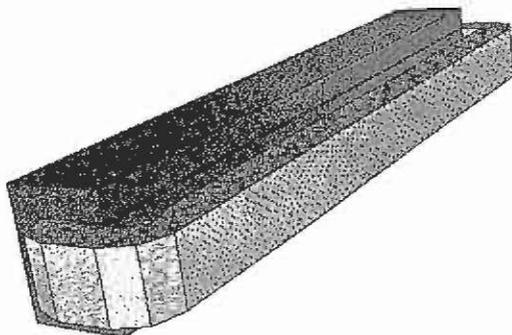
Fig. 5.113. Half symmetry of the inductor.

The design goal for the inductor is twofold: stabilise the current during the heating phase for the electrodes of the lamp and provide a defined over-voltage when the starter opens in order to ignite the fluorescent lamp.
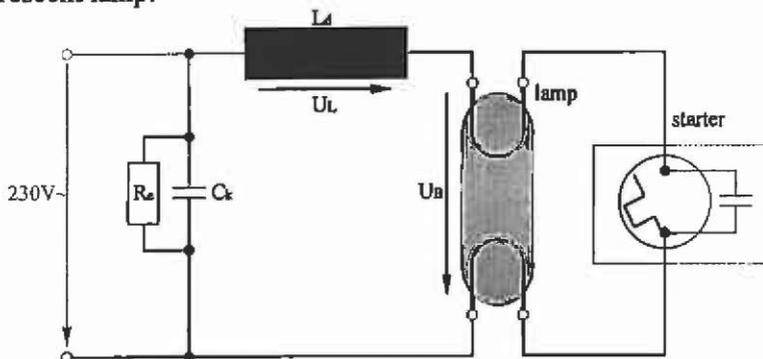
Fig. 5.114. Circuit of a fluorescent lamp and shape of the inductor.

The results of the inductance computation (Fig. 5.115) using both definitions clearly indicate the difference that occurs for the saturated operation points of the device. This difference is useful in this case as it allows determination of the linear operation range of the device. However, the value of the inductance should be determined based on the flux linkage.
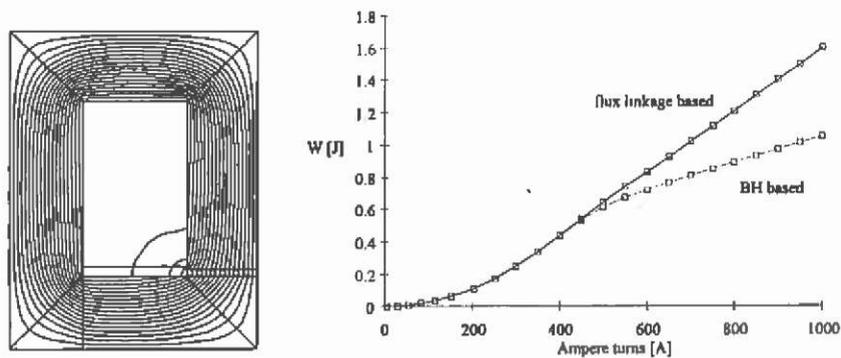
Fig. 5.115. Two-dimensional flux plot and the computed energy.

6 Coupled field problems

The term "coupled problem" is used in many numerical approaches and applications. Various coupling mechanisms in a different context, such as field problems with electrical circuits, methods in a geometrical or physical sense, couplings in time and/or coupled methods to solve a field problem, are meant with this term. For a proper classification of these problems and related solution methods a systematic definition is proposed. It can be used in the evaluation and comparison of solution methods for various problems.

A coupled system or formulation is defined on multiple domains, possibly coinciding, involving dependent variables that cannot be eliminated on the equation level (Zienkiewicz [123]). In the literature, this notion is often linked to a distinguishing context of various physical phenomena or methods, without further specification. This paper proposes a classification scheme in which the numerical models meeting the proposed definitions can be put. This may lead to the definition of a series of test problems for specified coupled problems and solution algorithms. A classification scheme can simplify the comparison of the various examples and approaches out of the literature that solve such coupled problems.

Next to "coupled problems" the terms "weak-" or "strong-coupled" will be discussed to propose a more homogenous terminology.

6.1   Coupled fields

To start with a definition of standards or a classification of technical physical problems, the properties and the interdependencies of such phenomena must be considered and discussed.

A general and simplified structure of considered field problems is drawn in Fig. 6.1. Here, the link between the single fields is determined by material properties depending on the corresponding field quantities. If the field blocks represent numerical methods to solve the single problem in two dimensions, further couplings to external equations such as

electrical circuits, magnetic or thermal equivalent circuit models are possible to complete the scheme.

The link between the drawn blocks is, in the context of coupled problems and its numerical solution, a computer model or method. The following question is in which way the physical phenomena have to be considered in an overall solution. From the idea of how to link the effects numerically, a classification of the methods in this sense can be performed.

The coupling of magnetic field equations, described by a partial differential equation (PDE) and the electrical circuit equations providing algebraic expressions for the electrical current densities, can be considered as a special type of coupled problem.
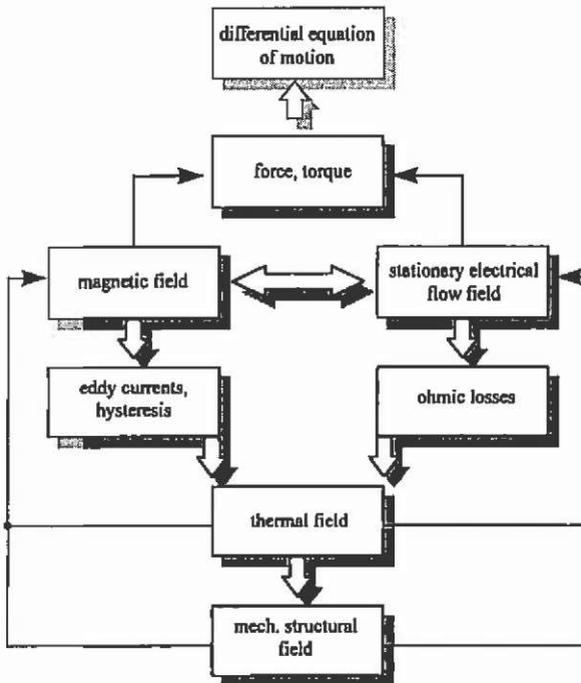
Fig. 6.1. Simplified structure of coupled fields.

In general, more than one independent physical field variable is involved. The field variables for stationary problems are present in a set of PDEs, or in the transient case in ordinary differential equations (ODE). The coupling is often non-linear and this results in a complicated numerical solution process.

Fields can be described by differential equations. A general form of a differential equation has to be studied to understand the parameter coupling between equations. Equation (6.1) represents the general form

of a differential equation with its possible coefficients in the particular terms. In coupled fields problems, such coefficients are field dependent and represent the link between the various field types, such as magnetic/thermal etc.

$$a\frac{\partial}{\partial t}f \ -\nabla\cdot(\lambda\nabla f) \ +\gamma\nabla f \ +\alpha f \ = g \tag{6.1}$$

1. parabolic, transient term
2. diffusion
3. convection
4. absorption
5. source
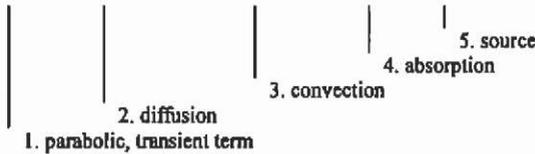
The first term characterises the equation being parabolic. Stationary equations do not contain this term ($\frac{\partial}{\partial t} = 0$). In Laplace's equation terms 2 and 5 are present. To obtain the Helmholtz equation, term 4 can be added. For these two types of equations, a variational formulation exists. The 3rd term is typical for problems considering motion effects, eq.(5.284).

The coefficients in (6.1) are usually derived from given material characteristics. For example, temperature dependent material properties of permanent magnet material can be used to define a coupled magnetic/thermal field problem. Within a field problem definition, the characteristics vary locally.

6.2 Strong and weak coupling

In general, it is possible to distinguish between the coupled problem in two ways, in its physical or its numerical nature. Very often a coupled problem is called either

- strong, or
- weak.

In the physical sense, the strong coupling describes effects that are physically strongly coupled and the phenomena can not numerically be treated separately. If numerical formulations exist, the coupling can be found in the governing differential equations due to the coupling terms. The weak coupling describes a problem where the effects can be separated. The problem with this definition is obvious: If coupled problems are studied, it is not very well known how strong or weak they are physically coupled; this is the desired answer expected from the analysis of the overall problem. For example if the material property describing parameters are non-linearly dependent on the field quantities,

the coupling, (strong/weak) can even change with varying field quantities and the field quantities are the result of the analysis. Therefore, the definition of strong/weak coupling should be chosen according to the numerical aspects instead of their physical nature. Choosing for the numerical aspects, it is possible to have a combined strong/weak coupling of field problems. This means that the strategy of coupling can vary, and thus the methods/models, while solving the problem.

Numerical strong coupling is the full coupling of the problem describing equations on matrix level. The equations of all involved and modelled effects are solved simultaneously. This implies that the coupling terms are entries in the coefficient matrix as well.

The numerical weak coupled problem is understood as a cascade algorithm, where the considered field problems are solved in successive steps and the coupling is performed by up-dating and transferring the field dependent parameters to the other field definition before solving again.

Since the problems cannot be distinguished by means of elimination, a bi-directional influence exists. The sensitivity of a sub-problem to changes of the variables of the studied problem can differ strongly. It is difficult to quantify a threshold for separation of both groups, and therefore the separation may be considered as somewhat subjective. In this respect, the time constants of the sub-problems play an important role. Usually the thermal and mechanical time constants are several orders larger than the electromagnetic time constants. So, on a short term, the problem with a larger time constant can be considered as weak coupled. But this is not true if the stationary solution is of interest.

6.2.1 FEM coupling of two fields

In this section the strong coupling FEM equation system of a magnetic/thermal problem is derived. For simplicity it is assumed that both field problems are defined on the same mesh. For a more realistic coupling, projection methods can be applied to enable the field definitions on different meshes. This approach results in additional coupling terms in the final coefficient matrix. For further simplicity, the material's properties v and k are assumed to be independent of A and T respectively. The coupling of the fields causes the remaining non-linearity by the loss mechanism.

The magnetic/thermal coupled problem is modelled by a set of two equations:

$$v\nabla^2 A - j\omega \cdot \sigma A = -J_0$$
$$k\nabla^2 T = -q_0 \quad .$$

$$(6.2)$$

It is assumed that the source term of the thermal equation consists only of joule losses:

$$q_0 = q_{joule} = \rho \cdot J^2(A) = \rho \cdot \left(J_0^2 + \frac{\omega^2 A^2}{\rho^2} \right) \quad . \tag{6.3}$$

The first term will appear on the right-hand side of the system. The second term, the eddy current losses, have to be linearised and represent the coupling term with a non-linear coefficient:

$$q_0 = \rho J_0^2 + \left(\frac{\partial q_{eddy}}{\partial A} \right) A = \rho J_0^2 + \left(2\sigma\omega^2 A \right) A = \rho J_0^2 + m(A)A \tag{6.4}$$

Written in matrix/vector notation eq.(6.4) is rewritten as:

$$\begin{bmatrix} \upsilon\nabla^2 - j\omega \cdot \sigma & 0 \\ m & k\nabla^2 \end{bmatrix} \cdot \begin{bmatrix} A \\ T \end{bmatrix} + \begin{bmatrix} J_0 \\ \rho J_0^2 \end{bmatrix} = 0 \quad . \tag{6.5}$$

There is a coupling present through the coefficients, although there is a zero entry in the off-diagonal of the magnetic equation. Applying the Galerkin approach results in an integral per element of the form:

$$\int_{\Omega_e} \left(\begin{bmatrix} N_A \\ N_T \end{bmatrix}^T \cdot \begin{bmatrix} \upsilon\nabla^2 - j\omega \cdot \sigma & 0 \\ m & k\nabla^2 \end{bmatrix} \cdot \begin{bmatrix} N_A \\ N_T \end{bmatrix} \begin{bmatrix} A_i \\ T_i \end{bmatrix} + \begin{bmatrix} N_A \\ N_T \end{bmatrix}^T \cdot \begin{bmatrix} J_0 \\ \rho J_0^2 \end{bmatrix} \right) d\Omega = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad . \tag{6.6}$$

For two-dimensional first order elements this yields six algebraic equations:

$$\begin{bmatrix} K_A & 0 \\ M & K_T \end{bmatrix} \cdot \begin{bmatrix} A_i \\ T_i \end{bmatrix} - \begin{bmatrix} F_A \\ F_T \end{bmatrix} = \begin{bmatrix} * & * & * & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 \\ + & + & + & * & * & * \\ + & + & + & * & * & * \\ + & + & + & * & * & * \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ T_1 \\ T_2 \\ T_3 \end{bmatrix} + \begin{bmatrix} \oplus \\ \oplus \\ \oplus \\ \otimes \\ \otimes \\ \otimes \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad . \tag{6.7}$$

The first three equations are complex, the last three real. The entries marked with an * are the same terms that would be found in the de-coupled problem. The terms marked with a + result from the eddy current heat source term.

6.2.1.1 Non-linear iteration A method of handling the remaining non-linearity, is the Piccard iteration or successive substitution. The block iteration scheme can be given by:

| step1 | solve the magnetic equation (with relaxation) | $A_i^{k+1} = \left(K_A^k \right)^{-1} \cdot F_i^k$ |
| step2 | calculate the heat source terms | $M^k A_i^{k+1}$ or $M^k A_i^k$ |

| step3 | solve the thermal equation (with relaxation) | $T_i^{k+1} = \left(K_T^k\right)^{-1} \cdot F_T^k - M^k A_i^{k(+1)}$ |
| step4 | check convergence | |

Every loop involves the solution of two systems of equations in successive steps. For the calculation of the heat sources, the magnetic solution from step1 can be employed (Gauss-Seidel-like algorithm) or the magnetic solution of the previous loop (a slower Jacobi-like algorithm). The relaxation of the iterations proves very important. An adaptive relaxation parameter can reduce the number of iterations significantly. A faster convergence can be expected applying Newton iterations.

6.3   Coupled problems

The overall term coupled problems considers the coupled fields and in addition includes the coupling of methods as well. The link between different methods, hybrid methods, to solve a field problem, for example using the combination of finite element and boundary element method, is understood as a coupled problem.
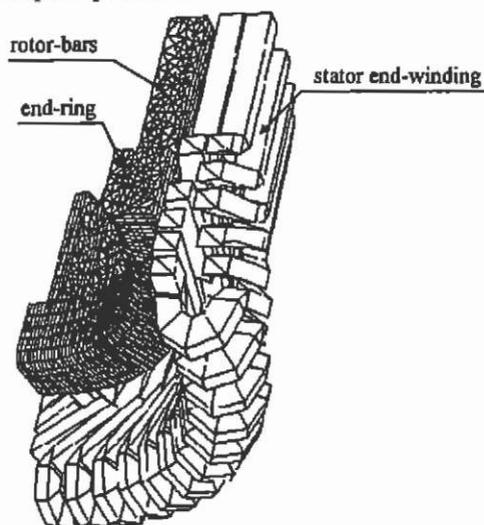
Fig. 6.2. Material mesh of end-ring and bar-ends and coil meshes of the stator end-winding of an induction machine.

Or the classical analytical machine theory delivers models that can be combined with a numerical technique in order to form an overall coupled model of higher accuracy. For example the computation of end-winding effects, using a three-dimensional FEM model, to extract the parameters for an equivalent circuit model, can be seen as an approach of coupled models as well (Fig. 6.2). With respect to computational efforts,

for example for dynamic simulations of motor models or observer models for the machine control, the coupling of those methods is advantageous to obtain an accurate but simple overall model of the machine.

Observing problems in the transient modelling of relative motion of machine parts such as in a rotating motor, a possible solution of this modelling problem can be a coupling of geometries by element types with special properties. Overlapping shape functions can be used to join different meshes of a FEM model and this can be understood as a coupled problem.

A further example of this type of problem, the coupling of measurements with a numerical model, can be given. The basic idea in this type of problem is to measure inaccessible parameters and to use them as input for the numerical field computation. Such parameters are mainly non-linearly dependent on the field quantities and their interdependency from them is unknown. For example material data obtained by measurements are approximated by interpolating polynoms and can be used in this numerical format for the field computations. Look-up tables with measured data samples are possible as well.

6.4 Classification of coupled field problems

After this first more or less subjective judgement of the various coupling mechanisms, in the following discussion the coupled problems are distinguished with respect to physical and numerical aspects. The single involved types or mechanisms of coupling the various fields are described here as sub-problems with specific properties. It will be concluded with a matrix systematic. The matrix entries distinguish between the problem, the model description, the coupling mechanism, a proposed iteration scheme and a proposed method for solving the overall field problem.

6.4.1 Sub-problem extent: domain/interface
The different interacting physical phenomena described by the coupled problem are defined on partially or totally overlapping domains. For example thermo-electromagnetic problems belong to this group. For the electromagnetic problem definition the surrounding air has to be modelled. The same domain is considered in the thermal problem by special boundary conditions such as heat transfer due to convection or radiation boundaries. By using the FEM, different meshes for each sub-problem can be used. The interaction takes place through interface equations. The involved field problems can be numerically strong, i.e. on matrix level, or weak coupled, computed in a cascade algorithm.

In a second class, the interaction of the fields is described by interface equations of the sub-problems. A heat transfer or cooling problem with different models for a heated object and the cooling fluid belongs to the second group. The group determined by overlapping domains is sometimes referred to as "class I" and the interface group as "class II".

6.4.2   Sub-problem discretisation methods: homogenous/ hybrid

It is sometimes advantageous to apply different discretisation methods for the involved fields. The methods used can be the FEM opposed to BEM or FEM methods with different types of elements to result in a hybrid method. Analytical models can be considered. The addition of algebraic equations originating from equivalent circuit models is possible as well. For example, a two dimensional FEM model to compute the temperature distribution inside an electrical machine can be extended by an equivalent thermal circuit model to consider the heat transfer in the axial machine direction. In this way, a quasi three-dimensional approach is obtained by the coupled methods. The combination of different FEM models with an additional analytical model is possible. External electric circuits can be coupled to consider the voltage or current-driven energy source.

6.4.3   Numerical iterative solution methods: full/cascade algorithms

Due to the nature of the physical sub-problems and the chosen discretisation method, differing numerical properties can be linked to the equations descending from the sub-problems. A variety of numerical methods can be chosen to solve the single sub-problem. Most of them can be regarded as block iterative schemes. It is possible to put all the sub-systems in a single matrix, with off-diagonal blocks mathematically describing the (linearised) coupling. This can be considered as a numerically strong and thus fully coupled approach.

On the other hand, several blocks can be solved separately with a well-suited equation solver. Not considering a possible parallellisation, the solution of the sub-problems is usually obtained in successive steps in a "cascade" algorithm. The newly obtained part of the solution can be used immediately in the next step of the iterative process. Other suitable solution techniques are domain-decomposition (DD) algorithms.

6.4.4   Classification matrix

The above remarks on the classification of coupled problems to build up a matrix systematic underline the difficulty of putting all the mechanisms with respect to their different nature into a single systematic. The developed matrix shows couplings between entries in the horizontal as

well as in the vertical direction (Table 6.1). Bi-directional links to other entries are possible as well.

The columns of the matrix represent the mentioned differences of the considered problems with its coupling mechanism. The rows of the systematic represent the proposed types of problem to put into the appropriate columns.

With respect to the geometry, in the first column the studied domains have different properties, such as strong differing material characteristics. The numerical sub-problems are described by partial differential equations (PDEs) and the coupling of the systems of equations is defined by its boundary conditions or interface equations. Depending on the condition of the single sub-problems, a full coupling and weak coupling by cascade algorithms is proposed. For example a hybrid FEM/BEM can be used to solve the overall field problem or in the case of strong differences in the condition of the sub-problems domain decomposition (DD) algorithms, a weak coupling can be employed. Here, an ambivalence of the overall problem can be noticed. Using a hybrid method can be considered as a coupled method and the DD as a weak coupling of physical systems.

The physical nature of the field sub-problems is considered in the second column. Examples of this are coupled magnetic/thermal or other field combinations. The fields can be described either by PDEs or by a combination of PDE and algebraic equations, if equivalent circuit models are used for one of the sub-problems. The coupling is mainly performed by the exchange of the material parameters and source terms or directly by the circuit equations; for example if external electric circuits are considered. For the solution, numerically strong and weak coupled iteration schemes can be applied.

Hybrid methods are put into the third column. The coupled phenomena have different numerical properties. All possible coupled methods such as FEM, BEM, magnetic-, thermal-equivalent circuits as well as the classical analytical field theory coupled to modern numerical techniques, are put to this matrix entry. The model description of the overall problem can be done by coupling PDEs, circuit equations, analytical methods or other methods.

The difference of behaviour in time of the coupled effects considers the last column of the matrix. Here, all the transient problems can be found. Simulations with respect to the differential equation of motion, an ordinary differential equation (ODE) are put into this matrix entry. Various methods are suited to solve such in time-coupled problems.

Table 6.1. Classification scheme for coupled field problems.

problem	geometrical	physical	methods	in time
	domain with different properties	different physical properties • magnetic/thermal • magnetic/thermal/dynamic • magnetic/mechanical • electric/thermal ...	different numerical properties • FEM • BEM • equivalent circuits • analytical models ...	different time constants of the coupled phenomena
model description	PDE/PDE	PDE/alg · PDE/PDE	PDE/PDE, PDE/alg, other couplings	PDE/alg and/or ODE · PDE/PDE and/or ODE
coupling	• boundary conditions • interface equations	• circuit equations • parameter source terms	• boundary conditions • interface equations	$\tau_1 \equiv \tau_2$ · $\tau_1 \gg \tau_2$
iteration scheme	full · cascade	full · cascade	full · cascade	full · cascade
proposed method	FEM/BEM · DD	FEM/circuit equations · FEM/FEM	FEM/BEM · DD, other methods	FEM transient, FEM/BEM transient · FEM/FEM

PDE partial differential equation, alg algebraic equation, ODE ordinary differential equation, DD domain decomposition method

7 Numerical optimisation

The design process of electromagnetic devices reflects an optimisation procedure. The construction and step by step optimisation of technical systems in practice is a trial and error-process. This design procedure may lead to sub-optimal solutions because its success and effort strongly depends on the experience of the design engineer (Fig. 7.1).
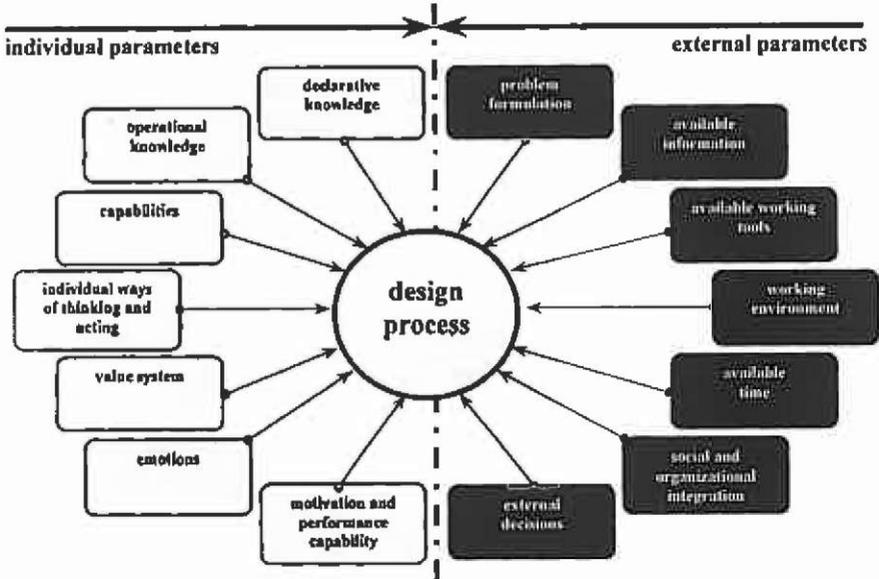
Fig. 7.1. Parameters affecting the design.

To avoid such individual parameters and thus to achieve faster design cycles, it is desirable to simulate the physical behaviour of the system by numerical methods. In order to get an automated optimal design, numerical optimisation is recommended to achieve a well defined optimum.

Optimisation of electromagnetic devices turns out to be a task of increasing significance in the field of electrical engineering. The term of Automated Optimal Design (AOD) describes a self-controlled numerical process in the design of technical products (Hameyer [49]). Recent developments in numerical algorithms and more powerful computers offer the opportunity to attack realistic problems of technical importance (Pahner [84]).

7.1 Electromagnetic optimisation problems

The distinctive feature of this type of optimisation problem is its complexity, which results from a high number of design parameters, a complicated dependence of the quality on design parameters and various constraints. Often the direct relation of the desired quality of the technical product on the objective variables is unknown. Stochastic optimisation methods in combination with general numerical field computation techniques such as the finite element method (FEM) offer the most universal approach in AOD. This section discusses methodology, characteristic features and behaviour of optimisation methods.

To be able to select the appropriate optimisation algorithms to form an overall design tool together with the numerical field computation, the properties of typical electromagnetic optimisation problems will be discussed (Rao [94], Pahner [84]). Electromagnetic design and optimisation problems reflect mainly the following categories:

- constrained
- problem type:
  - $\boxtimes$ parameter- or static optimisation, $f(x) \rightarrow \min$.
  - $\square$ trajectory, or dynamic problem, $f(\dot{x}, x) \rightarrow \min$.
- non-linear objective function
- design variables:
  - $\boxtimes$ real
  - $\square$ mixed real/integer
- multi-objective function
- interdependencies of the quality function and the design variables are unknown; no derivative information available
- the quality function is disturbed by stochastic errors caused by the truncation errors of the numerical field computation method.

In reality electromagnetic optimisation problems are constrained due to the various reasons mentioned in chapter 2.2. Nowadays optimisations are performed mainly as static problems. Numerical optimisations require

huge amounts of computation time. Therefore, the optimisation as aimed at here, combined with the FEM, of the dynamic system behaviour is not yet performed. For transient problems an evaluation of the quality function by numerical methods (FEM) is too time consuming. Considering mixed real/integer design variables results in long computation times as well. The tick boxes in the list that are not marked, represent developments for the future. The optimisation problems that can be solved will grow with increasing computer performance as well.

7.2   Optimisation problem definition

In general, optimisation means to find the best solution for a problem under the consideration of given constraints and it does not mean to select the best out of a number of given solutions. In other words the definition of an optimum is:

> *Define a point $x_0=(x_1, x_2, ..., x_n)^T$ with the independent variables x_1, x_2, ..., x_n in such a way that by their variation inside the admissible space the value of a quality function $Z(x_0)$ reaches a maximum or a minimum. The point x_0 is described as the optimum.*

This definition in mathematical terms:
Minimise a quality function
$$Z(\mathbf{x}) = Z(x_1, \ldots x_n) \to \min.$$

considering

$$g_j(\mathbf{x}) \leq 0 \quad j = 1(1)m$$

$$h_j(\mathbf{x}) = 0 \quad j = 1(1)p \quad .$$

(7.1)

The g_j are called inequality and the h_j equality constraints. Any constraint can be determined in one of these forms. Constraints represent limitations on the behaviour or performance of the design and are called behaviour or functional constraints, whereas physical limitations on the design variables (e.g. availability, manufacturability) are known as geometric or side constraints. If an optimisation problem with only inequality constraints $g_j(\mathbf{x}) \leq 0$ (Fig. 7.2) is considered, all sets of values $\mathbf{x}$ that satisfy the equation $g_j(\mathbf{x}) = 0$ form a (N-1)-dimensional hyper-surface of the design surface, the constraint surface. The constraint surface splits the design surface into two basic regions: the feasible or acceptable region with $g_j(\mathbf{x}) \leq 0$, and the infeasible or unacceptable region with $g_j(\mathbf{x}) > 0$. If, during the progress of the optimisation, a design vector lies on a particular constraints surface, this constraint is called an active constraint.
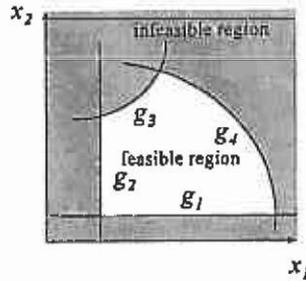
Fig. 7.2. Constraint surfaces in a hypothetical two-dimensional design space, with side constraints (g_1 and g_2) and behaviour constraints (g_3 and g_4).

The independent variables are the design parameter or object variables. Fig. 7.3 shows the shape of a two-dimensional quality function with the global optimum and difficulties such as saddle points and local extremum.
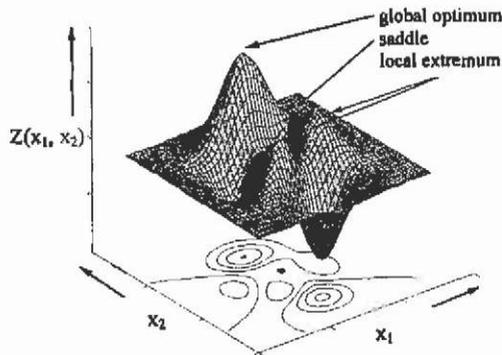
Fig. 7.3. Quality function with two object variables.

To obtain commensurable criteria for the generation of the design variations and to support a simplified formulation of the stopping criteria of the algorithm, the design variables should be transferred into a normalised form:

$$x_j = \frac{x_{j,d} - x_{j,l}}{x_{j,r}} \quad \text{or} \quad x_j = \frac{x_{j,d} - x_{j,0}}{x_{j,0}} \tag{7.2}$$

where $x_{j,d}$ is the original parameter with its given physical dimension, $x_{j,l}$ the lower bound of the parameter variation range, while $x_{j,r}$ denotes the actual parameter variation range. If no lower or upper bound of the parameter is given, the design variable can be normalised to its initial value $x_{j,0}$.

The appropriate formulation of the quality function represents a particular problem. All design aims must be formulated in this single function and all object variables must be implemented. Multiobjective

optimisation extends the optimisation theory by permitting multiple objectives to be optimised simultaneously. It is also known under different names, such as Pareto optimisation, vector optimisation, efficient optimisation, multicriteria optimisation, etc. One way of formulating a single objective function is a weighted linear combination of the q different objective functions:

$$f(\mathbf{x}) = \sum_{i=1}^{q} \gamma_i f_i(\mathbf{x}) \tag{7.3}$$

where γ_i denotes a weighting factor best formulated with the properties

$$\gamma_i \in IR, \quad 0 < \gamma_i < 1, \quad \sum_{i=1}^{q} \gamma_i = 1 \tag{7.4}$$

and $f_i(\mathbf{x})$ are the individual objective functions. In practice, the choice of the weighting factors may already influence the result of the optimisation. It is often not straightforward to select a single fixed weighting factor for each objective, especially if the objective function is erroneous or if no particular preference is given to one of the objectives.

7.3 Methods

In general, numerical optimisation algorithms are iterative methods, constructed to reach the desired optimum in successive steps. This is performed following particular rules to vary the object variables and to determine the search direction. The various algorithms differ only in the choice of step-length, determination of the search direction and in the choice of a stopping criterion. A general form of an optimisation algorithm can be given by applying:

step 0: Choose a start-vector $x^{(0)}$ in the admissible space and set the counter of iteration k=1.

step 1: Evaluate the solution-vector according to a quality function.

step 2: Check whether a stopping criterion is fulfilled. If yes, stop the optimisation; if not, set k=k+1.

step 3: Generate a new solution-vector by variation of the objective variables using a suitable step-length and search direction. Continue with step 1.

With the given properties of the electromagnetic optimisation problems, the requirements of the optimisation algorithms can be formulated. Numerical methods have to be examined with regard to the following criteria:

- reliability
- robustness
- insensibility to stochastic disturbances

- application range
- accuracy
- stable solutions
- performance.

Optimisation algorithms can be classified into:

- deterministic or stochastic and
- direct or indirect methods.

Deterministic methods are basically local optimisation methods, often based on the construction of derivatives or approximations of the derivative of the objective function (Fletcher [40], Bertsekas [10], Rao [94]). Such gradient based methods, e.g. Conjugate Gradient (CG), Newton, Quasi Newton, Broyden-Fletcher-Goldfarb-Shanno (BFGS), etc. are very popular, as they are effective and converge to the local optimum in a small number of steps. This low number of quality function evaluations would be ideal when applying a computationally rather expensive FEM analysis to evaluate the objective function. If no analytical objective function exists or the derivative is difficult to obtain, the use of these methods is not appropriate. Furthermore, these methods are very sensitive to stochastic disturbances, especially present in the derivative information they are based upon. Most deterministic methods additionally require the transformation of a constrained optimisation problem into an unconstrained one. In the case of a multimodal objective function, as is often the case in multiobjective optimisations, these methods are unable to find the global minimum (optimum).

An effective approach to compute the sensitivity information during a FE-analysis is introduced to field computation by Park et al. [86, 87]: the method of adjoint variables. This method was previously successfully applied in electronic circuit optimisation (Director & Rohrer [32], Vandewalle et al. [117]). Here, the sensitivity of the objective function with respect to a set of design parameters can be computed with only two solutions. This basically requires the development of a mesh generator and/or solver specialised for a particular optimisation task (Dappen [29], Ramirez & Freeman [93]).

In general, the human interaction involved in formulating an optimisation problem, in particular in finding the derivatives, is a considerable economical factor when evaluating the efficiency of any optimisation method. The preparation for such an optimisation task might require weeks, while the execution of the actual optimisation run is a matter of minutes. Over the past years, research has been carried out for achieving automatic differentiation of computer codes. The idea is to provide first and higher order derivatives of coded vector functions,

without human interaction. A variety of automatic differentiation software is already available, such as *ADIC* (Bischof et al. [14]), *ADOL-C* (Coleman & Jonsson [26]), *PCOMP* (Dobmann et al. [33]), etc. At present, software code contained in a single file with up to 10.000 lines of code can be automatically differentiated.

Stochastic optimisation methods, on the contrary, such as simulated annealing, evolution strategy and genetic algorithms, do not require derivative information. Any kind of design constraint can be implemented in a simple manner, by just rejecting a design that violates any constraint or by using penalty terms in combination with the objective function. These methods are capable of handling large dimensional optimisation problems and are less sensitive to stochastic disturbances of the objective function value (Kasper [61]). The major drawback of these methods is the large number of function evaluations required when compared to deterministic methods. This fact has, in a first view, an even greater impact when considering FEM based objective function evaluations. The first combinations of the finite element technique and stochastic optimisation methods considered partial models only (Preis & Ziegler [89], Mohammad [80]). One of the first publications reporting the application of a stochastic method to optimise an entire electrical machine is reported by Hameyer [49]. Since then, a large variety of optimisation problems have been solved using the combination of stochastic methods and finite element function evaluation (Palko [85]). Although the plain execution time of such optimisations is large when compared to deterministic approaches, the simplified set-up of the optimisation task and their ability to find the global optimum make such an overall optimisation procedure attractive.

The rather high computational expense of the FEM has always resulted in attempts to reduce the number of function evaluations by applying statistical methods to sample the search space efficiently. A variety of methods can be entitled as indirect, as the optimisation algorithms are executed on an approximation of the real objective function. The combination of the Response Surface Methodology (RSM) and Design of Experiments offers a whole set of statistical tools not only to optimise a design, but also to evaluate the main and interactive effects of the design parameters (Box & Draper [17]). Only a few applications of this method have been reported in electromagnetics in conjunction with FEM function evaluations (e.g. Brandiski et al. [19, 20]). A major drawback of these methods is the fact that due to the use of first or second order (global) polynomials, there is only a remote possibility of finding the global optimum in a search space with several local optima. This problem has recently been relaxed by the application of radial basis functions for

the approximation of objective functions. The first applications, employing the so-called General Response Surface Method (GRSM) have been introduced to the electromagnetics community by Alotto et al. [2, 3, 4]. The experience has shown, however, that these methods are applicable to rather low dimensional problems only, as their practical efficiency deteriorates with a high number of design variables. Other methods utilise the derivative information made available by the approximation based on radial basis functions, as reported in Suykens & Vandewalle [109]. These methods increase the probability of finding the global optimum present in the approximation.

To be able to choose the appropriate algorithm, in this section various methods will be discussed. Based upon the most likely classification of the electromagnetic optimisation problems discussed in the previous chapter, a pre-selection of optimisation algorithms has been derived. Direct search algorithms are selected. No algorithm requires derivative information of the objective function.

7.3.1   Non-stochastic direct search algorithms

In non-stochastic direct search algorithms, the search direction (parameter variation) and step lengths are fixed by a predefined scheme rather than in an optimal way. The advantage is that only the value of the objective function has to be available. No derivative information is required, nor does it need to be constructed from possibly erroneous objective function values.

7.3.1.1 Strategy of Hooke and Jeeves The strategy of Hooke and Jeeves (Schwefel [100]) is a direct pattern search method. This unconstrained optimisation method is characterised by a sequence of two kinds of moves:

- First, an exploratory move in each iteration, consisting of a sequence of single discrete steps per design variable (co-ordinate direction)
- This is followed by a pattern move, being an extrapolation towards an assumed favourable search direction (defined along the line from the initial design in the iteration and the best design encountered by the exploratory moves). An exact description of the extrapolation can be found in Schwefel [100].

Such an extrapolation does not necessarily lead to an improvement of the design. It is merely a guess, based on the pattern of the previously successful moves. This extrapolation determines the name: pattern search. The success of this last move is tested only after the following

exploratory move. The parameter variation step length δ_i of the exploratory moves is decreased in each iteration by the multiplication with a step length adjustment factor α ($0 < \alpha < 1$). Due to this sequential structure, a parallel implementation is impossible. The Hooke and Jeeves algorithm is characterised by the following properties:

- derivative free, unconstrained optimisation method
- dependency of the global convergence on the choice of the starting step length δ_i and step length factor α.

The value of the objective function is required only for detecting the best exploratory move per iteration. A constrained optimisation task might be transformed into an unconstrained optimisation. This is not always possible. If only side constraints are defined in the optimisation task, such a problem can be transformed in a "quasi-unconstrained" optimisation problem. This is achieved by setting the objective function to a very high value in case of violation of the constraints. This infeasible design does not need to be evaluated. This should not be confused with the penalty method. In the penalty method, a violation-dependent value (determined by a function) is applied after the design has been evaluated. If the violation of the constraints leads to a design that cannot be evaluated (e.g. invalid geometry), the application of the penalty method is prohibited.

The Hooke and Jeeves algorithm is favoured compared to supposedly better algorithms such as Rosenbrock's algorithms (Schwefel [100]) and the Nelder and Mead Simplex method, due to the following reasons:

- Nelder and Mead's Simplex method (an unconstrained method as well) requires the ranking of the evaluated designs in the n-dimensional simplex following their objective function value. This ranking decides on the new search direction. The "quasi-unconstrained" mode described above must fail, as all infeasible designs are equally valued.
- Rosenbrock's algorithm may include inequality constraints. However, the approach of the constraint must already be detected, not always being possible, especially in the case of non-linear constraints.

7.3.1.2 A theoretical optimisation example To illustrate the convergence of the selected optimisation methods, and also to allow a comparison of the different algorithms, a two-dimensional optimisation

problem is chosen as described in Dappen [29] and Alotto et al. [4]. This particular optimisation problem is defined by:

$$\text{minimize} f(\mathbf{x}) = \sum_{i=1}^{2}(-0.01((x_i + 0.5)^4 - 30x_i^2 - 20x_i)) \qquad (7.5)$$

subject to the constraints:

$$-6 \le x_i \le 6, \quad i = 1,2 \qquad (7.6)$$

This function is chosen as an example as its objective function can be visualised (Fig. 7.4) including the search paths of the different algorithms. There are four local minima of approximately equal value. The global minimum is located at (-4.454,-4.454). Such an objective function resembles to a large extent typical objective functions found in engineering applications. The starting point of the visually presented optimisations are always (0,0). This point is located on the slope towards the local optimum O3 (Fig. 7.4). Any gradient-based method would converge to O3. However, the stochastic algorithms are restarted from different starting points.
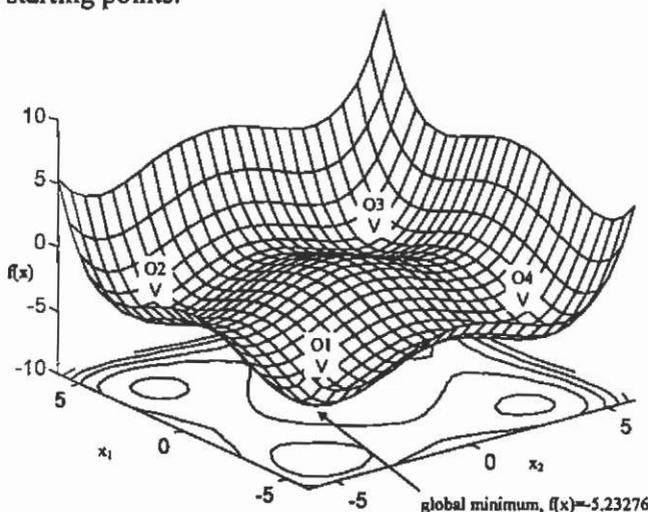
Fig. 7.4. Visualisation of the objective function surface of the optimisation problem defined in (7.5), (7.6) with the optimal solution at $x_{1,best} = x_{2,best} = -4.45377$.

Fig. 7.5 illustrates a successful optimisation run. The initial step size is 6.0 and the step length factor $\alpha = 0.5$. The path connects the best trials per iteration.

Fig. 7.6 illustrates the problem with the Hooke and Jeeves algorithm: there is no guarantee for global convergence, as it entirely depends on a "good" choice of the strategy parameters and the starting point.
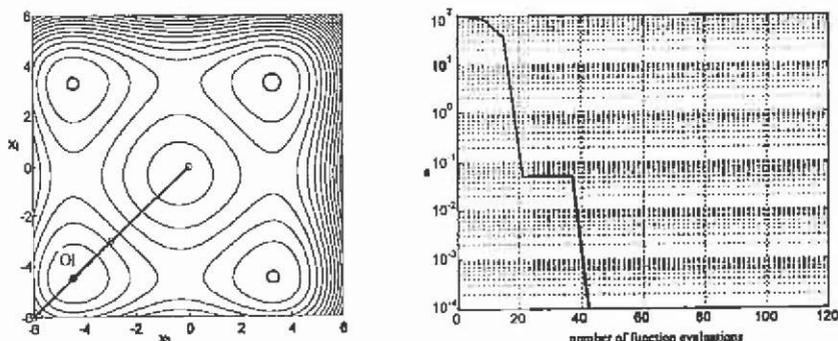
Fig. 7.5. Path of a successful (global convergent) optimisation using the Hooke and Jeeves algorithm and convergence of the error ($x_0(0,0)$, $\delta_0 = 6.0$, $\alpha = 0.5$).
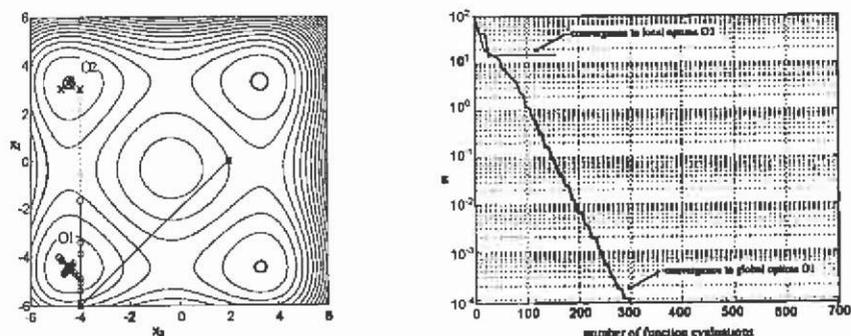
Fig. 7.6. Influence of the choice of the initial step length to find the global optimum. Start at $x_0(2,0)$, $\delta_0 = 6$, $\alpha = 0.5$ (local optimum O2), and $\alpha = 0.85$ (global optimum O1).

7.3.2 Stochastic direct search algorithms

Stochastic search algorithms have steadily gained interest over the past years due to the increase in the computing power available. Various algorithms have already been developed and applied to a wide variety of problems in different fields of science, technology and economics. Simulated annealing serves mainly as a basis for comparison, while the two evolution strategy algorithms are valuable optimisers in combination with finite element function evaluations.

7.3.2.1 Simulated annealing The technical process of annealing solids inspires simulated annealing algorithms. A slow and controlled cooling of a heated solid ensures proper solidification (highly ordered crystalline structure), the state of the minimal internal energy. Rapid cooling causes defects; the system is not in a state of a thermal equilibrium (Rao [94]). In

terms of the optimisation theory, a sequence of steps can be formulated, defining a minimisation process by random changes of the design variables and a probability acceptance criterion for them. An evaluated design vector is accepted or rejected following the Metropolis criterion (Metropolis et al. [79]):

- accept design if $f(\mathbf{x}_{i+1}) - f(\mathbf{x}_i) \leq 0$ and set $\mathbf{x}_{i+1} = \mathbf{x}$,
- otherwise accept with the probability

$$P(f(\mathbf{x}_{i+1}) - f(\mathbf{x}_i)) = e^{-(f(\mathbf{x}_{i+1}) - f(\mathbf{x}_i))/k_b T},$$

with the scaling factor k_b called the Boltzmann constant and T the temperature. Algorithms applying the above probability distribution are called Boltzmann annealing algorithms. It has been proven (Geman & Geman [41]) that such algorithms are guaranteed to find the global optimum if the reduction of the temperature is taken to be not faster than:

$$T_k = \frac{T_0}{\ln k}, \tag{7.7}$$

with T_k the temperature at iteration k. Applying such a logarithmic cooling schedule results in practice in an infinite lasting optimisation process. Therefore, other cooling schedules are chosen, the majority being based on experimental studies for selected types of problems, rather than being mathematically derived (Ingber [57], Hajek [48]). These faster schedules are sometimes called simulated quenching to express the fact that they do not satisfy the sufficiency condition (7.7) to converge to the global minimum. However, they are usually a good trade-off between a fast convergence and high global convergence probability. The most common temperature schedule is a simulated quenching algorithm with the exponential cooling schedule:

$$T_{i+1} = cT_i \qquad 0 < c < 1. \tag{7.8}$$

However, the optimal choice of the annealing factor c is not obvious and depends on the problem type. Typical values vary between 0.98 - 0.80, with decreasing probability of finding a global optimum in non-convex feasible spaces.

The simulated annealing starts at a high initial temperature T_0. It follows the evaluation of a sequence of design vectors, either purely randomly generated or following a scheme, as for instance the evolution strategy. This is continued until equilibrium is reached: the average value of f converges towards a stable value (e.g. by means of an error bound on the standard deviation of f) as i increases. During such a period of constant temperature, a step length vector δ is adjusted periodically. The best design vector is stored. If equilibrium is reached, the temperature is lowered and the above sequence is repeated starting from the active

optimal point. This process is stopped if a sufficiently low temperature is reached and no further improvement of the objective function can be expected.

Except for the definition of the scaling factor k and the scheme to lower the temperature, also the formulation of the probability to accept or reject the evaluated designs have undergone investigations in the past few years in order to accelerate the algorithm by maintaining the property of global convergence. Codes such as Adaptive Simulated Annealing must be named here (Ingber [57]). From the viewpoint here, however, it must be stated that these improvements of the global convergence are of a theoretical nature and have a significant effect only on very large dimensional optimisation problems (100 and more design variables). Therefore, the simulated quenching methods are preferred here.

The advantages of simulated annealing are:
- simple implementation
- slow cooling guarantees high probability of global convergence,
- possibility of solving mixed integer, discrete or continuous problems
- insensitive to stochastic disturbances of the objective function
- simple combination with other direct search algorithms.

The disadvantages are:
- no optimal scheme to determine a sufficiently high T_0 without knowledge of the feasible region
- high number of function evaluations.

The implemented version of the simulated annealing algorithm tries to tackle one of the disadvantages. By adjusting the step length vector δ of the parameter variation automatically in such a way that approximately 50% of all designs will be accepted during one temperature step, the first disadvantage is relaxed. During each temperature step, a number of j step-length adjustment cycles are performed. Each cycle with constant step length consists of N-design variations, changing one design parameter i at a time. A new design is generated by applying a random change (based on the step length vector δ) to the previously accepted design vector x_{pacc}:

$$x_{j,j}^{(k)} = x_{pacc,j}^{(k)} + \delta_i^{(j)}(2 \cdot z(0,1) - 1) \tag{7.9}$$

with $z(0,1)$ a uniformly distributed random number of the interval $[0,1]$. The adjustment of the step length vector per parameter i in (7.9) is based on the design acceptance ratio r per cycle:

$$r_i = \frac{n_{accepted,i}}{n_{cycle}} \tag{7.10}$$

with $n_{accepted,i}$ the number of accepted designs based on changes of variable i and n_{cycle} the predefined number of cycles. The step length δ_i is adjusted after each cycle to:

$$\delta_i = \begin{cases} \delta_i\left(1+\alpha\frac{(r_i-0.6)}{0.4}\right) & if\, r_i > 0.6 \\ \dfrac{\delta_i}{\left(1+\alpha\dfrac{(0.4-r_i)}{0.4}\right)} & if\, r_i < 0.4 \end{cases} \tag{7.11}$$

with α the step length factor, usually equal to 2.0.

To illustrate the performance of the implemented simulated annealing algorithm, the optimisation problem (7.5), (7.6) is solved. The strategy parameters of the SA are chosen as: step length factor $\alpha=2.0$, $c=0.85$, two iterations with two step length adjustment cycles before temperature reduction ($2*2*N$ trials per temperature step). The optimisation is terminated if the difference of the best function value of the present temperature step to the best overall value is less then 5.0e-4, and this is true for two successive temperature steps. A typical optimisation history is shown in Fig. 7.7. The optimisation was repeated thirty times, and the algorithm always converged to the global minimum in O1 requiring typically 600-850 function evaluations.
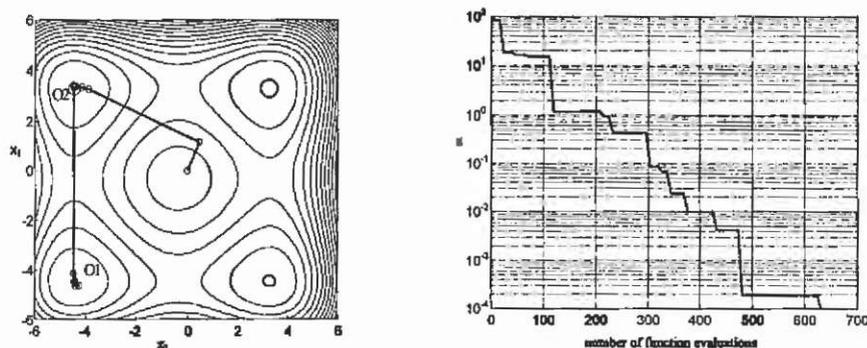
Fig. 7.7. Path of the connected best trials per temperature step. The path illustrates one strength of the algorithm, as it is able to escape from the local optimum O2 that is reached first.

A parallel implementation of SA is possible, but not attempted here, as the implemented SA mainly serves as a comparison tool. Furthermore, a simple parallelisation of the above implementation of the algorithm is

not likely to be efficient, as reported in (Ingber [57]). A better efficiency is reported for clustered algorithms, where an annealing is performed at different temperatures per computational node.

7.3.2.2 Self-adaptive evolution strategy Evolution strategies imitate in a simplified way biological mutation and selection. Evolution strategies slowly started to become popular with the work of Rechenberg [95] and Schwefel [100]. Especially during the past decade, accelerated by the rapid increase of the available computer power, a vast number of developments have been reported (Bäck et al. [5]). In contrast to genetic algorithms, evolution strategies are directly based on real valued vectors when dealing with continuous parameter optimisation problems. While Rechenberg [95] developed a theory for the simple two member (single recombinant) evolution scheme, the theoretic framework of multimember multirecombinant (also called multiparent) evolution strategies, including self-adaptation, is still weak. Bäck, Hammel and Schwefel comment on this in Bäck et al. [5]: "We know that they work, but we do not know why". However, first attempts have been made (Beyer [11, 12]). Even though in these publications the superior progress rate of multirecombinant evolution strategies is highlighted, the single recombinant schemes are still present in engineering research. This is mainly due to their simple implementation, but they are certainly no longer justified when seeking a high performance algorithm.

The implemented evolution strategy is basically the multirecombinant scheme developed in (Schwefel [100]), slightly adapted in (Hameyer [49]). An evolution scheme is based on a population of designs of size λ. The members of this population are created by recombination and mutation from a set of the μ parent design vectors $\mathbf{x}_p^{(t)}$. These parent design vectors are selected following their fitness from the λ offsprings of the previous iteration. Instead of randomly selecting only one of these parents to generate one offspring, multirecombinant schemes are based on the contribution of ρ parent vectors. The factor ρ is therefore termed the sexuality. Our experiments and recent publications (Beyer [11]) have indicated that a sexuality $\rho = \mu$ promises the highest progress rate of the algorithm. Mutation is independently applied to each design vector element as:

$$x_{m,i}^{(k)} = x_{p,i}^{(k)} + \delta_m^{(k)} s_i^{(k)} \quad \text{with} \quad i = 1(1)N \tag{7.12}$$

with k the generation index, p the parent index, m denoting a population member index of the interval $[1,\lambda]$ and $s_i^{(k)}$ the random search direction. The parental designs can be selected either from the present population only ($\rho/\mu,\lambda$-"comma"-strategy) or from the population and the previous

parent set ($\rho/\mu+\lambda$-"plus"-strategy). While in the comma strategies, individual parameter sets are mortal (existing only for one generation), they may survive for the whole optimisation in a plus strategy. Plus strategies are more likely to get trapped in local optima. On the other hand, they feature a higher progress rate (Beyer [12]). Using an intermediate recombination scheme, the step length $\delta_m^{(k)}$ for each population member is constructed on the basis of ρ step lengths of the μ parents. As indicated above, it is advisable to choose $\rho = \mu$:

$$\delta_m^{(k)} = \frac{1}{\rho} \sum_{j=1}^{\rho=\mu} \delta_{\rho,z_j(1,\mu)}^{(k)} \tag{7.13}$$

with $z_j(1,\mu)$ a uniform distributed random integer from the interval $[1,\mu]$ being newly generated for each j.

Furthermore, instead of constructing an individual step length for each design parameter, only one step length is used for the whole set. This requires a normalisation of the parameter. Tests have shown that this single step length speeds up the algorithm, provided there are no strong dependencies among the object parameters. It must be noted, that (7.13) actually defines this evolution scheme to be self-adaptive of σSA-type (σSA stands for σ-self adaptive, with σ the mutation strength) as defined in Beyer[12]. In self adaptive schemes, the mutation strength σ as a strategy parameter is individually coupled with each set of object parameters. Thus the individual step length is stored with the object parameters. To evolve through the course of the optimisation, an additional learning parameter is specified. It determines how quickly and accurately the self adaptation is performed. Beyer[12] reports this self adaptive scheme to achieve optimal convergence velocity for $(1,\lambda)$-strategies if the mutation strength is mutated following either a lognormal distribution or a symmetrical two-point distribution. Here, the intermediate step length (mutation strength) is adjusted following a two-point distribution:

$$\delta_{m,i}^{(k)} = \begin{cases} \delta_{m,j}^{(k)}\alpha & u(0,1] \leq 1/2 \\ \delta_{m,j}^{(k)}/\alpha & \text{if} \quad u(0,1] \geq 1/2 \end{cases} \tag{7.14}$$

with $u(0,1]$ sampling from the random uniform $(0,1]$ distribution and α the step length adjustment factor. The optimal adjustment factor α is problem dependent, usually chosen between 1 and 1.7 with 1.2-1.3 being a good choice for problematic multiminima problems. The adjustment by (7.14) helps in exploring the feasible space in two distinct directions: locally by narrowing the search space for approximately 50% of the offsprings and globally by widening it for the remaining sets. Hameyer [49] supports the latter by enforcing the final step length to follow the

Maxwell distribution. This makes larger step lengths more likely than smaller ones, and should avoid trapping in local optima.

As already indicated above, the choice of the strategy parameters is problem dependent. Based on a large number of tests, involving numerous technical optimisation problems, a possible choice of the strategy parameters is collected in Fig. 7.8. It should be clear, however, that these are empirical data and not necessarily the optimal choice for a particular optimisation problem, nor do they indicate the bounding lines any limits. The choice of the number of parents depending on the population size for optimal progress is supported by the data given in Beyer [11].
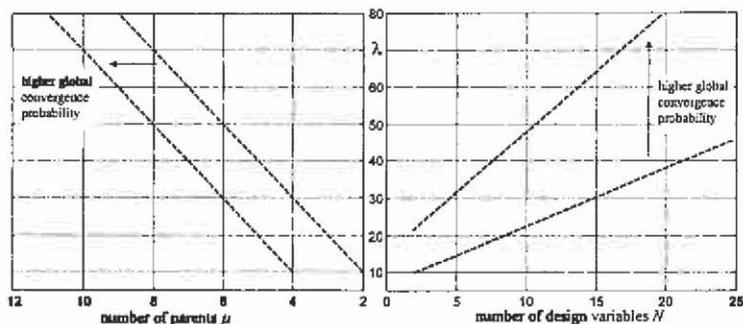
Fig. 7.8. Typical choice of the strategy parameters for the multirecombinant evolution strategy (population size λ related to the dimensionality of the optimisation problem N, and number of parents related to population size). These empirical data are valid only for typical technical problems, featuring locally smooth, non-convex feasible spaces.

Selecting the strategy parameter according to Fig. 7.8 must be accompanied by an appropriate initial step length. The initial step length should be chosen to be larger than the Euclidean distance of the object parameter space. If the design parameters are normalised to the interval $[0,1]$, this is $\sqrt{2}$. If one assumes a well-conditioned optimisation problem, where small changes to the object parameters cause small changes in the value of the objective function, then the change of the step length can be taken as a stopping criterion.

To illustrate the global convergence of the self-adaptive evolution scheme, a sequence of optimisations with different strategies (30 runs each) are performed on the optimisation problem (7.5),(7.6). While the (2/2,8)-strategy globally converges in only 37% of the tests, a (4/4,15)-strategy does so in 95% of the runs. For this theoretical model, the step length adjustment factor α is 1.6.
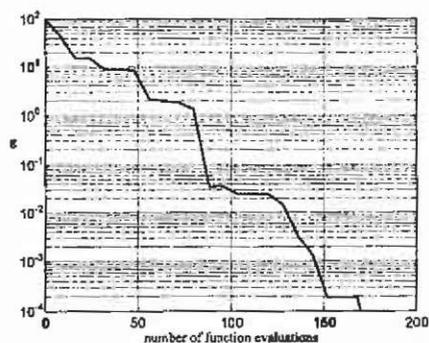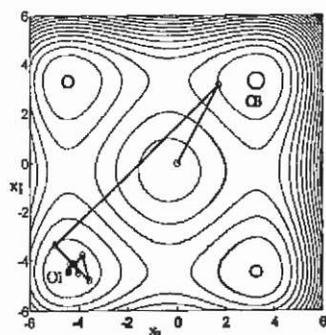
Fig. 7.9. Convergence history (best offspring per iteration) for a successful run with (2/2+8)-strategy.
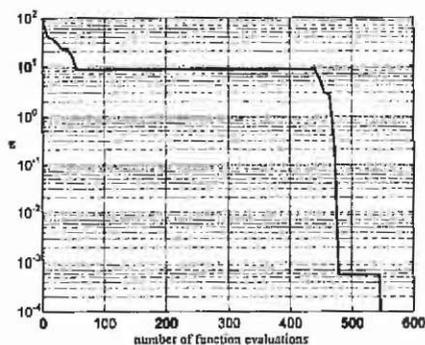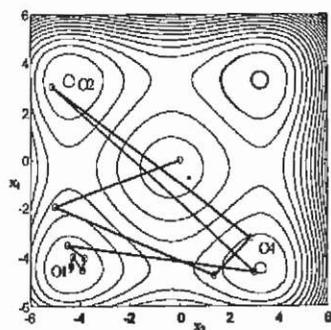
Fig. 7.10. Convergence history of a successful but slow test with a (2/2+8) strategy. The algorithm is trapped in a local optimum (O4) for a considerable number of iterations.
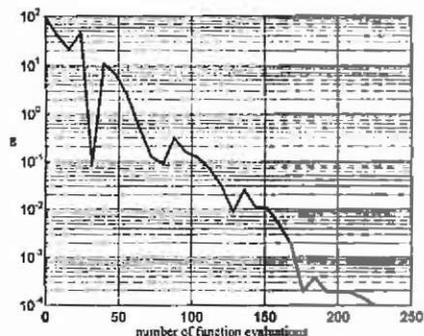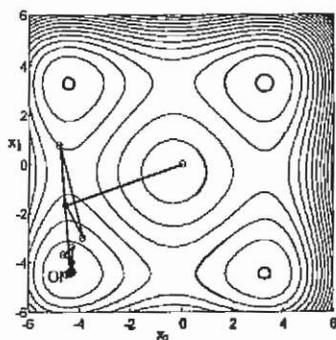
Fig. 7.11. Convergence history of a successful test with a (2/2,8)-strategy. This strategy globally converges in only 37% of the test runs.
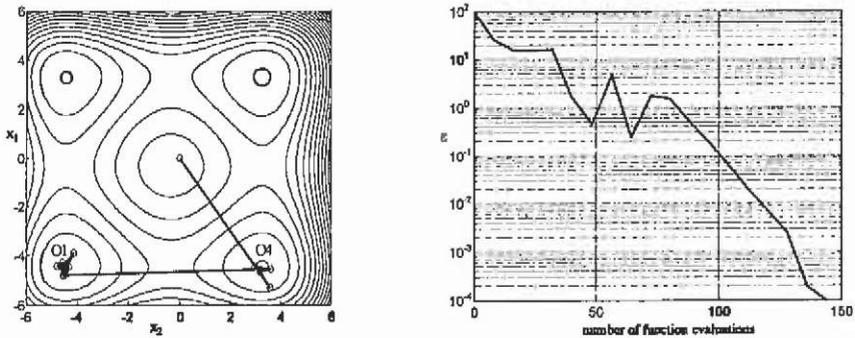
Fig. 7.12. Convergence history of a test with a (4/4,15)-strategy. This strategy globally converged in 97% of the test runs.

The evolution strategies are well suited for parallel implementation, as all individuals in the population can be constructed and evaluated independently. One important property of the evolution strategy must be pointed out, as it forms the base of a sequencing of these algorithms with other algorithms: the evolution strategies converge very fast during the initial steps of an optimisation but they are poor in fine tuning the parameter set. Such an algorithm is well suited for finding the most interesting region of the feasible space at first, being followed by a faster local algorithm to fine-tune the parameter set.

7.3.2.3 Differential evolution Differential evolution (DE) (Storn & Price [107, 108]) is a rather recent approach for the treatment of real-valued multiobjective optimisation problems. As is typical for stochastic search algorithms, differential evolution does not require any prior knowledge of the variable space, nor of the derivatives of the objective functions towards the design variables. The algorithm is very simple, requiring only two control parameters, and is inherently parallel. Differential evolution is a self-adaptive evolution scheme clearly deviating from the σSA scheme outlined in the previous section.

Consider a N-dimensional vector of design variables x:

$$\mathbf{x} = (x_0, x_1, ..., x_{N-1})^T. \tag{7.15}$$

In the initial step, a population of size λ of randomly chosen designs x_i is constructed such that the initial population covers the entire parameter space uniformly. Practically, this could be achieved by defining an initial step length of design variations δ_0, being applied randomly to a given start design $\mathbf{x}_0$:

$$\mathbf{x}_{m,i}^{(0)} = \mathbf{x}_0 - \delta_0 + 2 \cdot \delta_0 \cdot s_i(0,1), \tag{7.16}$$

with $m=1(1)\lambda$, $i=1(1)N$ and the randomly chosen $s_i \in [0,1]$. In the process of the optimisation, DE generates new parameter vectors by adding the weighted difference between a defined number of randomly selected members of the previous population to another member. In its basic strategy, this is the difference of two vectors added to a third:

$$\mathbf{v}_m^{(k+1)} = \mathbf{x}_{r_1}^{(k)} + \alpha \cdot (\mathbf{x}_{r_2}^{(k)} - \mathbf{x}_{r_3}^{(k)}), \tag{7.17}$$

for $m = 1(1)\lambda$ with k the generation index, $r_1, r_2, r_3 \in [1, \lambda]$, randomly chosen and mutually different and $\alpha \in IR, \alpha > 0$. The difference term $(\mathbf{x}_{r_2}^{(k)} - \mathbf{x}_{r_3}^{(k)})$ is obviously similar to the step length δ in the previously described σSA scheme, as well as α finds its equivalence as the adjustment (or learning) factor. The difference is that this step length is not taken from a fitness-selected set of individuals, but rather from a randomly selected individual of the previous population. No explicit step length information has to be maintained. Also, the basis vector $\mathbf{x}_{r_1}^{(k)}$ does not need to be fitness selected. To increase diversity in the population, crossover is introduced, leading to a new parameter vector of the form:

$$x_{m,i}^{(k+1)} = \begin{cases} v_{m,i}^{(k+1)} & \text{for } i = \langle n \rangle_N, \langle n+1 \rangle_N, ..., \langle n+L-1 \rangle_N \\ x_{m,i}^{(k)} & \text{for all other } i \in [1, N] \end{cases} \tag{7.18}$$

The brackets $\langle \ \rangle_N$ denote the modulo function with modulus N. The index n is a randomly chosen integer from the interval $[1,N]$. L defines the number of parameters that are to be exchanged and is taken from the interval $[1,N]$. L is chosen in such a way that the probability $P(L \geq v) = (p_r)^{v-1}, v > 0$, with the crossover probability $p_r \in [0,1]$. In pseudo-code, the determination of L can be expressed as:

```
L=0
do
  L=L+1
while ((rand(0,1)<p_r) and (L<N))    .
```

This recombination scheme is different from the one defined by (7.13) and (7.14). The new parameter vector $\mathbf{x}_m^{(k+1)}$ is checked for violation of any constraints. If it violates one of the constraints, this parameter vector is rejected and the construction of a new vector is repeated. The selection process now has similarity to a tournament selection process. If this resulting design vector yields a better value of the objective function than its predecessor $\mathbf{x}_m^{(k)}$, the new design replaces the old one in the population. If not, the old vector is retained. This

selection scheme is the most distinctive difference from the σSA-evolution strategy introduced in the previous section. Here, fitness is tested against the direct predecessor, whereas the σSA-strategy constructs a fitness selected set of parents from the whole population. Several deviations of this algorithm can be defined, depending on the choice of the vector to be perturbed, the number and choice of parameter vectors considered for the computation of the difference vector and the crossover method (Storn & Price [108]). A good choice for non-critical technical problems is a strategy that increases the greediness of the algorithm by using the best parameter vector from the previous population:

$$\mathbf{v}_m^{(k+1)} = \mathbf{x}_{best}^{(k)} + \alpha \cdot (\mathbf{x}_{r1}^{(k)} - \mathbf{x}_{r2}^{(k)}), \qquad (7.19)$$

or applying mutation to the direct predecessor:

$$\mathbf{v}_m^{(k+1)} = \mathbf{x}_m^{(k)} + \alpha \cdot (\mathbf{x}_{best}^{(k)} - \mathbf{x}_m^{(k)}) + \alpha \cdot (\mathbf{x}_{r1}^{(k)} - \mathbf{x}_{r2}^{(k)}). \qquad (7.20)$$

The construction of the new population member $\mathbf{u}_{i,G+1}$ involves in any case the crossover scheme as outlined in (7.18). The stop criterion for the σSA scheme is the variation of the average step length δ of the selected parents. Such an average step length has to be constructed explicitly for differential evolution, to serve as a stopping criterion. The L_2-norm of the difference vector between each population member and its predecessor is taken.

The experiments conducted in the scope of this book have indicated that the (DE/best/1)-strategy defined in (7.19) is to be favoured for most technical problems. The two remaining strategy parameters can be chosen as $\alpha = 0.5$ and $\gamma = 0.9$. The influence of the population size has been found to be less critical than in σSA schemes. A minimum of $\lambda > 15$ should be chosen always if $N > 2$. However, problems with up to 25 parameters have been successfully solved with population sizes between 30 and 40. Storn & Price [108] advises one to choose $\lambda \approx 10*N$. Here, the authors have found that such a large population is not necessary for the problems classified in section 7.1. Storn & Price [108] use the algorithm to solve artificial optimisation test problems.

DE has been applied to the optimisation test problem (7.5), (7.6) as well. The results are presented in Fig. 7.13. The strategy using $\lambda=10$ is successful in only 68% of the test runs, while tests with $\lambda=20$ succeeded in 94% of the cases within 300 trials. Differential evolution is well suited for parallel implementation.

Both evolution strategies can be used in cascade coupled with the Hooke and Jeeves algorithm. The optimisation is started using an evolution strategy, until a defined accuracy bound or step size variation is reached. Then the Hooke and Jeeves algorithm is invoked for the fine-tuning of the parameter set, assuming that the evolution strategy had

approached the global optimum, allowing a local optimiser to take over. This reduces the overall number of function calls, but might not be permitted in the case of a strongly constrained optimisation problem.
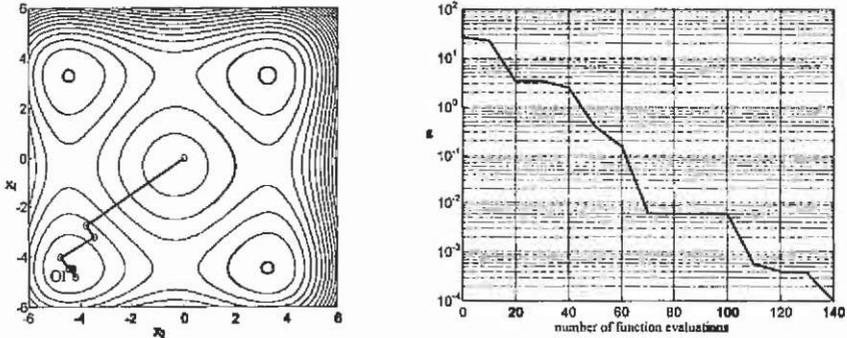
Fig. 7.13. Convergence history of a successful test using the (DE/best/1)-strategy with $\lambda=10$, $\alpha=0.5$ and $p_\gamma=0.9$.

7.3.3   Indirect methods
While in direct search methods the optimisation algorithm samples the objective function directly, in indirect methods the optimisation algorithm is applied to an approximation of the real N-dimensional feasible surface. A trial on a fitted surface is computationally comparatively cheap if it replaces, e.g. a finite element analysis.

7.3.3.1 Response surface methodology and design of experiments The response surface methodology (RSM) (Box & Draper [17]) is such an indirect method. The classical RSM is based on first or second order polynomials. The general form of a second order response surface is:

$$f(\mathbf{x}) = b_o + \sum_{i=1}^{N} b_i x_i + \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} b_{ij} x_i x_j + \sum_{i=1}^{N} b_{ii} x_i^2 \qquad (7.21)$$

with b_0, b_i, b_{ij} and b_{ii} the unknown coefficients that can be found using the method of least squares and N the number of variables (in RSM also called factors). The b_{ij} are called the interaction coefficients, whereas the b_i and b_{ii} are referred to as the coefficients corresponding to the main effects. The immediate advantage of the computation of the coefficients is the information they provide about the significance of individual interactions and main effects. They provide a tool for reducing the dimensionality of the problem by excluding non-significant design parameters from the design optimisation task. This is usually performed using first-order response surfaces. To fit any second-order surface, the number of trials to be computed must be at least equal to the number of unknown coefficients.

There exists a theory to define the sample points to consider. This can be found under the name design of experiments (Box & Draper [17], Montgomery [81], Clarke & Kempson [25]). Experimental designs should ensure an efficient sampling of the region of interest and provide information on the accuracy of the response in the region. To find an accurate second-order fitting surface, three level designs are required (three values per parameter range). This leads to 3^N-samples to be taken for one surface. The research on statistical methods advises a number of samples that are suitable for response surface fitting and require considerably fewer function evaluations than the full 3-level factorials (Box & Draper [17], Clarke & Kempson [25]). Furthermore, they give equal precision for the responses at equal distance from the centre of the factor space (rotatability property). Such recommended designs are:

- Central Composite Design (CCD) (Box & Draper [17])
- Small Composite Design (SCD, special case of CCD) (Draper [34])
- Box-Behnken Design (BBD) (Box & Draper [17]).

Taking the axis experiments a distance α_r from the centre of the region ensures rotatability of the design. But α_r is determined for CCD as:

$$\alpha_r = \left(n_f \right)^{\frac{1}{4}} \tag{7.22}$$

with n_f the number of runs for the 2^N-factorial, leading for the $N=2$ to $\alpha_r = \sqrt{2}$. If the parameters are normalised and bounded in the interval [-1,1], this would be outside the feasible region of the optimisation task. In this case, one takes the outlying axis runs to fit within the feasible space (Fig. 7.14).
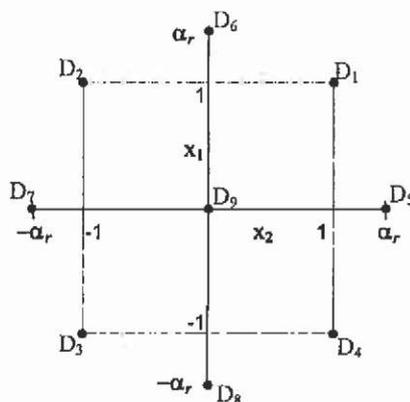
Fig. 7.14. Central composite design on normalised parameter space of dimension $N=2$ consisting of four factorial experiments or runs (D_1-D_4), four axis runs (D_5-D_8) and one centre run (D_9).

A CCD is a five-level design, as it requires the coding of the parameter levels $-\alpha_r$, -1, 0, 1 and α_r. It should be noticed that a CCD for $N=2$ (with one centre run) appears to be simply a rotation of the full 3^N factorial design (Fig. 7.15a). However, going to higher dimensions, the advantage of 2^N+2N+n_c over 3^N experiments becomes obvious.

In the case of the objective function defined by eqns.(7.5) and (7.6), Fig. 7.15 and Fig. 7.16 represent the second order response surface obtained by fitting CCD and full 3^N-factorial designs. The CCD is constructed by taking the 2^N factorial designs (also called runs), $2N$ axial or star runs and n_c centre runs (Fig. 7.14).

The Box-Behnken designs are formed by combining two-level factorials with incomplete block designs (Box Draper [17]). In the two-dimensional case, this would reduce to a two-level factorial with one added centre point, requiring only 5 runs. However, 6 coefficients have to be determined, leaving the least squares problem underdefined. Box-Behnken designs are very effective from 3 factors upwards.
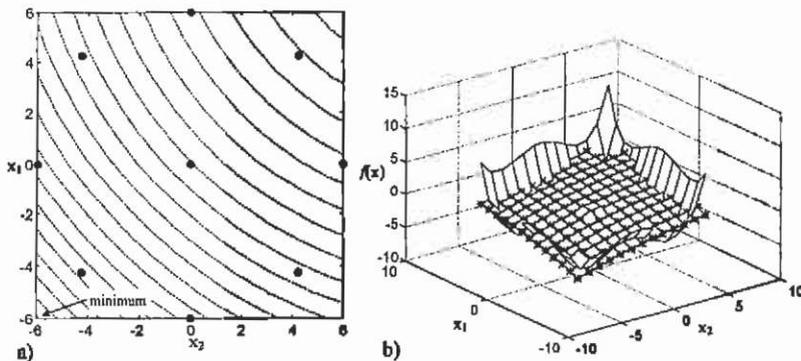
Fig. 7.15. Second order response surface using a central composite design (CCD). Part a) shows a contour plot of the response surface and indicates the positions at which samples are taken, while b) visualises the agreement between response surface and original curve.

As the response surface is constructed by a second order polynomial, an optimisation algorithm can be derived by applying a four-step scheme:

step 1: Construct a second order response surface using designed experiments.

step 2: Use a gradient based algorithm to find the minimum of the response surface.

step 3: Contract the active design space by a defined factor around the new optimum.

step 4: Stop if stopping criterion fulfilled or go back to step 1.
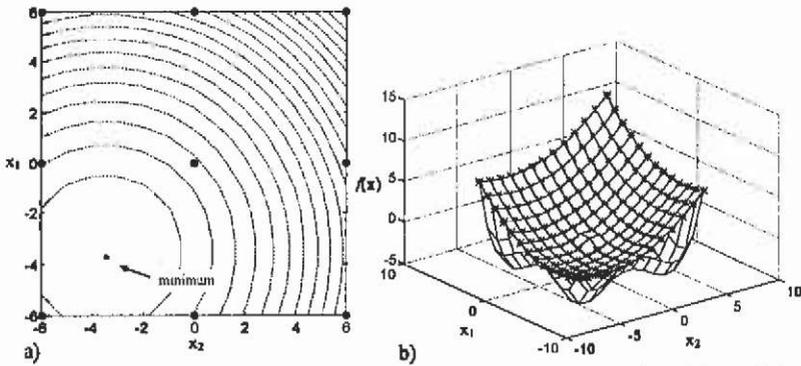
Fig. 7.16. Second order response surface based on a full 3-level factorial. Part a) shows a contour plot of the response surface and indicates the positions at which samples are taken, while b) visualises the agreement between response surface and original curve.

Examination of Fig. 7.15 and Fig. 7.16 reveals the problem of the response surface methodology: the actual fit of the response surface can be very poor. In our case, the minima of both response surfaces are closer to the global one than the initial centre point. However, this cannot be guaranteed, if multiple minima reside in different sized regions of the feasible domain. Brandiski [19, 20] has examined the usefulness of different approaches for multiobjective optimisation in the design of small permanent magnet machines using FE-analysis, such as:

- Select one main objective and reformulate the other objectives as constraints.
- Apply an Euclidean distance function to minimise the distance of the global optimum to all single objective optima (this is comparable to the idea of minimising the sum of the single normalised objectives).
- Apply a desirability function to each single objective (normalising the response to values of the interval [0,1]) and apply a geometric mean to achieve a single scalar value.

Brandiski concludes that both last-mentioned methods are useful, with advantages of the approach using desirability functions. The concept of desirability functions shows similarities to applying the fuzzy set theory to formulate a multiobjective function. These approaches do not solve the inherent weakness of the response surface methodology regarding multiminima functions.

As a conclusion, the response surface methodology using first or second order polynomials has its advantages if the objective function surface features a low curvature, probably having a single optimum only.

Its success in finding the global optimum depends on the choice of the limits of the search region and its contraction scheme. It is interesting, however, to examine the main effects of the variables and their interactions, which can be used to reduce the dimension of the optimisation problem by excluding less significant design variables or to detect unwanted dependencies.

7.3.3.2   General response surface methodology The inherent problem of the response surface methodology to provide "good" fitting response surfaces has led to the introduction of the general response surface methodology (GRSM) capable of finding the global optimum of a multimodal design space (Alotto et al. [3, 4]). The basic idea is the same as in the RSM: find a response surface function and apply an optimisation algorithm to find the optimum of this function. The new feature of the GRSM is the usage of multiquadrics (a special form of radial basis functions) to achieve a response surface including multiple minima. Due to the possibility of including multiple minima, gradient based optimisation algorithms are not permitted. Alotto proposes the use of simulated annealing instead.

GRSM uses approximations of the objective function at any point x_i of the form:

$$f(\mathbf{x}_i) = \sum_{j=1}^{M} c_j h\big(\|\mathbf{x}_i - \mathbf{x}_j\|\big) \qquad (7.24)$$

with c_j the approximation coefficients, M the number of experiments and a possible radial basis function $h(\|x-x_j\|)$ chosen to be:

$$h\big(\|\mathbf{x} - \mathbf{x}_j\|\big) = \sqrt{\|\mathbf{x} - \mathbf{x}_j\|^2 + s} . \qquad (7.25)$$

(In geometric terms $\|x-x_j\|$ denotes the distance (radius) between the two points in the N-dimensional space defined by the vectors x and x_j. This is the origin of the term radial basis functions.) The shift factor s is a parameter defining the curvature of the approximated N-dimensional surface. Alotto presents a statistical method (Bootstrapping (Alotto et al. [4]) to define s in a near optimal way. However, experiments have shown that choosing s smaller than the average spacing of the sample points is sufficient for most applications.

If one substitutes the interpolation condition $f(x_i)=y_i$ in (7.24), the matrix equation for the unique coefficients c_j is obtained:

$$\mathbf{Hc} = \mathbf{y} , \qquad (7.26)$$

with the coefficients of the matrix $H_{ij}=h(\|x_i-x_j\|)$. The matrix $\mathbf{H}$ is a full matrix with all diagonal elements zero. As long as the number of sample points is relatively small (up to a few hundred points) and singularities due to duplicate points are avoided, this system can be solved by any

pivotation method. The approximation function fits exactly all sample points. This is in contrast to the response surface methodology based on second order polygons, in which the least square difference is minimised. However, the GRSM does not give any information on the main and interaction effects as does the RSM, due to the fact that the coefficients c_j are not related any more to a particular design parameter.

The attraction of the classical RSM method is based mainly on the statistical methods that provide interpretable fitted first or second order polynomials. The experimental design developed for this purpose provide the tool to construct these curve fittings efficiently for a well defined polynomial order. The curve fitting based on radial basis functions does not have such a well-defined global polynomial order. Alotto et al. [4] proposes to combine the multiquadrics approximation with the theory of design of experiments with the goal of reducing function calls in comparison to the RSM. This can be successful only if higher level designs are used, resulting in a larger number of experiments. The same number of experiments as designed for the RSM will deliver not much more information with the GRSM (Fig. 7.17).

The advantage of the multiquadrics approximation (on a multimimina objective function) arises with higher factorial designs (Fig. 7.18) or by accumulating the sample points in successive zooming steps. More detail of the original curve is present in the approximation, however at the expense of significantly more sample points. Such a resolution cannot be achieved by a global second order polynomial approximation.
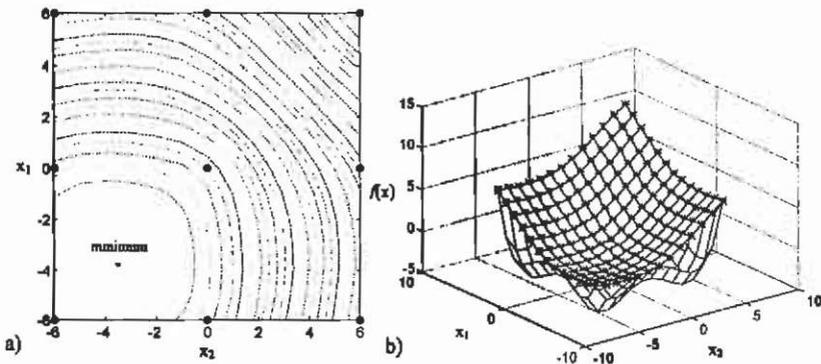
Fig. 7.17. Multiquadrics response surface based on a full 3^N-factorial; a) shows a contour plot of the response surface and indicates the positions where samples are taken, while b) visualises the agreement between response surface and original curve. The result is similar to Fig. 7.16.

Comparing RSM and GRSM leads to the conclusion that the advantage of GRSM is the more accurate approximation that can be achieved when accumulating the sample information over the course of

the optimisation or by higher level factorial designs. Additionally, the approximated surface passes through all sample points taken. The accuracy of the fit depends on the number of sample points if no prior information about the curvature of the design space is available. A global optimum might still be missed if the sample grid is chosen unfavourably or the successive zooming steps narrow the search region too fast. Furthermore, the statistical methods introduced by the design of experiments method, to evaluate main and interaction effects, are mostly not applicable to radial basis functions.
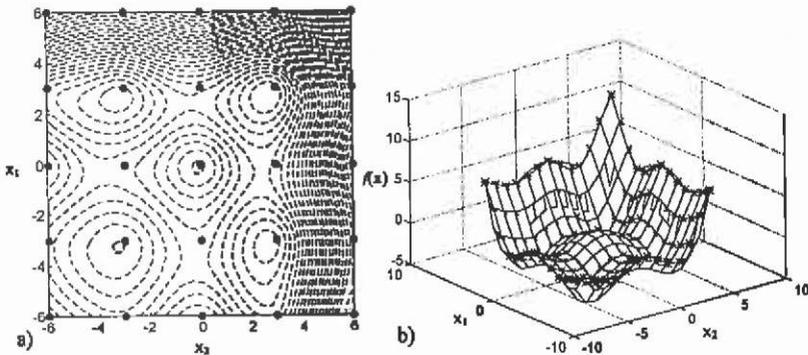
Fig. 7.18. Multiquadric approximation based on a full 5^N-factorial (25 samples); a) shows a contour plot of the response surface and indicates the positions at which samples are taken, while b) visualises the agreement between response surface and original curve.

7.3.4 Adaptive coupling of evolution strategy and multiquadrics approximation

Arriving at this point, the question of a possible combination of the interesting features of the presented methods arises. One of the most significant features of the evolution strategies is their ability to converge fast to the region of interest. The fine tuning of the design, on the other hand, requires a high number of evaluations of the objective function. The presented approximation method based on radial basis functions, on the other hand, is able to approximate multiminima functions with a high accuracy. The accuracy of the fit is not determined by a predefined spacing of the sample grid as in RSM, but rather by sampling "in the right place".

Here, a new approach is introduced, that could be characterised as an adaptive coupling of evolution strategy and multiquadrics approximation. The idea is to combine the above-mentioned features of both methods to reduce the overall optimisation time when using

computationally expensive function calls (e.g. FEM analysis). The basic steps of this method are:

---

step1:   *Start an evolution strategy optimisation using the objective function (direct search iteration).*

step 2: *Construct a multiquadrics approximation f'(x) after each iteration, including only points within a defined radius from the active optimum. This radius is a function of the active step length, ensuring an automatic contraction of the active approximation space.*

step 3: *Compute the next evolutionary iteration using the objective function, but determine the predicted value from the approximation f'(x) as well.*

step 4: *Record each point within a maximal radius $k \cdot \delta$ from the active optimum and construct an updated approximation f'(x). If more than a defined ratio of predicted experiments are accepted function during an iteration, go to step 5, otherwise return to step 3.*

step 5: *Start a new evolutionary iteration, but now using the approximation f'(x) only (indirect search iteration).*

step 6: *Depending on the acceptance ratio determined in step 4, continue an adaptive number of evolutionary iterations using the approximation f'(x) only.*

step 7: *Stop if the stopping criterion is fulfilled; if not return to step 3 and start sampling on the real objective function again.*

This algorithm can be applied to any basic evolution strategy. Here, it has been employed to the differential evolution strategy (DE). There are three levels of adaptivity determining the algorithm:

1. The contraction or zooming of the approximated region is adaptive to the progress of the optimisation by considering a search space with a maximum radius of $k\delta^{(i)}$. The factor k is empirically chosen as:

$$k = k_1(\alpha) \cdot 10.0 \tag{7.27}$$

with α the step length factor of the evolution strategy. For the DE-strategy, $k_1(\alpha) = \alpha/0.5$, ensuring that the approximation region is wide enough for any progress during the indirect iterations (Fig. 7.21). If the step length is rapidly decreased ($\alpha = 0.5$), a smaller approximation region is permitted. If α is larger, the probability that the evolution strategy reaches the boundary of the approximation region in case of bad fitting, is higher. Therefore the chosen region must be large enough. The reason for this special measure is the bad approximation near the boundary of the active region. The points are determined by the evolution strategy instead of a regular grid as in the RSM or

GRSM. This inherits a higher accuracy of the approximation towards the centre of the search space when using the evolution strategy. At the same moment, the shift factor s for the multiquadrics approximation eqn.(7.24) is defined by the average step length δ in the last iteration.

2. The acceptance of the approximation is determined based on the variance of the objective function value of the iteration underlying the active approximation. Accepting predicted experiments with an error less than 10% of the variance has given good results. The error is computed by:

$$\varepsilon(\mathbf{x}_j) = \left| \frac{f'(\mathbf{x}_j) - f(\mathbf{x}_j)}{f(\mathbf{x}_j)} \right|, \tag{7.28}$$

with the trial being accepted based on the variance of the objective function during the last iteration:

$$\varepsilon(\mathbf{x}_j) < 0.1 \cdot \frac{1}{(m-1)} \sum_{k=1}^{n} \left(f(\mathbf{x}_k) - \frac{1}{l} \sum_{i=1}^{n} f(\mathbf{x}_i) \right)^2 \quad . \tag{7.29}$$

3. The number of indirect search iterations only depends on the acceptance ratio achieved with the active approximation. A higher acceptance ratio allows a larger number of iterations on $f'(\mathbf{x})$. Tests indicate that a non-linear dependency of acceptance ratio and number of indirect search iterations n_h is required to ensure a high accuracy of the optimisation throughout the approximation period:

$$n_h = \left(\frac{\dfrac{n_a}{\lambda} - 0.5}{0.5} \right)^2 \cdot 10 + 1 \tag{7.30}$$

with n_a the number of accepted trials per iteration and λ the population size.

Due to the choice of the DE-strategy, particular attention has to be paid to avoid singularity of the matrix $\mathbf{H}$ to solve the approximation equation (7.26). Such singularity is caused by duplicate points in the pool of recorded experiments. In DE, this is possible if the crossover probability is high. In this case, an identical replicate of a previous sample is generated for the new population. Such points are not taken into the sample pool to construct the approximation. The contraction of the approximation region is necessary primarily to reduce the number of unknowns in the matrix equation (7.26).

A further problem that is inherent to all evolution strategies appears when returning from an approximation step: the selection and mutation generates new population members entirely based on members

originating from an approximated function. Their approximated function value is erroneous, probably even lower then the global minimum of the objective function value. The greedy criteria inherent to all evolutionary algorithms could cause these approximated results to survive. To prevent this wrong development, all population members of the last approximated iterations must be mortal, and thus not allowed to survive this first iteration using the objective function evaluation. This causes the new method to require slightly more steps than the classical DE alone would need to converge to the optimum. However, a large percentage of these steps is taken on the approximated function.

The performance of the method is demonstrated on the test example in eqns.(7.5), (7.6). Using the same DE-strategy settings as in section 7.3.2.3 (Fig. 7.13), only 60 objective function evaluations are required, whereas 80 function calls are performed on the approximated function. In Fig. 7.19, the convergence history of the error is shown.
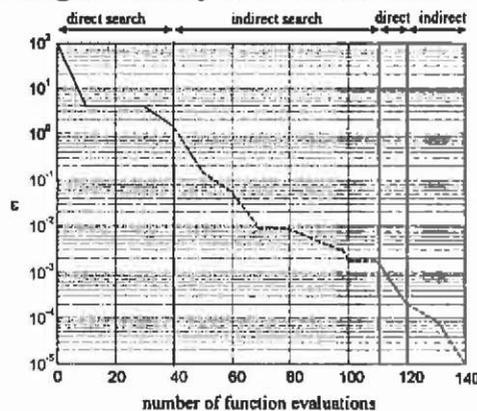
Fig. 7.19. Convergence of the error using the new coupled method based on a (6/10,0.5,0.9)-DE-strategy.

In Fig. 7.20 to Fig. 7.21 the approximation surfaces during the progress of the optimisation are shown. Only 5 multiquadric approximations are calculated. After three iterations (2,3 and 4) the approximation is accurate enough for accepting 80% of the predicted function values. Four iterations using this approximation follow. During this period, the approximations are naturally not updated. These are followed by two direct search iterations, 10 and 11. After these two steps, the updated approximation is accurate enough to return to the approximation again. It must be noted that the probability of finding the global optimum is entirely dependent on the basic strategy parameters of the evolution strategy. However, choosing the population size larger than when using the classical DE-strategy offers the advantage of finding a

better approximation with less direct, and computationally expensive, iterations.
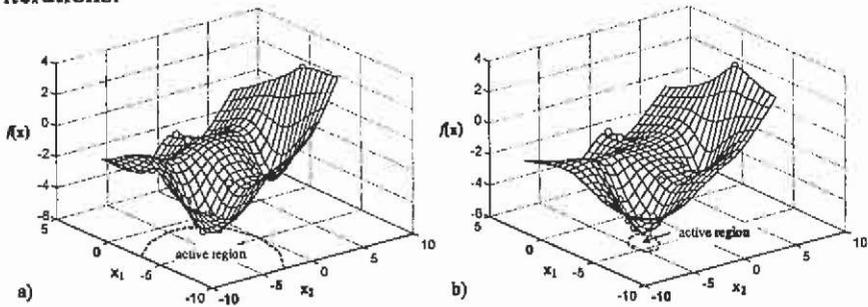
Fig. 7.20. Updated approximations during the initial direct search iteration
a) after 2 and b) after 4 iterations of differential evolution. The approximations
span approximately 10 to 20 times this active search region.
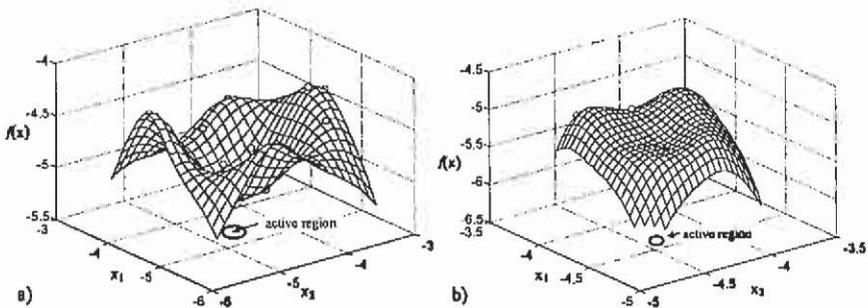
Fig. 7.21. Updated approximations during the final two direct search iterations,
a) after 10, and b) after 11 iterations of the differential evolution algorithm.

The reason for choosing the approximated region considerably larger then the active search region is visible in Fig. 7.21. The approximation accuracy outside the active region is poor. If the evolution strategy reached this region during the indirect sampling, the optimisation might fail. Taking the approximation region wider than the active sampling region of the evolution strategy, as proposed in eq.(7.27), is essential and effectively prevents failure of the method.

A remarkable feature of this proposed algorithm is that the efficiency depends on the curvature of the objective function. Tests with simple second-order surfaces have shown reductions of objective function calls of up to 80%.

It must be mentioned at this point that this algorithm performs well for low dimensional optimisation problems (tested up to 5 on the above objective function and other test functions). Further studies are required to improve and extend the algorithm, especially the adaptation criteria, to provide equal performance for higher dimensional multimodal problems.

8 Linear system equation solvers

There are various methods known to solve a system of linear equations. In this section, we will give an overview of the most important methods suitable to solve equation systems obtained from finite element models of electromagnetic field problems with its special properties. We will consider a system of the form

$$\mathbf{Ax} = \mathbf{b} \qquad . \qquad (8.1)$$

The matrix $\mathbf{A}$ is derived from a linear elliptic partial differential equation (PDE) by using a finite difference or finite element discretisation. Most of the numerical methods that will be discussed in this section are applicable to a wide range of problems. Here, we will focus on the standard example of Poisson's equation $\nabla^2 \mathbf{x} = -\mathbf{q}$.

Discretised by the finite element method, matrix $\mathbf{A}$ will usually be extremely sparse. For the present problem, the matrix is also

- symmetric
- positive definite
- structured.

These are properties that are not always satisfied for more general problems. The solution of linear systems derived from a PDE can be approached from two different points of view:

- algebraic:
  One forgets about the nature of the original problem and deals with (8.1) as an abstract matrix equation. Algebraic methods are usually very robust. They tend to be sub-optimal, however, but easy to apply, and widely available in standard software packages.
- geometric:
  One tries to employ all available knowledge about the mesh and the nature of the PDE and boundary conditions. The system is seen as a collection of discrete finite element equations, whose molecules (stencils) reflect the connectivity of the unknowns and their relative strengths. Geometric methods are more specialised towards one particular class of problems. They are usually of

more limited applicability, often hard to implement but possibly extremely effective.

To study the theoretical aspects of the different equation solver, refer to the literature (Briggs [23], George & Liu [43], Heath [55], Saad [97]). In this chapter it is intended to give a rough overview of possible methods and with the properties of the particular algorithms discussed here, the reader is enabled to choose an appropriate equation solver for his class of problem.

8.1   Methods

The various methods to solve a linear system of equations are traditionally distinguished by direct and iterative methods.

8.1.1   Direct methods
- Gaussian elimination (LU/Cholesky factorisation)
  ⇒ dense factorisation
  ⇒ band factorisation
  ⇒ sparse factorisation
- Fast direct solvers (fast Poisson solvers)
  ⇒ fast Fourier transform techniques (FFT)
  ⇒ cyclic-reduction techniques (CR, FACR)

Direct methods can be characterised as follows:
- They produce a solution to the system of equations in a finite number of operations.
- They are mainly of algebraic nature, and widely applicable under extremely weak conditions (e.g. non-singularity). Some direct methods are geometric, e.g. most fast direct solvers. Certain methods use the geometric information reflected in the matrix adjacency graph, e.g., to determine an optimal renumbering of the unknowns.
- The accuracy of the solution depends on the conditioning of the problem, the numerical stability of the solver and the precision of the computer arithmetic. The accuracy can sometimes be improved by an iterative refinement technique.
- They usually require large amounts of memory, and may be very time-consuming.
- They do not require an initial estimate for the solution. On the other hand, they cannot take advantage of good initial approximations, which are sometimes available.

- The user has little control over the accuracy of the solution. In some cases, a low accuracy solution suffices. This is especially the case for PDE problems, where the discretisation error is often in the range of 10^{-2} (or even worse). Direct methods cannot take advantage of this fact.

8.1.2 Iterative methods

- Stationary (or linear) methods with the iterations:
  $$x^{(k+1)} = Gx^{(k)} + c$$
  $\Rightarrow$ The classical methods:
  - Richardson
  - Jacobi
  - Gauss-Seidel
  - SOR/SSOR

  $\Rightarrow$ PDE-specialised methods:
  - Alternating Direction methods (ADI, LOD)
  - Multigrid (geometric, algebraic, ...)
  - Domain decomposition (Schwarz, Schur, ...)

  $\Rightarrow$ Specialised preconditioners:
  - Incomplete factorisation (ILU, IC, ILQ,...)
  - Approximate inverse preconditioners
  - Polynomial preconditioners

- Non-stationary ('acceleration') methods with the iterations:
  $$x^{(k+1)} = F_k(x^{(k)}, x^{(k-1)}, \ldots)$$

  $\Rightarrow$ Chebyshev iteration/acceleration

  $\Rightarrow$ Krylov subspace projection methods
  - Ritz-Galerkin methods: CG, FOM, ...
  - Petrov-Galerkin methods: CGN
    $\Rightarrow$ Minimum residual methods: GMRES, MINRES, ORTHODIR, …
    $\Rightarrow$ A/AT methods: BCG, QMR, ...
  - Hybrid methods: CGS, BiCGSTAB, TFQMR, FGMRES, ...

Iterative methods have the following qualities and drawbacks:

- Usually, they require little memory: storage for the nonzero matrix elements, the right-hand side, the solution, and perhaps a few additional vectors for temporary storage. In the geometric approach no matrix has to be stored explicitly. The storage for the matrix elements can be integrated into the mesh data structure.

- Often such methods can be implemented as soon as a procedure is available that computes the product of the matrix with an arbitrary vector. Hence, such methods can be applied in cases where one does not want to construct the matrix. This occurs, for example, when solving non-linear problems with the Newton procedure (construction of Jacobian matrix can be avoided).
- When optimally tuned, they usually require much less work than direct methods. For good performance they may require accurate estimation of various problem-dependent parameters, such as the extreme-eigenvalues of the discretisation matrix. They may lose much of their potential performance, or even stop to converge, if the parameters are not chosen carefully.
- They are less robust than direct methods. They often require different conditions on the matrix to be satisfied (e.g., symmetry, positive definiteness,...).
- A small perturbation to the system matrix or to its structure may have a fatal effect on the convergence. In certain cases, one may want to resort to two-level iterative schemes, with the classical iteration used within a block-level iteration.

8.2   Computational costs

Table 8.1 summarises the computational cost of solving an elliptic boundary value problem on a mesh of size $k \times k$ (2D) or $k \times k \times k$ (3D). The formulas are valid for the Poisson equation on the unit square, discretised with second order central differences. For more complicated PDEs, domains, boundary conditions or discretisation schemes some of the methods may not be viable options (e.g. FFT based techniques).

The complexity of a method is often written as a function of n, where n is the order of the coefficient matrix. Table 8.2 states the exponents of n in the computational cost of solving a 2D or 3D elliptic problem. Note that the cost of a dense direct solver increases as n^3, while multigrid solves a problem at a cost proportional to n.

Table 8.1. Computational complexity of direct and iterative methods for the solution of the problem on a rectangular mesh with k mesh-points in every direction (Heath [55], p.344).

method	2D	3D
dense Cholesky	k^6	k^9
Jacobi	$k^4 \log k$	$k^5 \log k$
Gauss-Seidel	$k^4 \log k$	$k^5 \log k$
band Cholesky	k^4	k^7
optimal SOR	$k^3 \log k$	$k^4 \log k$
sparse Cholesky	k^3	k^6
conjugate gradient	k^3	k^4
optimal SSOR	$k^{2.5} \log k$	$k^{3.5} \log k$
preconditioned CG	$k^{2.5}$	$k^{3.5}$
optimal ADI	$k^2 \log^2 k$	$k^3 \log^2 k$
cyclic reduction	$k^2 \log k$	$k^3 \log k$
FFT	$k^2 \log k$	$k^3 \log k$
multigrid V-cycle	$k^2 \log k$	$k^3 \log k$
FACR	$k^2 \log \log k$	$k^3 \log \log k$
full multigrid	k^2	k^3

Table 8.2. Exponent of n (total number of mesh-points) in the computational complexity of direct and iterative methods applied to the model problem Poisson equation (Heath [55], p.345).

method	2D	3D
dense Cholesky	3	3
band Cholesky	2	2.33
sparse Cholesky	1.5	2
conjugate gradient	1.5	2
preconditioned CG	1.25	1.17
full multigrid	1	1

9 Modelling of electrostatic and magnetic devices

The modelling of a technical device is a process of neglecting and simplifying in order to describe the physical device in a mathematical model. The type and quantity of the simplifications influences the accuracy of the solution. In most cases, a choice has to be made between the required accuracy and the numerical size of the model.
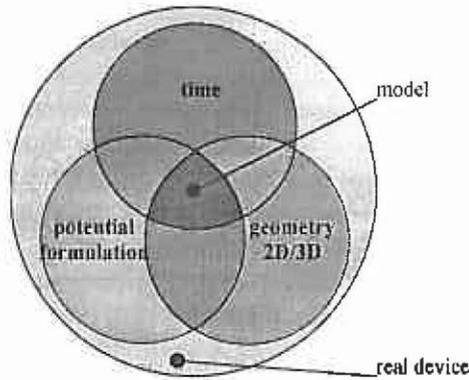
Fig. 9.1. Assumptions.

To define an appropriate field problem, the model of a real electromagnetic device, simplifications have to be made on three levels (Fig. 9.1). The recommended predictions concerning the behaviour of the real device must be translated into a model representation. Thus, assumptions have to be made concerning different issues.

The time dependence of the problem can be of the type:
- static
- stationary changing
- periodically changing, or may be
- transient.

The problem can be considered geometrically as a
- 2D problem that is constant in the z-direction, or as an

- axially symmetrical problem, only consisting of a tangential component, or as an entire
- 3D model.

In the mathematical formulations, simplifications are already embedded. The most important effects of the device must be present in the particular potential formulation for example such as:
- non-linearities,
- hysteresis effects
- couplings to other field types
- couplings of external circuits, ...

Decisions have to be taken to define the model and thus to describe the real device with a maximum of accuracy.

9.1　Modelling with respect to the time

To choose the appropriate solver module of a numerical field computation software package the dependence with respect to the time of the field problem has to be considered. Table 9.1 represents a systematic for choosing an appropriate solver. Time-harmonic or periodic problems can be handled with a transient solver as well. The disadvantage of this approach is the huge computational cost.

Particular attention must be paid to the analysis of electromagnetic devices such as electrical machines. By studying different modes of operation, different problem types have to be defined.

In the following sections the model properties of the various field types are discussed.

Table 9.1. Selection of appropriate solver modules.

electromagnetic fields			
time independent		time dependent	
$\dfrac{\partial}{\partial t} = 0$		$\dfrac{\partial}{\partial t} \neq 0$	
no eddy currents		considers eddy currents	
static	quasi static	periodic (quasi stationary)	transient
static solver		time-harmonic solver	transient solver

9.1.1 Static problems

A static field is a field with constant boundary conditions and is excited by DC currents or voltages. If a DC voltage is applied, the DC current depends on the voltage only by the DC resistance. No induced effects such as eddy currents or induced voltages are present in the model. In a static model, the flux is constant in time and thus the reluctivities are constant as well.

Regarding the static computation as an experiment in the laboratory, a probably varying temperature during the experiment can represent a 'time' dependent factor, if the stationary temperature is not reached yet. Material properties such as the conductivity of a field-exciting conductor change with varying temperature and thus the resulting field as well if a voltage is imposed as excitation. This effect can be taken into account by defining the material at the particular temperatures in the different materials assuming a stationary temperature during the simulation.

9.1.2 Quasi static problems

A quasi-static field is a time-varying field where no eddy currents are involved. The field solution does not depend on the time-derivative term in the differential equations. It can be regarded as a static field for a particular instant of time. The calculation of the field is performed for a certain instant of time, and therefore the flux density and the resulting inductances are calculated for this specific instant of time.

The problem is excited by imposed currents. After the calculation of the coil inductances out of the solution, the induced voltages can be derived.

If the field is driven by imposed voltages, a time-dependent solver has to be chosen to fulfil the voltage law. The applied voltage is the sum of all induced and resistive voltage drops.

An example of a quasi-static field calculation is an instantaneous no-load calculation of a device operated by time-dependent currents where eddy current effects can be neglected, such as an induction machine at synchronous operation.

9.1.3 Time-varying problems

When eddy currents are involved and have to be considered, a time-derivative term appears in the differential equations. The varying field generates induced voltages and currents. The eddy currents are influencing the field.

Each form of time variation can be modelled either in the time domain or in the frequency domain. A time-stepping solver is called a transient solver. When the field is periodic with one or a limited number

of frequencies, it is more efficient to perform a field calculation in the frequency domain instead of in the more convenient time domain. If only one frequency is involved, this solver is called a time-harmonic solver.

9.1.3.1   Transient, time domain A transient solver starts from given starting conditions. Depending on actual field quantities, considering the imposed sources and boundary conditions, the differences of the field quantities are calculated. Adding these differences to the previous field quantities leads to the field at the new instant of time. The time dependence of the field is approximated by a Taylor expansion.

$$A(t + \Delta t) = A(t) + \left. \frac{\mathrm{d}A}{\mathrm{d}t} \right|_t \Delta t + \dots \tag{9.1}$$

9.1.3.2   Frequency domain A solver in the frequency domain represents the field quantities, applied sources and boundary conditions as a summation of phasors rotating at a given angular frequency ω.

$$A(t) = \mathrm{Re}\left\{ \sum_{\substack{n=0 \\ k=1,n+1}}^{n=0} \overline{A}_k . \mathrm{e}^{jk\omega t} \right\} \quad . \tag{9.2}$$

From each phasor, the magnitude and the phase, or the real and imaginary component is calculated. When non-linear materials are considered, the material characteristic varies at given intermediate frequencies.

$$\lambda(t) = \mathrm{Re}\left\{ \sum_{\substack{n=0 \\ k=2n}}^{n=0} \overline{\lambda}_k . \mathrm{e}^{jk\omega t} \right\} \quad . \tag{9.3}$$

9.1.3.3   Time-harmonic problems In this problem class, all field quantities, imposed sources and boundary conditions are assumed to be sinusoidal varying with respect to the time.

$$A(t) = \mathrm{Re}\left\{ \sqrt{2}.\overline{A}_k \mathrm{e}^{jk\omega t} \right\} = \sqrt{2} \cdot \left| \overline{A}_k \right| \cos\left(\omega t + \arg\left(\overline{A}_k\right)\right) \quad . \tag{9.4}$$

For all material properties it is assumed to be constant in time.

$$\lambda\left(\left|\overline{A}_k\right|, t\right) = \lambda_0\left(\left|\overline{A}_k\right|\right) \quad . \tag{9.5}$$

For non-linear material, an effective material characteristic is used. This characteristic gives a sinusoidal averaged value in terms of the rms value of the field quantity.

Regarding the magnitude of the magnetic potentials, it can be noticed that independent sources will cause a potential peak whereas short-connected conductors will try to keep a constant potential. The boundary conditions and the external conditions for the conductors

determine the relative magnitudes of the potentials. The potential is smoothly distributed over the regions without current.

9.2   Geometry modelling

Devices have a three-dimensional geometry. Very often it is a complicated shape or the device contains moving parts. It is possible to reduce the geometrical dimensions to build a FEM model with sufficient accuracy. To obtain a model, symmetries in the device can be used or transformations on the geometry can be performed in order to describe the problem with a simpler discretisation.
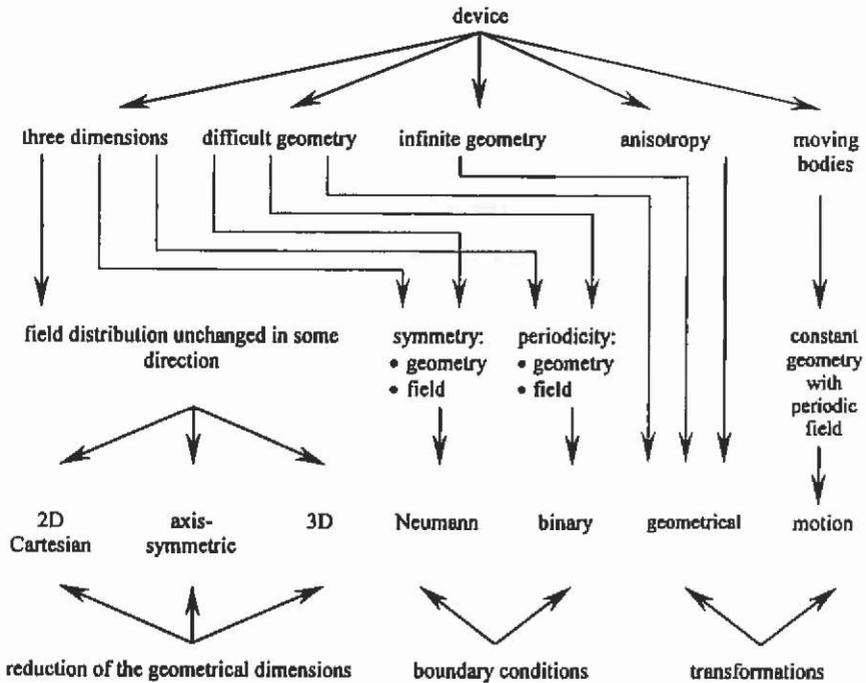
Fig. 9.2. Systematic to reduce the geometrical dimensions of numerical models.

9.2.1   Reduction of the geometrical dimensions

Two-dimensional Cartesian model:

If the device owns the following properties:
1.   the dimensions of the geometry compared to those of its cross-section are long,
2.   the shape of the cross-section remains the same along the length of the model, and

3. it can be assumed that all flux lines are present in the cross-section,

a two-dimensional Cartesian model (Fig. 9.3) can approximate the geometry. Neglecting the end-region effects, typical examples for such models are cylindrical electrical machines or long inductors.
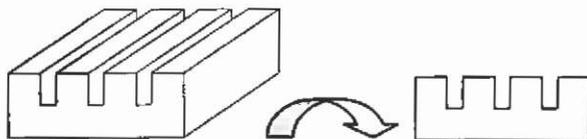
Fig. 9.3. Geometrical reduction from 3D to 2D Cartesian.

Two-dimensional axis-symmetrical model:

If the properties of the device's geometry are true,
1. the geometry has a cylinder symmetry, and
2. the field is axis-symmetrical and is not periodic,

an axis-symmetrical model (Fig. 9.4) can approximate the geometry of the device.
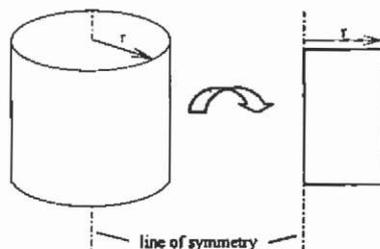
Fig. 9.4. Geometrical reduction from 3D to 2D axis-symmetry.

Entire three-dimensional model:

If it is not possible to reduce the geometrical dimension of the model, the field problem must be solved using a three-dimensional model. Fig. 9.5 shows, as an example, the geometry of an electrostatic micro motor. An axis-symmetrical model can not approximate this geometry because the field is periodic due to the voltage excitation at the stator electrodes. With known periodicity and employing appropriate boundary conditions the 3D model can be reduced.

9.3   Boundary conditions

Symmetric geometry and field

If at the axis of symmetry of a model on both sides the same material is defined, the same sources are in both parts present and the

same boundary conditions are applied, the field and geometry are symmetrical according to the line. Considering a magnetic field, the flux crosses this line orthogonally. The symmetric part can be omitted and replaced by a homogenous Neumann boundary condition (Fig. 9.6). The same idea can be followed in a three-dimensional model (Fig. 9.7).
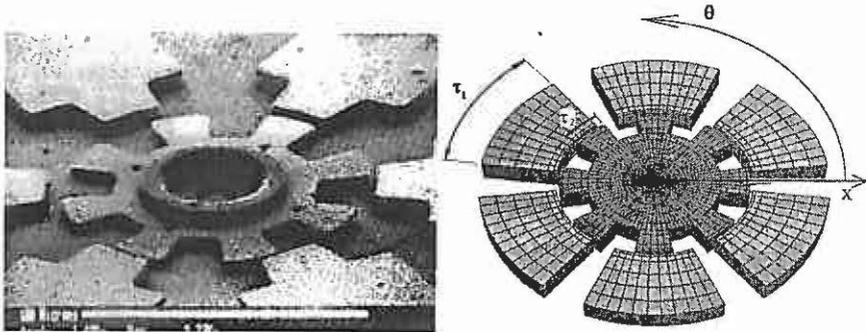
Fig. 9.5. Electrostatic micro motor and entire three-dimensional model.
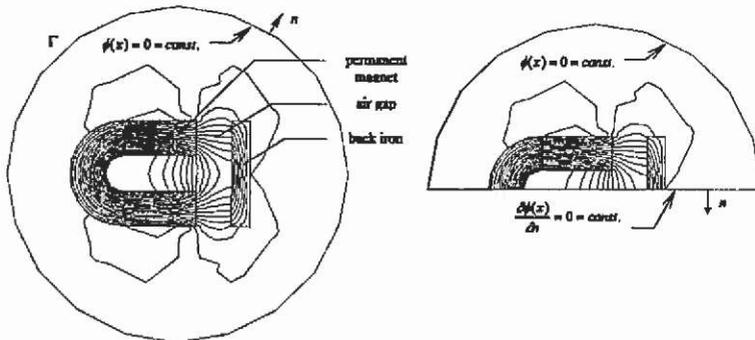
Fig. 9.6. The Neumann boundary condition represents a line of symmetry.

Periodic geometry and field

If the geometry, the applied sources and the boundary conditions are periodic with a given spatial period, the field is also periodic. The smallest common symmetry has to be modelled. The spatial connections are replaced by binary boundary conditions (Fig. 9.8). In Fig. 9.8 the field is periodic with one pole pitch. This is the smallest common symmetry with respect to the field and geometry.

9.4   Transformations

Geometric transformations map real geometrical shapes to arbitrary ones. Depending on the type of differential equation and on the transformation functions, the differential equation can change.
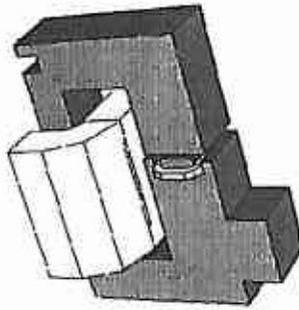
Fig. 9.7. Half-symmetry of the model of a magnetic contactor.
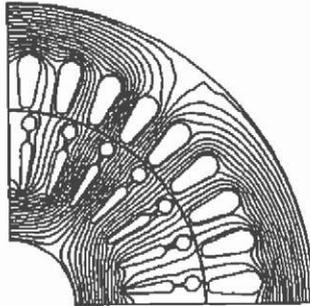
Fig. 9.8. Binary boundary condition applied to reduce the model size to ¼ of the entire cross-section.

Three common reasons for transforming the geometry are:

- to avoid difficulties with the automated mesh generation or refinement,
- to consider anisotropic materials and
- to compute the field of an infinite geometry.

Geometry

When a model causes difficulties during automated mesh generation, the shape of the model can be mapped into a simpler one (Fig. 9.9). The back transformation is done immediately after meshing or after solving the problem. If the problem is solved using the arbitrary shapes rather than the real geometry, the system of differential equations of the model undergoes the same transformation.

Fig. 9.9. Transformation of the mesh of a discretised geometry.

10 Examples of computed models

In this chapter, the modelling of realistic technical devices will be demonstrated on a selected set of examples. All simplifications applied to the models will be motivated and discussed.

10.1 Electromagnetic and electrostatic devices

10.1.1 Synchronous machine excited by permanent magnets

One of the most popular types of electrical machine used for servo drives is the permanent magnet-excited synchronous motor. The properties of this type of machine are a high efficiency and dynamic which combined with controlled inverters offer advantages when compared to other drive systems. High-energy permanent magnet material, such as the rare earth grades Samarium-Cobalt (SmCo) and Neodymium-Iron-Boron (NdFeB), enable various designs that have already been discussed in the literature (Fig 10.1, Fig. 10.2).

magnet material
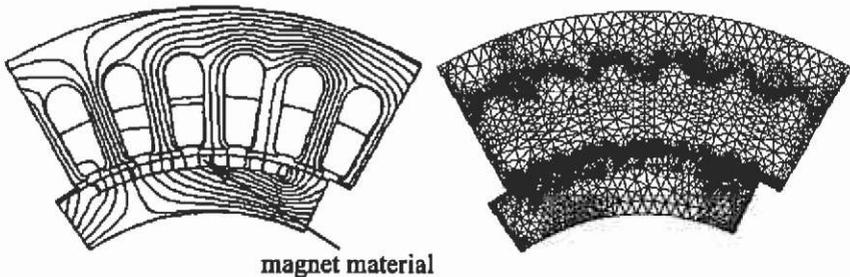
Fig. 10.1. Synchronous machine designs with corresponding FEM mesh of a 6-pole machine with inset magnets.

The previous chapters have shown that the appropriate choice of model simplifications depends on the aim of the analysis. The manufacturer of the machine might be interested in the local saturation level inside the machine, whereas the application engineer requires the

lumped parameters to define the controlled machine's behaviour in the dynamic drive system.
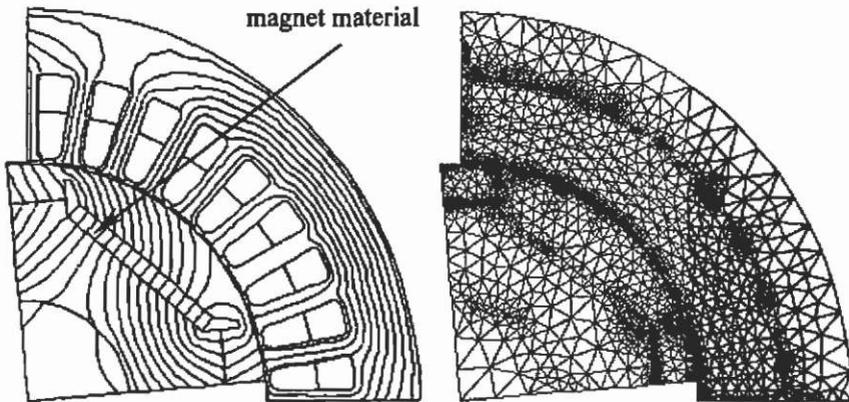
Fig. 10.2. Synchronous machine designs with corresponding FEM mesh of a 4-pole machine with buried magnet system.

It is assumed here, that the d/q-axis theory in the analysis of permanent magnet machines is known. Fig. 10.3 shows the phasor diagram of the studied machine.
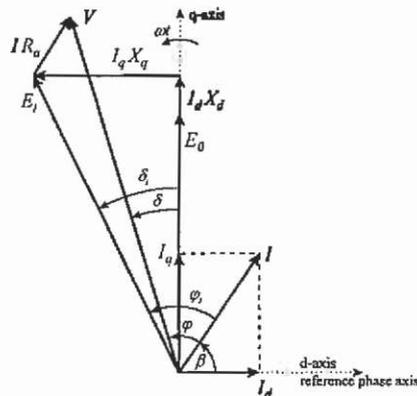
Fig. 10.3. Phasor diagram of the studied permanent magnet machine.

The analysis of permanent magnet excited machines using FEM is performed assuming imposed winding currents. Therefore, end-winding reactances do not have to be considered in a FEM model. Such effects must be taken into account in analytical models which use the lumped parameters that are determined by a FE-analysis.

In this example, the magnetic and geometric periodicities allow the model to be reduced to one pole pitch. This is achieved by applying

binary constraints (periodic constraints) to connect both sides of the pole structure (Fig. 10.2).

10.1.1.1 Static analysis The simplest analysis procedure for this machine is a single two-dimensional non-linear magnetostatic analysis. The aim is to obtain the torque at a certain instant of time for a given current. Induced currents due to the relative motion of stator and rotor are neglected.

Since the stator mmf is in alignment with the winding phase carrying the maximum current, it is convenient to choose the time $t = 0$ in correspondence with the magnetisation axis of this phase. In this case, the initial phase angle β of the stator current is equal to the electrical angle between the stator and rotor mmf. For the steady state conditions at constant load, voltage and frequency, the electrical angle β remains constant. Changing the initial angle can simulate different load situations β. The following data can be derived from the single FEM solution:

- local field quantities such as flux density, magnetic field strength
- the torque (using Maxwell stress tensor method)
- the flux linkage with the stator winding at this instant of time
- stator core iron-loss.

Iron losses

The iron losses can be approximated knowing the loss characteristics, the losses per weight at different flux density levels $P(B)$. Such properties are usually available from the material manufacturer. The total iron loss in the stator for a given frequency can be determined by an integration over the stator core volume (10.1).

$$P_{fe} = \int_{V_{core}} P_r(B) \cdot m_v dV \tag{10.1}$$

In a 2D-problem definition this simplifies to a summation over all 'element losses' of the core cross section:

$$P_{fe} = \sum_{k=1}^{z} P_k(B_k) \cdot \Delta_k \cdot \sigma_k \cdot l_z \tag{10.2}$$

with l_z the length of the device, Δ_k the area of the element k and σ_k the mass density of the material of element k.

Another possibility for the estimation of the losses at different frequencies is provided by:

$$P_{fe} = \sum_{k=1}^{z} \left(c_1 \cdot \frac{f}{50} + c_2 \cdot \left(\frac{f}{50} \right)^2 \right) \cdot B_k^2 \cdot \Delta_k \cdot \sigma_k \cdot l_z \tag{10.3}$$

with f the synchronous frequency, c_1 and c_2 the factors determining the loss density components (hysteresis losses, eddy current losses). The material manufacturer must supply these factors. Estimated values can also be found in the literature.

Torque

As the rotor and the stator magnetic field are in synchronous operation, one can assume that the field pattern is the same for each instant of time.

The relative motion of rotor and stator is not considered. The torque consists of the synchronous torque, the reluctance torque (if d- and q-axis reactance of the machine are different) and a cogging component caused by the interaction of rotor magnets and stator slots. The cogging torque cannot be analysed by a single solution. It is possible to examine this torque component by a small sequence of steady state problems, where the rotor angle and the initial phase angle are changed simultaneously (using electrical angles) over one slot pitch. Cogging torque analysis is very sensitive to errors in the flux density computation in the air gap region. It is therefore recommended to ensure a very fine discretisation (Fig. 10.4) or/and using higher order elements in the air gap region.
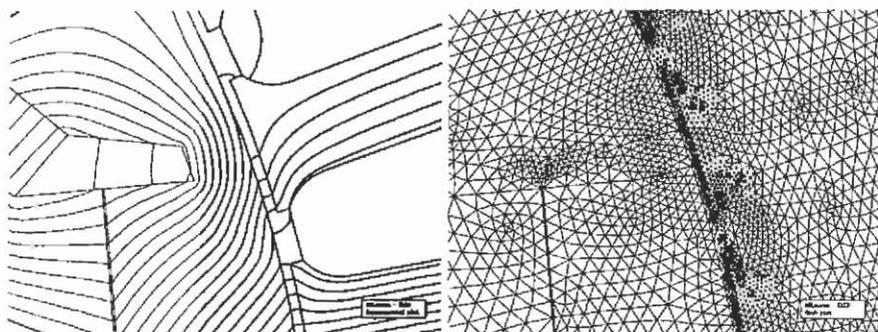
Fig. 10.4. Equipotential plot (load condition) and discretisation close to and inside the air gap of a buried magnet synchronous motor.

The loss computation using the above equations neglects the local variation of the flux density. The described analysis assumes a constant level of saturation at all times. In reality the saturation is shifted locally with the rotation of the magnetic field as well. Resulting higher harmonic fields may considerably increase the iron losses in the stator but are not considered in this approach.

10.1.1.2 Sequential analysis Next to the single analysis, a sequence of steady-state analyses is suggested. This approach is not equal to a

transient analysis including motion. The type of analysis remains the same, non-linear and magneto static.

The mesh used in the single steady-state analysis remains unchanged for this approach. A sequence of analyses has to be prepared, in which the rotor position and the initial phase angle are varied simultaneously (in electrical angles). The current density (according to a defined load situation and instant of time) is applied to the stator winding regions. To generate a sequence of meshes, the following tools can, if available, be used:

- parameterisation of both geometry and excitation with automated re-meshing of the model
- the use of special air gap elements (elements with overlapping shape functions, sliding boundaries) to avoid the re-meshing of the model for every new rotor position
- BEM/FEM coupling; the boundary elements are in the air gap, and the non-linear ferromagnetic parts of the machine are modelled by the FEM
- scripting facilities, which allow creation of new models using combinations of commands normally entered by the user.

As a sequence of analyses is carried out, it is highly recommended that the post-processing is automated as well.

Induced voltage at no-load operation

To evaluate the induced winding voltage, generate a sequence of models with different rotor positions and the permanent magnet excitation only. The offset in rotor positions must be chosen smaller than the slot pitch to consider the slotting effects. Compute the flux linkage with each phase at each rotor position using (5.344) or (5.345). From this result, the induced voltage in the stationary winding is given by:

$$e_i(t) = -\frac{d\psi(t)}{dt} \quad . \tag{10.4}$$

$\psi(t)$ is the flux linked with the entire phase winding, considering all poles. If the FEM model consists only of one pole (Fig. 10.2), this must be taken into account. The rms value of the induced voltage E_i may be determined by integrating the induced voltage time form over one electrical period. The induced voltage may be compared to a no-load generator experiment. The machine is driven by another motor at constant speed and the terminal voltage is measured. The induced voltage from the permanent magnets at no-load is different from under load conditions. The saturation of the flux paths is different in this case. This voltage may therefore not be used to determine the d-axis reactance under load conditions.

Flux linkage

The flux linkage with the stator winding over one electrical period under load condition can be evaluated by generating a sequence of models which covers half a period of the stator current excitation. For the computation of the induced voltage time form apply (5.344), (5.345) and (10.4). The inner torque angle is given by:

$$\delta_i = \beta + \phi_i - \frac{\pi}{2} \tag{10.5}$$

where the phase angle ϕ_i is determined from the phase difference between the induced voltage and the given current (Fig. 10.3). As illustrated in the phasor diagram, the d- and q-axis components of the induced voltage are defined by:

$$E_{iq} = E_i \cdot \cos\delta_i = E_0 + I \cdot \cos\beta \cdot X_{md} \tag{10.6}$$

$$E_{id} = E_i \cdot \sin\delta_i = I \cdot \sin\beta \cdot X_{mq} \tag{10.7}$$

where E and I are the effective values of voltage and current. Whereas obtaining X_{mq} from (10.7) is straightforward, E_0 under load condition cannot be assumed to be equal to E_i at no-load. This assumption leads to very inaccurate results.

Skewing effects

The models described above neglect skewing. The influence of skewed stator slots can be considered in the following way. However, at the expense of higher computational cost, skewing can be taken into account by a rather simple concept:

Analyse a sequence of models that have the equal stator current excitation, but the rotor position is changed over fractions of the skewing angle. Each model serves as a partial model of the entire machine with length $l=l_z/n$, with n the number of models generated per skewing angle.

Cogging torque

While the sequence of models under load conditions is evaluated, the torque may be computed as well. The resulting torque time form allows the evaluation of the cogging torque, provided that the steps between two models in the sequence are chosen smaller than the slot pitch.

Iron losses

The actual iron losses depend on the local change of the flux density. The results given by (10.1) and (10.2) assume a single frequent sinusoidal change of the flux density with respect to the time. The realistic local change of the flux density however is different. Whereas

the vector of the flux density in the tooth head describes an elliptic path over one electrical period, the flux density in the tooth shaft is always parallel with the tooth but changing its direction. This behaviour can be examined by reporting local flux density vectors (describing selected sub-volumes of the iron core) over a sequence of loaded models. The hysteresis losses may then be computed by the summation of the prescribed area in the hysteresis characteristic for all the taken sub-volumes.

10.1.1.3 Loading method The above procedure for the determination of the lumped parameter model of the studied permanent magnet machine is computationally expensive. The saturation-dependent value of the induced voltage generated by the magnets (10.6) cannot be determined accurately.

Therefore, the combination of the finite element method with the analytical calculation of this type of machine is an interesting concept. This method is called the loading method, as all parameters are determined under load conditions, considering the mutual influence between the direct and the quadrature axis fields.

Equation (10.6) is underdefined, as E_0 and X_{md} are unknown. The idea consists of a linearisation around the operating point at a given load. A small change is applied to the load (the stator current), assuming that a small change does not influence the saturation level of the machine. With this assumption the reactance and the induced voltage generated by the magnets do not change. A second equation can be obtained from this linear solution:

$$E_{i_q}^{\cdot} = E_i^{\cdot} \cdot \cos \delta_i^{\cdot} = E_0 + I^{\cdot} \cdot \cos \beta \cdot X_{md} \quad . \tag{10.8}$$

From (10.6) and (10.8), E_0 and X_d can be calculated. Combining this approach with the sequential approach above, E_i and δ_i can be derived from the time form of the induced voltage and the given current time form. Two static problems per time instant have to be prepared for the analysis at a given load. The first one is a non-linear static model, equal to the models prepared for the single steady-state analysis. The accurate alignment of the rotor and stator magnetic axes is very important. The problem is solved and the solution is saved for later post-processing. At the same time, the local values of the permeability (reluctivity) in all elements are saved. A new linear problem is defined, with a small change in the stator current (around 5%). The model is solved using the (fixed) permeabilities from the non-linear solution.

The combination with the sequence approach increases the computational cost. At each instant of time, two FEM models must be solved. However, E_i and δ_i can be determined in a faster way. Both values

can be determined by using a Fourier series of the vector potential A along a contour inside the air gap (near the stator surface).

$$A(\vartheta) = \frac{a_0}{2} + \sum_{\nu=1}^{m}\left(a_\nu \cos\left(\frac{\nu\vartheta}{2\tau_p}\right) + b_\nu \sin\left(\frac{\nu\vartheta}{2\tau_p}\right) \right) \tag{10.9}$$

with ν the ordinal number of the harmonic, m the maximum harmonic considered (higher than the slot harmonics) and τ_p the pole pitch in degrees. Due to the symmetric design of the machine, the a_0-term vanishes. The angular dependency can be transformed into the time domain considering the synchronous speed of the machine. With the alignment of rotor and stator magnetic axis, the cosine term coefficient a_1 in (10.9) represents the quantity of half the q-axis flux per pole and unit depth. The sine term coefficient b_1 represents the quantity of half the d-axis flux per pole and unit depth.

The following data can be extracted from both solutions. The resultant flux per pole is given by:

$$\phi_m = 2 \cdot l_z \cdot \sqrt{a_1^2 + b_1^2} \quad . \tag{10.10}$$

The inner torque angle is:

$$\delta_i = \arctan\left(\frac{b_1}{a_1}\right) \quad . \tag{10.11}$$

The effective value of the induced voltage can now be evaluated using the analytical expression:

$$E_{1i} = \frac{2\pi}{\sqrt{2}} \cdot f_s \cdot \phi_m \cdot N \cdot k_w \tag{10.12}$$

with f_s the synchronous frequency, N the number of series turns per phase and k_w the stator winding factor for the fundamental harmonic (product of winding distribution factor and pitch factor). It should be noted that ϕ_m is a magnetic quantity and represents a magnitude value, but E is an electric quantity representing an effective value. Skewing can be considered by an additional skewing factor that is multiplied with the winding factor.

The d-axis reactance and induced voltage from the magnets under load conditions can be determined by applying (10.9) to (10.12) to both FEM solutions. Inserting the known values into (10.6) and (10.8) yields the value for E_0 and X_{md}. The leakage reactance of the end-winding may be determined analytically or with a 3D end-winding model.

The q-axis reactance is determined by using (10.7) and adding the end-winding leakage reactance.

The entire input power for the m-phase machine is given by:

$$P_1 = m\left(E_0 I_q + I_d I_q (X_d - X_q) + I^2 R_1\right) \tag{10.13}$$

with R_l the ohmic resistance of the stator phase winding at operating temperature.

The electromagnetic power can be obtained by employing:

$$P_e = m\left(E_0 I_q + I_d I_q \left(X_{md} - X_{mq}\right)\right) \tag{10.14}$$

and the electromagnetic torque can be calculated by:

$$T_e = p \cdot \frac{P_e}{2\pi f_s} \tag{10.15}$$

with p the number of pole pairs.

Ohmic losses are considered in the input power (10.13) whereas the iron losses are neglected there. They can be estimated using (10.2) and added to (10.13). Rotor losses are neglected.

It must be noted that this method can not be applied for the overall motor characteristics. X_d and X_q can not be evaluated when I_d and I_q respectively are zero. The underlying equations would require a division by zero.

A harmonic analysis can be performed by introducing harmonic ordinal numbers and dependencies (winding factors for the harmonics). Combining the loading method with the sequencing approach requires a different error discussion as the application of the loading method alone.

Combining the loading method and sequencing approach requires that the post-processing for the loading method is applied after the sequence of solutions is obtained. The computation of the time form of the induced voltage requires the differentiation of the time form of the linked flux. Effects of induced currents are neglected in the analysis.

By applying the loading method only, the quantitative influence of the relative motion of rotor and stator can not be considered because the parameters are determined from one rotor-stator alignment only. However, as the slot harmonics are present in the Fourier-series expression of the potential, they will be present as well in the voltage time form derived from this expression. The exact value of the derived quantities will be erroneous.

The linearisation around the working point, which is the underlying idea of this method, imposes an error. Imposing the permeabilities to the state at the given load decreases this error. The changes in the stator current for the linear model should not exceed 10% of the load current.

The leakage field crossing the slots is not considered using the introduced approach. It could be examined along contours inside the slots, integrating the penetrating flux through the contours. This flux has to be subtracted from the flux computed with equation (10.10). The advantage of the loading method is its low computational expense when compared to the sequence method or to a transient analysis.

10.1.1.4 End-winding reactance The end-winding inductance is a lumped parameter that cannot be computed by a 2D approach. It is also addressed as the end-winding leakage reactance, because the flux linked to this part of the winding is a leakage field. Theoretically, it is possible to compute the end-winding leakage from the difference between the measurements and the computed results using the flux linkage approach. The end-winding leakage inductance can be determined from the comparison of measurements and two-dimensional finite element computation employing:

$$L_{\sigma 3D} = \frac{1}{2\pi f_s} \cdot \frac{(U + IR_a)}{I} \cdot \tan(\delta - \delta_i) \tag{10.16}$$

with δ the measured and δ_i the computed torque angle. As the end-winding leakage reactances are typically small compared to the main reactances (4-10%), this requires a high accuracy of the measured data. Furthermore, for the comparison of measurements with computed operating points, the excitation current, in particular I_d and I_q , are computed from the measured data, resulting in the analysis of a slightly different operating point. The measurement error is amplified.

Another possibility is the determination of the end-winding leakage inductance via the stored magnetic energy in a 3D model of the end-winding region (Fig. 10.5a). Taking advantage of symmetries in the machine, only one pole (or multiples of poles) have to be modelled. The inductance is determined from the stored magnetic energy by:

$$L_{\sigma e} = \frac{2}{3} \cdot \frac{W_{mag}}{i^2} \cdot \frac{2 \cdot 2p}{n_p} \tag{10.17}$$

with n_p the number of poles included in the 3D analysis, $2p$ the number of poles of the machine. While the post-processing is rather simple, the generation of the 3D model requires considerable interactive time and computer resources.

A linear, magneto static analysis can be performed, with only one coil system excited. The face of the non-modelled iron core towards the end section of the machine is constrained in order to generate the effect of infinite permeability of the iron core (Fig. 10.5b) and to reduce the model size. The remaining boundaries are symmetry boundaries, and thus the flux is enclosed. This is an imperfect modelling of the shielding effect of induced currents in the aluminium or cast iron frame of the machine. Furthermore, the accuracy depends on the quality of the discretisation.

10.1.2   Induction motor
The finite element model used for the induction motor analysis is usually two-dimensional (Fig. 10.6) and describes a part of the cross-section of

the motor. Three-dimensional models are applied to compute particular details of the machine such as the end-winding parameters.

a)   b)
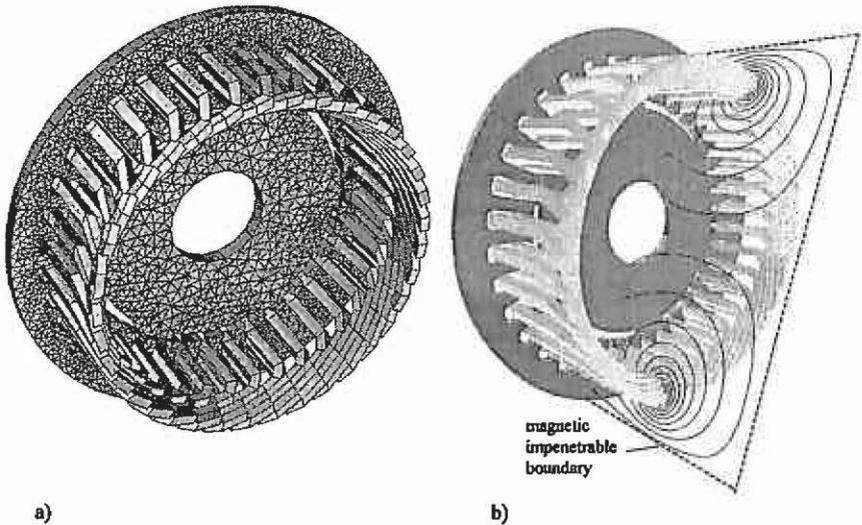
Fig. 10.5. a) FEM mesh of the end-winding of the studied machine with a double layer, chorded winding and b) schematic of the applied boundary conditions using a 2D field solution on a planar cut in the axial direction.
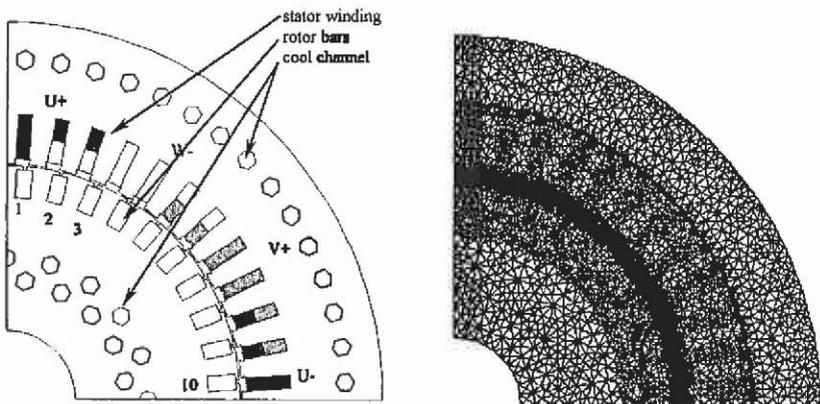
Fig. 10.6. Cross section, winding layout and 2D mesh of the studied 4-pole induction motor.

10.1.2.1 Non-linear time-harmonic problem It is essential for the induction motor analysis to consider both effects, saturation and eddy currents, simultaneously. This requires a computationally expensive transient solution or can be approximated using a time-harmonic solution

in combination with a modified magnetisation curve (method of effective reluctivities) or a combination of static and time-harmonic solutions.

Effective reluctivity

In this method, a time-harmonic solution is recommended. To consider the non-linear characteristic of the ferromagnetic material, the reluctivities are adapted in an iterative process of successive time-harmonic solutions. Using Newton iterations yields in a faster non-linear time-harmonic solver; a relaxation approach is possible as well. Here, the reluctivities are adapted following the iteration scheme:

$$v^{(k+1)} = v^{(k-1)}(1-\alpha) + v^{(k)}\alpha \tag{10.18}$$

$0 < \alpha \leq 1$ is the relaxation factor.

It is assumed that the magnetic field strength is sinusoidal varying $H = H_m \sin(\omega t)$. To consider the AC excitation of the magnetic field and thus the time dependent reluctivities, instead of using the regular BH characteristic, the values of an effective magnetisation curve is chosen.

To calculate the effective characteristic it is assumed that the stored magnetic energy over one period of H must be equal to the energy by using the effective characteristic.

$$\frac{1}{2}\int_0^{H_m} H_m \, dB_m = \frac{1}{T}\int_0^T \left(\int_0^B H(b) \, db\right) dt \tag{10.19}$$

Another possibility for determining the effective magnetisation characteristic is to choose the average value over one period.

$$v = \frac{1}{T}\int_0^T v(t) \, dt \tag{10.20}$$

Combination of time-harmonic and static solution

An alternative approach to determine the element reluctivities for a non-linear time-harmonic problem consists in employing a magneto static solution. With the static solution, the saturation of the operational point is approximated. To define the static problem, the exciting currents and voltages have to be imposed. Their values can be obtained by a time-harmonic solution equipped with reluctivities reached in the static computation. This approach represents an iterative process (Fig. 10.7).

10.1.2.2 Problem definition The time-harmonic problem is defined by:

$$\nabla \cdot (v \nabla A) - j\omega \cdot \sigma A = J_0 \tag{10.21}$$

with $\omega = 2\pi f$ the angular frequency of the problem. In an induction machine the stator currents vary with the stator frequency f_1 while the rotor currents vary with slip frequency:

$$f_2 = s f_1 \tag{10.22}$$

```
        ┌──────────────────────────┐
   ┌───▶│  solve a time-harmonic    │
   │    │      problem P_th          │
   │    └──────────────────────────┘
   │    ┌──────────────────────────┐
   │    │ extract the current densities│
   │    │   ¼ period shifted in time │
   │    └──────────────────────────┘
   │    ┌──────────────────────────┐
   │    │     solve two static       │
   │    │ problems P_real, P_imaginary│
   │    └──────────────────────────┘
   │    ┌──────────────────────────┐
   │    │  extract the reluctivities │
   │    │  ν = (ν_real + ν_imaginary)/2│
   │    └──────────────────────────┘
   │              │
   │         ◇ convergence? ◇
   │              │
   └──────────────┘
              ▼
         ( end of iteration )
```
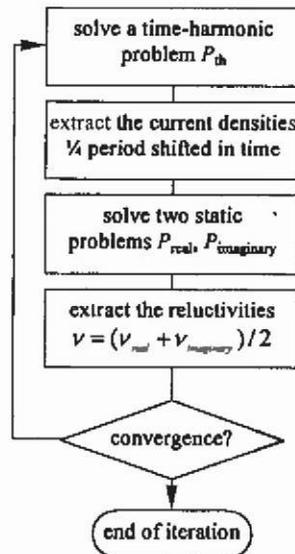
Fig. 10.7. Combined static and time-harmonic solutions to consider non-linear
time-harmonic problems.

Slip frequency

Using imposed currents at slip frequency with given phase angles
can define the problem

$$\nabla \cdot (\nu \nabla A) - j \cdot 2\pi f_2 \cdot \sigma A = J_0 . \tag{10.23}$$

If a voltage-driven problem with rotor frequency has to be defined,
some changes are recommended to obtain a correct solution. The terminal
voltage has to be transferred by multiplying by the slip s and the
conductivity of the stator winding by the factor $1/s$.

Stator frequency

Defining as a function of the stator frequency, the problem is:

$$\nabla \cdot (\nu \nabla A) - j \cdot 2\pi \cdot f_1 \cdot s\sigma A = J_0 \tag{10.24}$$

The solutions obtained with rotor slip frequency f_2 and conductivity
σ deliver the same solution as the problem defined by the stator
frequency f_1 and transferred conductivity $s\sigma$. The latter approach is
preferable because only one parameter, the rotor conductivity, has to be
changed. No additional changes are necessary. The rotor resistance
increases by the decreased conductivity, but the rotor reactance increases
with the same factor due to the higher frequency. The product of
frequency and conductivity determines the induced currents. Losses in
the rotor are a factor $1/s$ considered too low.

10.1.2.3 External circuits To form the short-cut loops of a rotor cage winding, a lumped parameter model representing the end-ring impedances (Fig. 10.8) has to be coupled to the 2D FEM model. Z_B is the end-bar impedance and Z_R the impedance of the end-ring. The resulting external circuit equations are simultaneously solved with the field problem.
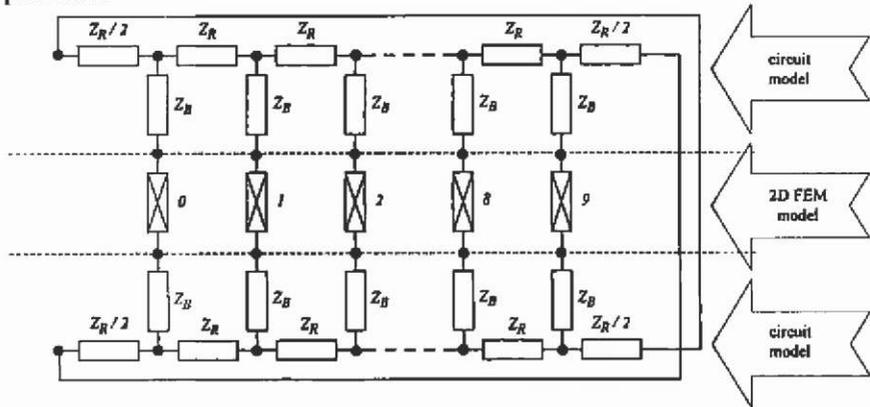
Fig. 10.8. Coupled FEM - external circuit rotor model.

The parameter of the external circuit must be calculated analytically before the field computation or computed by a numerical method.

10.1.2.3.1 End-ring parameter During load operations, the end-ring resistance has the most significant influence on the motor behaviour. Although the resistance of a ring-segment R_R is much smaller than the bar resistance R_B (about 1 % of the bar resistance), its influence can be up to 20 - 30 %.

$$R_{eqB} = R_B + \frac{R_R}{2\sin^2\left(\dfrac{p\pi}{N_2}\right)} \qquad (10.25)$$

The equivalent bar resistance R_{eqB} is the resistance to be used to include the Joule-losses in the ring when calculating the rotor losses due to the bar currents. N_2 is the number of rotor bars and p is the number of pole pairs. To calculate the ring resistance R_R of the entire ring:

$$R_R = \frac{2\pi D_R}{\sigma t(D_R - D_i)}K_R \qquad (10.26)$$

with

$$K_R = \frac{1+\left(\dfrac{D_i}{D_R}\right)^{2p}}{1-\left(\dfrac{D_i}{D_R}\right)^{2p}}\left(1-\left(\dfrac{D_i}{D_R}\right)\right) \tag{10.27}$$

is applied. D_i is the inner diameter of the ring and t the thickness. Current redistribution in the end-ring is not considered. The end-bar resistance is given by:

$$R_B = \frac{\ell_B}{\sigma S} \tag{10.28}$$

with S the cross-section If the length ℓ_B of the end-bar. To describe the influence of the end-ring, the inductance:

$$L_c = \mu_0 \frac{N_2}{3p}\left(\frac{1}{3}\left(l_B - l_{fe}\right) + k\tau'\right) \tag{10.29}$$

is used. With l_B the bar length, l_{fe} is the length of the iron lamination, τ' is the pole pitch diameter in the middle of the ring $\tau' = \dfrac{\pi D_R}{2p}$, $k = 0.18$ for $p = 1$ and $k = 0.09$ for $p > 1$ (Liwschitz-Garik [69]). (10.29) expresses the influence of the end-ring per bar. To compare the analytical formula with the 2D and 3D calculations, expression (10.29) is multiplied by $2\sin^2\left(\dfrac{p\pi}{N_2}\right)$ to refer it to the ring, as done in (10.25) and multiplied by N_2 in order to obtain an expression for the ring inductance L_R:

$$L_R = \frac{\mu_0 N_2^2}{3p}\left(\frac{1}{3}\left(l_B - l_{fe}\right) + k\tau'\right)\cdot 2\sin^2\left(\frac{p\pi}{N_2}\right) \quad . \tag{10.30}$$

The value of the ring inductance of the motor under consideration using (10.30) is $L_R = 0.84\ \mu H$.

10.1.2.3.2   2D/3D computed end-region parameters

Two-dimensional FEM model

For the 2D finite element calculation, a model is made of the axial cross-section of the motor end-region (Fig. 10.9). A 400 kW traction motor is modelled. The accurate modelling of the cross-section is difficult to obtain since it contains a number of different materials with different, often unknown and generally anisotropic magnetic properties. Also for the correct modelling of the boundary between bearings and frame or bearings and shaft, questions arise as to whether to model them as good

or bad magnetically permeable. To overcome these questions, a number of models are built using different boundary conditions and materials in order to find the influence of each component of the end-region.
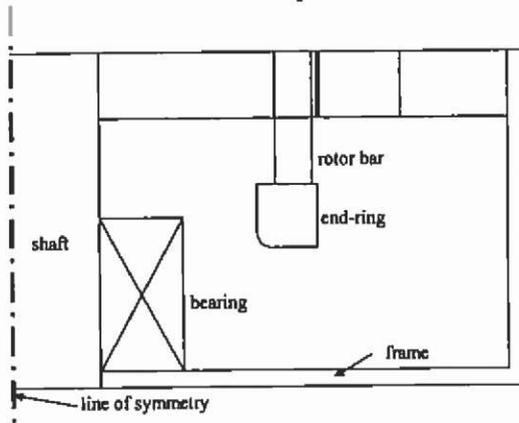
Fig. 10.9. Axial cross-section of a squirrel-cage induction motor end-region.

Two extreme situations are considered, one where the boundary of the frame, bearing and shaft is considered to be a flux line (Dirichlet condition), the other where the boundary is considered perfectly magnetically permeable. The finite element problem is described as axis-symmetric. A unit current is considered to flow in the end-ring. The problem is defined as time-harmonic, neglecting saturation. The following figures (Fig. 10.10, Fig. 10.11) show flux plots for some of the models considering the different boundary conditions and material properties.
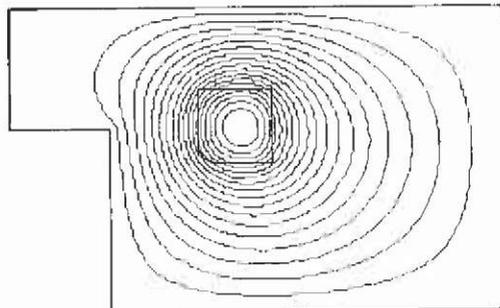
Fig. 10.10. Only the air around the end-ring is modelled and considered being a flux line.

The inductance values obtained from the computations Fig. 10.10 and Fig. 10.11 are 0.35 µH and 0.42 µH. When all boundaries are considered to be magnetically conducting, only the air surrounding the ring has to be modelled (Fig. 10.12).
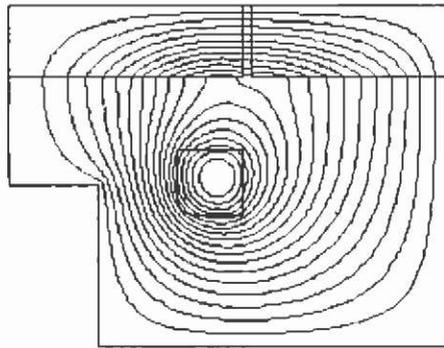
Fig. 10.11. Ferromagnetic stator and rotor iron considered.
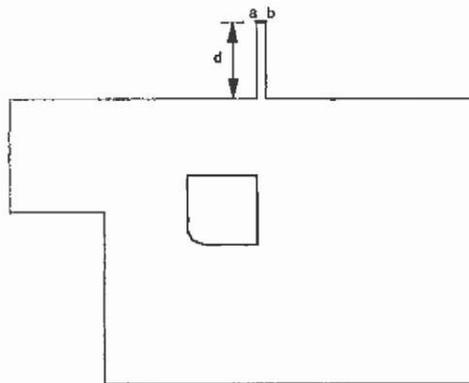
Fig. 10.12. Outline of the finite element model to study the magnetically conducting situation.

Line $\overline{ab}$ in Fig. 10.12 is constrained to be a flux line. It is necessary to have at least one piece of the boundary considered to be a flux line. If not, there is no physical interpretation for the problem. Examining the solutions of this model, a large variation in the inductance is found when the distance d is varied. The model requires d being equal to half the core length. Only then is $\overline{ab}$ a realistic flux line. The inductance value obtained in this case is 13.05 μH. This is a large difference when compared to the other extreme situation where all material boundaries are considered to be a flux line (Fig. 10.10). Therefore, from the 2D approach it is not obvious which value for the ring inductance has to be used as lumped parameter in the coupled finite element-circuit model.

The 2D approach has some additional drawbacks:

- The correct excitation of the ring via the bars can not be accounted for in a 2D axis-symmetric approximation. Therefore, the inductance of the bar-ends outside the iron core are not included in the calculations.

- By exciting the ring in an axis-symmetric problem definition, a flux through the shaft is introduced. Since only the leakage part of the ring-inductance is needed, the coupling with the stator end-winding has to be considered. It is obvious that all flux through the frame links both end-ring and end-windings. Therefore, it is part of the mutual inductance and not part of the leakage components.

To conclude, separating the mutual and leakage components can be performed only using a three-dimensional model of the details in the end-region.

Three-dimensional FEM model
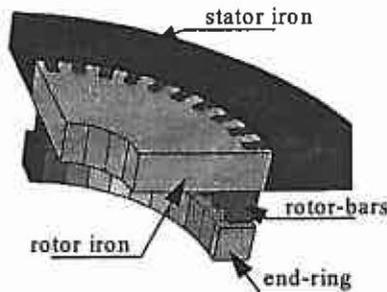
Fig. 10.13. Three-dimensional material mesh generated by extrusion in axial direction.

The 3D model is built using an extrusion-based mesh generator. Due to symmetry, only one fourth of one end-region has to be modelled. The 3D model consists of a material mesh and a set of coil meshes required for the currents. Both meshes are generated separately allowing a different extrusion direction for material mesh and coil meshes. It can be noticed that a part of the iron core is modelled as well (Fig. 10.13).

The stator end-winding is not incorporated in the material mesh. The end-winding is modelled as a set of current-driven coils in air. This is feasible since current redistribution due to skin effect is negligible in the stranded stator end-winding. In the end-ring and the rotor-bars, skin effect cannot be neglected. Therefore, they are considered in the material mesh. Fig. 10.14 shows part of the material mesh (end-ring and bar-ends).

Referring to Fig. 10.14, it is obvious that the generation of the stator end-winding coils requires a more complex extrusion procedure when compared to the modelling of the material mesh. Therefore, the building of such complex models using extrusion techniques is only possible if material and coil meshes can be built separately. Because the stator-end winding is not modelled in the material mesh, the coil meshes have to

represent the actual end-winding geometry as accurately as possible. This is not required for the coil meshes used for the excitation of the end-ring and bar-ends. Because the end-ring and bar-ends are modelled in the material mesh, the coil meshes for exciting them only have to be inside the materials and to provide a path for the current to flow. The current occupies the full material available considering skin effects.
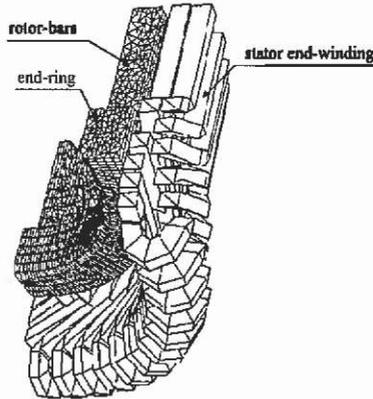
Fig. 10.14. Material mesh of end-ring and bar-ends and coil meshes of the stator end-winding.

Both stator and rotor coils are defined as current driven. In the 3D model, 9 full ring-segments and 2 half-segments are present (Fig. 10.13, Fig. 10.14). Therefore, 11 rotor coils are required for the current excitation. Twenty-two current-driven coils represents the stator winding. Only 2 end-winding coils are completely inside the model; the other 20 coils are cut off at the boundaries of the model (Fig. 10.14). Fig. 10.15 shows two of the coils used for the end-winding excitation, the two coils which are completely inside the model (coil 1 and coil 2), together with three other coils which are cut off at the boundary of the model. When referring to the cross-section of Fig. 10.13, coil 1 occupies the upper half of the first stator slot (the slot in the upper left corner) and the lower half of slot 11; coil 2 occupies the upper half of the second slot and the lower half of slot twelve. In the real motor, each of the stator coils contains four turns. The currents for both rotor and stator coils are obtained from a two-dimensional finite element analysis.

Five inductances can be calculated:

L_R end-ring inductance
L_s end-winding inductance
M   mutual inductance between end-ring and end-winding
$L_{\sigma R}$ end-ring leakage inductance
$L_{\sigma s}$ end-winding leakage inductance
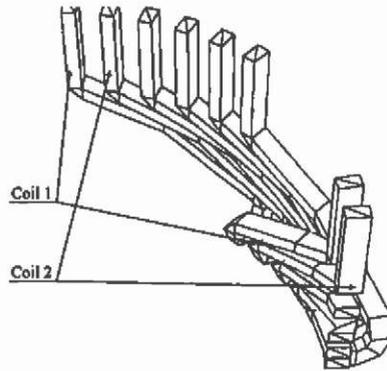
Fig. 10.15. Coil meshes representing the end-winding model.

The inductances are calculated based on the stored energy in the model. To determine both leakage components and the mutual inductance, three problems have to be solved: one having only the stator coils excited, one having only the rotor coils excited and one with both stator and rotor coils excited. Considering the stored energy in these problems to be W_{m1}, W_{m2} and W_{m3}, the following equations can be given:

$$W_{m1} = \frac{m}{2}\frac{1}{8}L_s i_1^2$$

$$W_{m2} = \frac{1}{2*4}L_R i_R^2 = \frac{m}{2}\frac{1}{8}L'_R (i'_R)^2 \tag{10.31}$$

$$W_{m3} = \frac{m}{2}\frac{1}{8}L_s i_1^2 + \frac{m}{2}\frac{1}{8}L'_R (i'_R)^2 \pm \left(\frac{m}{2}\frac{1}{8}M_{Rs} i_1 i'_r + \frac{m}{2}\frac{1}{8}M_{sR} i_1 i'_R \right)$$

where

L'_R	end-ring inductance referred to the stator,
i_R	ring current,
i'_R	ring current referred to the stator,
i_1	stator current,
M_{Rs}, M_{sR}	mutual inductance between stator end-winding and rotor end-ring,
m	number of phases (m = 3).

Since the rotor values are referred to the stator, $M_{Rs} = M_{sR} = M$. The value W_{m3} is less than $W_{m1} + W_{m2}$ (the negative sign has to be applied in the expression for W_{m3}) if the flux caused by the rotor excitations opposes the stator flux. If the rotor flux supports the stator flux, W_{m3} is larger than $W_{m1} + W_{m2}$. Under regular conditions, the rotor flux opposes the stator

flux, resulting in $W_{m3} < W_{m1} + W_{m2}$. The division by eight or four in (10.31) considers that only one fourth of the end-region or one fourth of the end-ring is modelled.

In the case where only the mutual inductance has to be calculated, it is sufficient to solve two problems, the first with the rotor flux opposing the stator flux, and the second with the rotor flux supporting the stator flux. The difference of both stored energy values is only a function of the mutual inductance and the applied currents. Introducing the end-ring leakage referred to the stator, L'_{or}, the different inductances can be calculated by:

$$
\begin{aligned}
L_s &= W_{m1} \frac{2}{m} \frac{8}{i_1^2} & L_{\sigma s} &= L_s - M \\
L_B' &= W_{m2} \frac{2}{m} \frac{8}{(i'_B)^2} & L'_{\sigma B} &= L'_B - M \\
L_B &= W_{m2} \frac{2*4}{i_B^2} & L_{\sigma B} &= L_B \frac{L'_{\sigma B}}{L'_B} = L_r \left(1 - \frac{M}{L'_B} \right) \\
M &= \left(W_{m3} - W_{m1} - W_{m2} \right) \frac{8}{3 i_1 i'_B}
\end{aligned}
\tag{10.32}
$$

Because the three-dimensional calculations performed are assumed to be linear, the number of calculations is reduced. The calculations with only the rotor- or only the stator-coils excited, are performed only once with a unit current. From both calculations, L_s and L_R are obtained. Using (10.31), the values for the stored energy W_{m1} and W_{m2} for the actual currents i_1 and i_R (or i'_R) are obtained. Therefore, for each studied slip value, only one additional calculation is required with both rotor and stator coils excited.

Some results of the calculations of the different inductance components are collected in Table 10.1. It can be noticed that both mutual and leakage components are strongly slip dependent. It can be stated that during no-load operation, the full stator end-winding inductance L_a represents a leakage reactance. Therefore, the end-winding leakage inductance $L_{\sigma S}$ varies from 0.49 mH (at no-load) to 0.28 mH (at standstill). The mutual inductance increases when the slip increases; the end-winding leakage inductance and the end-ring leakage inductance decrease with the same amount. Furthermore, the end-ring leakage referred to the stator, L'_{or} is found to be of the same value as the end-

winding leakage inductance $L_{\sigma S}$. This already is an indication that the influence of the-end ring leakage is not negligible at all load situations.

Table 10.1. Calculated inductance components for different slip values.

inductance [mH]	slip [%]		
	0.34	1.42	100
L_s	0.49	0.49	0.49
L_R'	0.46	0.46	0.46
L_R	0.65e-3	0.65e-3	0.65e-3
M	0.13	0.19	0.21
$L_{\sigma R}'$	0.33	0.30	0.25
$L_{\sigma R}$	0.47e-3	0.42e-3	0.35e-3
$L_{\sigma s}$	0.36	0.30	0.28

To predict the behaviour of the induction machine accurately, particular attention to the calculation of the resistance of the end-region of the squirrel-cage is recommended. It was assumed to have uniformly distributed current in the end-bar and end-ring (10.26). Computations using a 3D model result in a higher accuracy, and point out that the assumption of a uniformly distributed current is not valid (Fig. 10.16). It can be noticed that eddy current effects are present in the parts of the bars located inside the iron core and that a non-uniform current density distribution is computed in the end-rings.
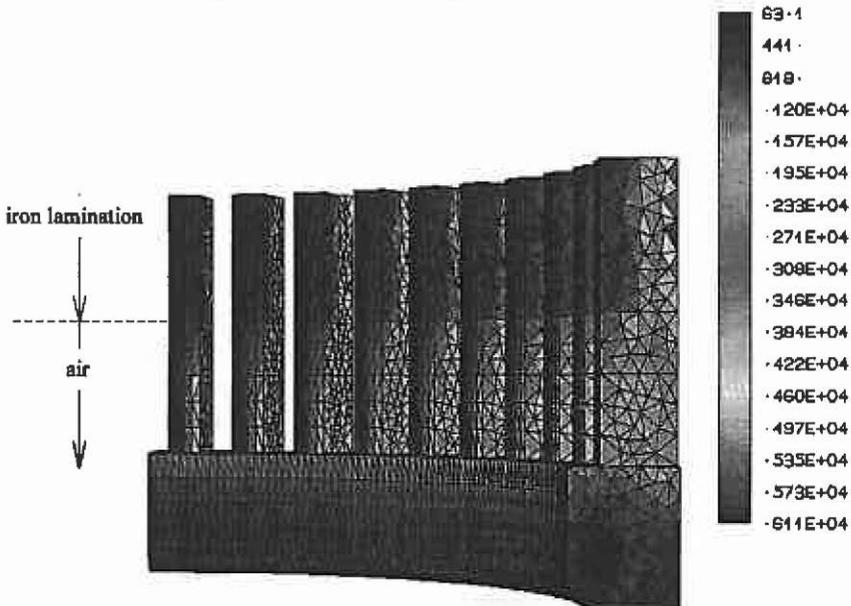
Fig. 10.16. Current density distribution in the end-ring region at start-up.

10.1.3 Electromagnetic shielding

Electromagnetic compatibility (EMC) is becoming more and more important in the design of electromagnetic devices. The product must satisfy international and national regulations and standards. The developer of a product has to be aware of two items: the generation of electromagnetic waves by his product, and the influence from outside fields to his product. This chapter will discuss the second issue, the problem of electromagnetic shielding. Magnetic fields can be shielded by applying two phenomena:

- high permeable material and/or
- induced currents.

Both problems will be discussed here. Studying shielding problems with the FEM analysis raises the difficulty of a sufficiently fine discretisation of the considered domain. The skin depth (penetration depth) of magnetic fields into high permeable or conducting material plays an important role in the pre-processing of the FEM model. The choice for 2D or 3D modelling will not be discussed, as this decision is based mainly on the model geometry.

Depending on the type of shielding (permeable material/eddy currents), the penetration depth of the field into the structural elements must be estimated beforehand. The estimation of the penetration depth can be performed by:

$$\delta = \sqrt{\frac{2}{\omega\mu\sigma}} \tag{10.33}$$

with the permeability μ, conductivity σ and the angular frequency ω. If the penetration depth can be discretised sufficiently finely to allow a potential variation that approximates the exponential change of the field, the full structure can be modelled. Practically, this is the case if at least 3 first-order elements can be defined in the skin depth. The best choice is a FEM analysis using mesh refinement. However, sometimes it is impossible to discretise the model finely enough to cover large extensions of the modelling domain. For instance, in the case of a transformer shielding, the model dimensions are in the range of metres, whereas the dimensions of shielding walls are in millimetres. In this case, special formulations may be applied, like thin-iron-plate elements or impedance-boundary conditions.

10.1.3.1 Thin iron plate elements If the thickness of an iron wall is too fine to be discretised with finite elements, thin iron plate elements provide a way to model such structures. In 2D these are line elements and

in 3D they are surface elements (triangular-shaped using tetrahedrons as the 3D finite elements). With such elements a material characteristic and a thickness is associated. Within the thin-plate element there is no component of magnetic flux density perpendicular to the plate. In reality this will be the case if a low permeable material such as air surrounds a high permeable material such as iron. For the time-varying field analysis, the material is assumed to be non-conducting.

The magnetic shielding of a monitor serves as an example. The analysis is performed to compare measurements taken in a test rig. This test installation employs a set of Helmholtz coils to generate a homogeneous field of defined direction in which the monitor is placed (Fig. 10.17).
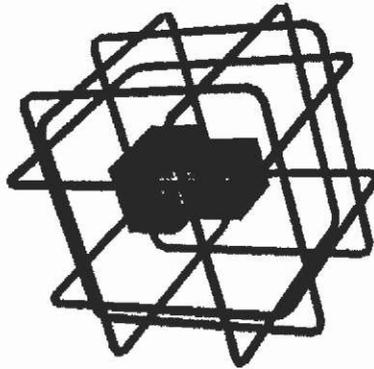
Fig. 10.17. Monitor located inside a set of 6 coils generating a homogeneous field of defined direction.

The shape and position of the set of coils requires a 3D analysis. The extent of the volume span by the set of coils is 1 metre. The surrounding air is modelled up to 6 metres. The iron-plates are modelling the surface of the monitor housing. They are assumed to have a thickness of 3 mm and are surrounded by air. A magnetostatic 3D analysis is performed. The Helmholtz coils are excited with currents producing a homogeneous field directed in parallel to the view-axis into the screen of the monitor. This is the worst case, as the screen area cannot be shielded.

Local field quantities outside the thin plates can be derived in the usual way using the form functions of the element types used. Field quantities inside the thin plate elements require a special post-processing and the results may be viewed using a surface mapping technique.

No field component perpendicular to the plate is considered inside the element. This could introduce large errors if the thickness of the material is in the range of the other dimensions of the plate, or if the surrounding material's permeability is of similar value.
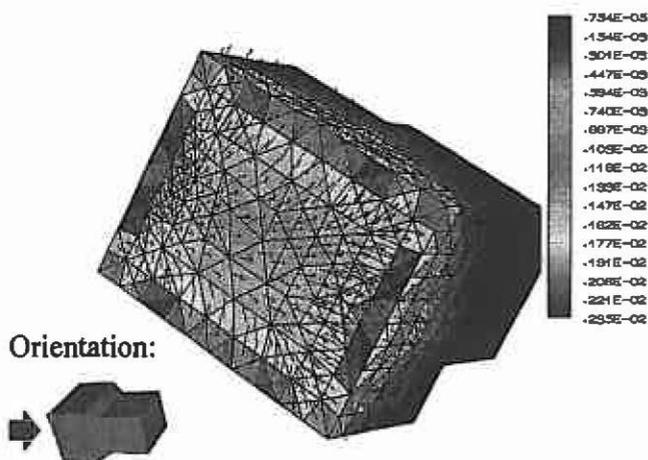
Fig. 10.18. Magnitude of the flux density inside the monitor nearby to the shielding.

10.1.3.2 Impedance boundary condition Whereas the thin-plate elements can be used only for non-conducting, highly permeable material, impedance boundary conditions may be used to model conductors with a small skin depth. They can be applied only for time-harmonic or transient problems.

The idea is to provide a boundary condition specifying the ratio of the electric to the magnetic field, at a surface, to be equal to a complex number. It is assumed that the actual distribution of the field inside the material, which is replaced by this boundary, is not of interest. Such boundary conditions may be used when the skin depth is relatively small compared to the size of the conductor.

The value of the impedance to replace the conductor can be estimated by:

$$Z_s = \frac{1+j}{\sigma \cdot \delta} \tag{10.34}$$

with σ the conductivity of the conductor and δ the skin depth estimated with (10.33). This can be performed automatically by the FEM program, if implemented, or can be provided by the user.

Local field quantities inside the model can be derived in the usual way by using the form functions of the element types used. Field quantities inside the replaced conductor require special post-processing and the results may be viewed using a surface mapping technique. It is possible to compute the ohmic losses in the replaced conductor as well.

10.1.4   Permanent magnet mini-motors

The design of mini motors requires the use of advanced three-dimensional field analysis methods to obtain the field distribution and subsequently the elements of the equivalent circuit and the torque.

Very small motors based on the electromagnetic principle are excited by high-energy rare earth permanent material such as NdFeB. The overall dimensions of such motor devices are found in a range of some millimetres.

10.1.4.1   4-pole motor with block shaped magnets The studied motor is from the axial flux type, equipped with an etched planar double layer winding in a double stator system (Fig. 10.19). In order to avoid cogging torques an air gap winding is used.
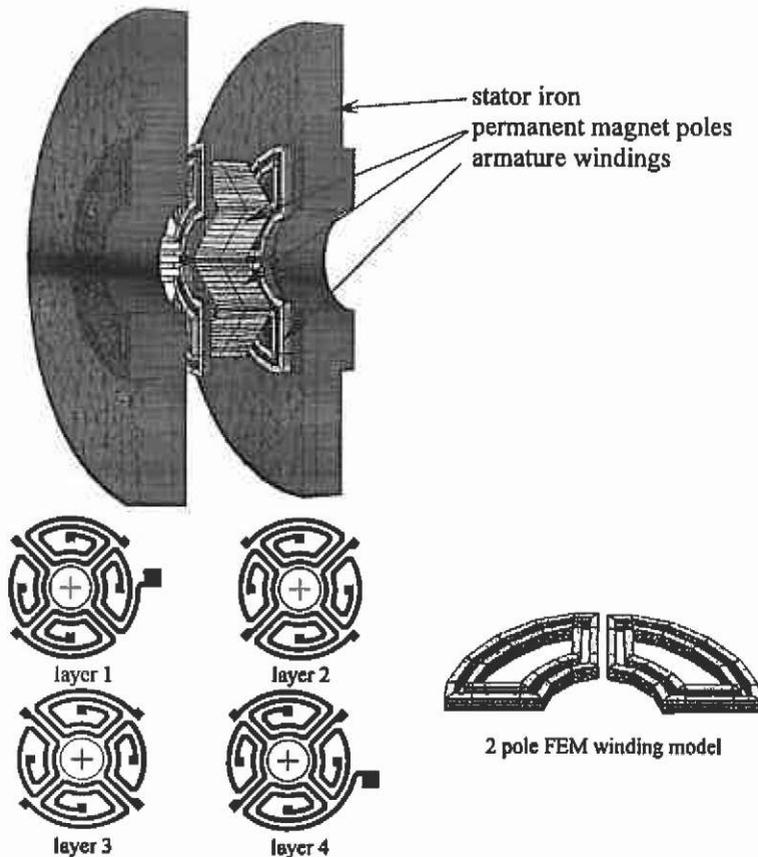
Fig. 10.19. Electromagnetic mini motor with 3D finite element model and armature winding layout.

In this type of application, the supply source is an essential part of the system. Due to the non-linearities of the ferromagnetic parts of the machine, the link with the time pattern of the supply voltage cannot be simulated using superposition. The motor is operated as a brushless DC motor. Constant DC currents are switched to the armature winding in the stator according to the signals of a position-sensing system equipped with hall sensors. The rotor is constructed with NdFeB permanent magnet blocks of the dimension 2x2x2 mm. The use of high-energy permanent magnet material can lead to significantly improved efficiency and performance of small electrical machines. The high remanence and coercivity at room temperature makes this material particularly attractive for this type of machine. However, the sensitivity of the coercivity of NdFeB to high temperatures calls for increased attention to the thermal aspects of a design. Integrated designs using NdFeB magnets are cost-effective for fractional and sub-fractional horse power motors.

In the design stage, the target is to obtain reliable results predicting the operational behaviour of this device. Macroscopic parameters, reactances and reluctances, describe the technical physical properties of the machine. Due to the presence of ferromagnetic materials, the calculations have to account for the non-linearities.

To extract the parameters for a simplified equivalent circuit, the inductance is found from the stored field energy after replacing the permanent magnets by air:

$$W_{magnetic} = \frac{LI^2}{2} \tag{10.35}$$

The torque is found from the virtual work:

$$T = \frac{\partial W_{magnetic}}{\partial \vartheta} \tag{10.36}$$

The evaluation of the torque as a function of the rotor position is performed by a sequential analysis. Once the device-dependent parameters are determined, the equivalent circuit is modelled and in combination with the characteristic values of the supplying energy source (Fig. 10.20), the overall system is modelled, simulated and analysed.

The computed flux density distribution (Fig. 10.21) shows that the ferromagnetic material of this design is not saturated. A motor construction with less material could be designed in a following design approach.
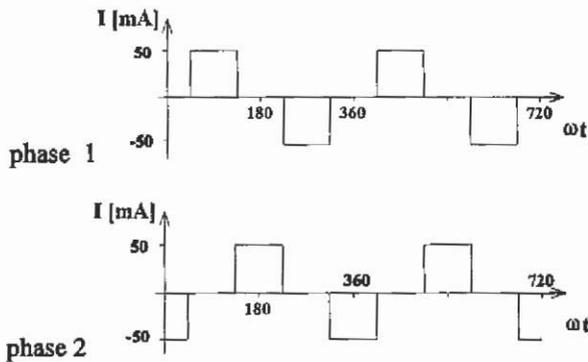
Fig. 10.20. Switching scheme for the stator winding currents of phase 1 and 2.
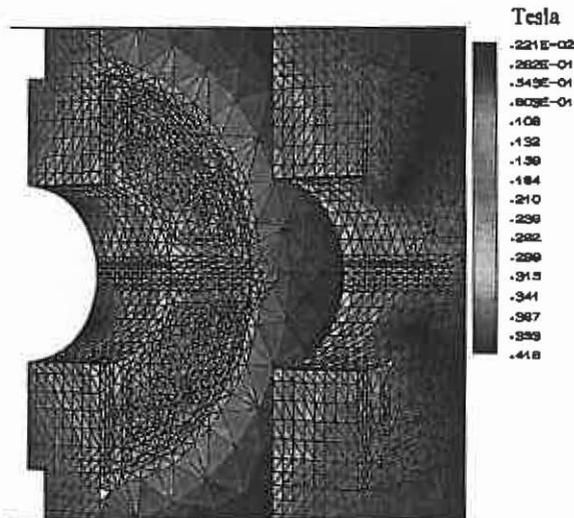
Fig. 10.21. Magnetic flux density at no-load operation.

10.1.4.2   Mini disc-type motor The motor in this example is an axial field disc-type motor excited by permanent magnets. The outer diameter is about 45 mm, axial length about 15 mm (Fig. 10.22). The stator back-iron consists of two discs of ferromagnetic material. The armature winding is placed on the two stator sides.

The required torque recommends a multi-layer winding in order to realise a sufficiently high current layer. Therefore, each winding consists of four layers. On each stator side, eight windings are installed and connected to form two phases. The rotor consists of a thin disc constructed with a sintered NdFeB material magnetised in the axial direction in a multipole arrangement (Fig. 10.23).
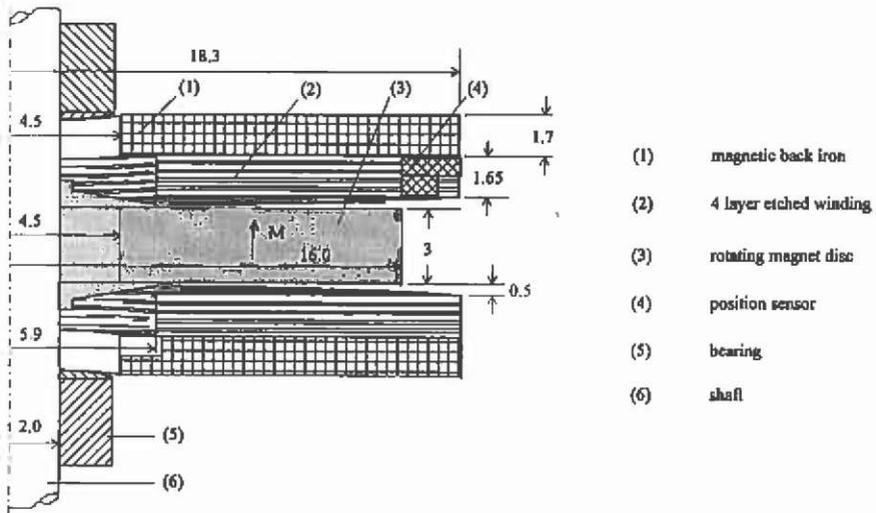
Fig. 10.22. Main dimensions of the mini disc-type motor.

(1)	magnetic back iron
(2)	4 layer etched winding
(3)	rotating magnet disc
(4)	position sensor
(5)	bearing
(6)	shaft

1. armature winding layers
2. permanent magnet poles
3. stator back iron
4. shaft
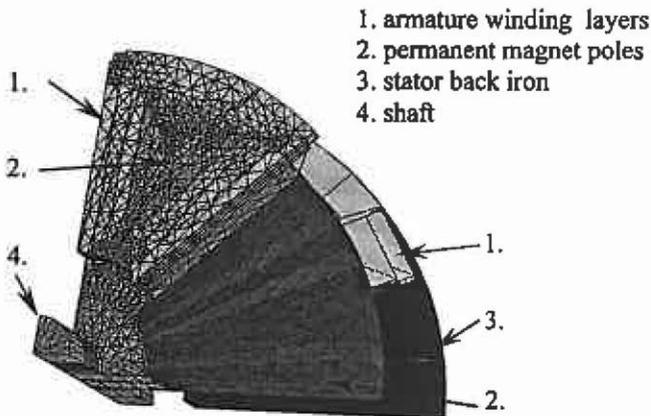
Fig. 10.23. Three-dimensional finite element model of the disc-type motor.

The supplying current source operates in the same way as introduced for the permanent motor (Fig. 10.19). To compute the induced voltage, Ψ_w , the flux generated by the magnets and coupled with one winding, which is not carrying current, is computed by a sequential approach as a function of the rotor position (Fig. 10.24). The torque can be computed for each instant in time by integrating it along the current-carrying conductors of the winding (Fig. 10.25) and superimposing the torque generated by the single winding phases.
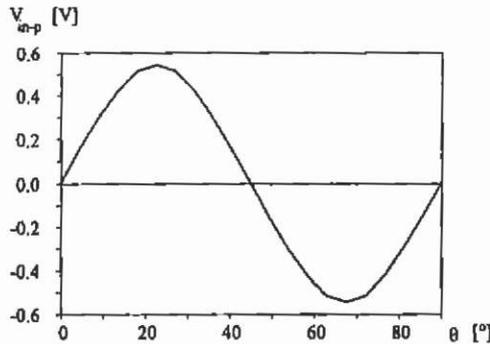
Fig. 10.24. Computed voltage induced in one phase at a rotor speed of 1000 rpm.
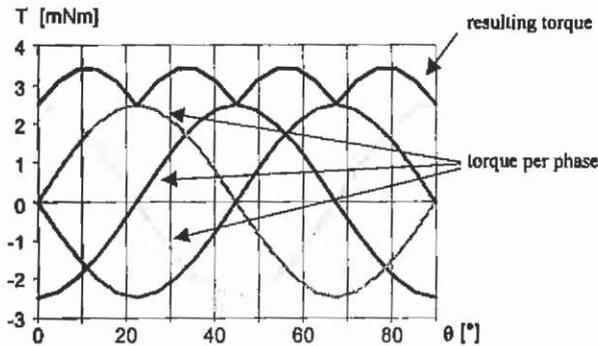
Fig. 10.25. Torque characteristics for the disc-type motor (winding coils operated with 0.5 A).

10.1.5 Design of electrostatic micro motors

Scaling analysis shows that as size is reduced, electrostatic designs become advantageous over the electromagnetic versions that dominate at dimensions starting in the millimetre range. The electrostatic micro motors studied here are based on the principal of variable capacitance. The operation principle is very simple. A voltage on the stator electrodes induces a charge on a conducting rotor and in response the rotor moves to minimise the electrostatic field energy.

The most inexpensive fabrication technology of electrostatic micro machines is a thin film process for planar structures. Therefore, such rotating actuators are extremely flat and the generated forces are very low. The motor with its outer dimensions is shown in Fig. 10.26. Fig. 10.27 shows its corresponding three-dimensional finite element model. In this case of geometrical symmetry, the mesh is extruded in an angular direction to build up the three-dimensional structure (Fig. 10.27b).
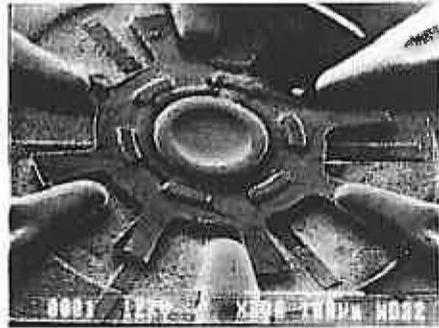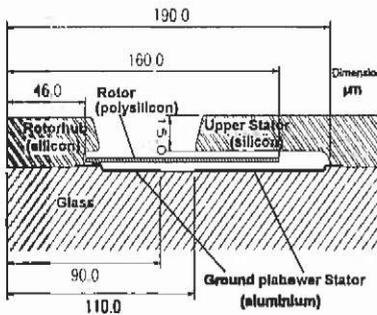
Fig. 10.26. Detailed construction and outer dimensions of the studied axial field electrostatic micro motor.

b)   c)
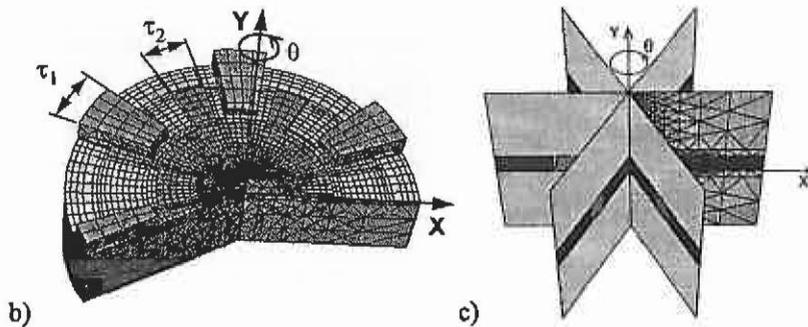
Fig. 10.27. a) Axial field electrostatic micro motor model and b) the base planes rotated in angular extrusion direction.

Radial field type machines are also feasible. When the same height of the machine is considered, the surface that contributes to the interaction between stator and rotor is much smaller. However, the problem is that only very small forces can be generated. Using a radial type of interaction and the LIGA production technique, allowing the fabrication of higher microstructures, results in higher torque values. This technique is very expensive. More inexpensive alternatives are developed but are not capable of supplying the same depth of the rotor. Fig. 10.28 shows the three-dimensional finite element model of a radial field micro motor.

However, both types of motor can be analysed in an analogue way. The electrostatic energy stored in the model is evaluated and serves as data to obtain the parameters of an equivalent circuit.

$$W_{electrostatic} = \frac{CV^2}{2} \tag{10.37}$$
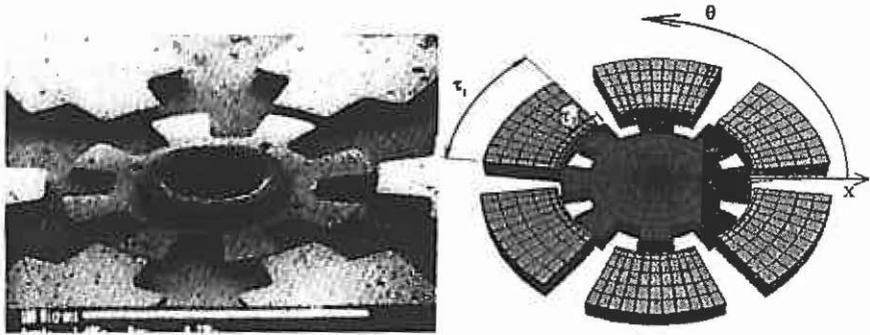
Fig. 10.28. Radial field electrostatic micro motor with FEM model.

The use of this equivalent circuit model enables calculation of the forces of the motor operated with various voltage cycles without new computationally expensive FEM analyses.

The desired parameters in the equivalent circuit are the values C of the capacity between the single components of the geometry as indicated in Fig. 10.29. The equivalent circuit in Fig. 10.29 consists of 12 capacitances, twice the number of stator electrodes. The capacitance of each capacitor varies with the rotor position.
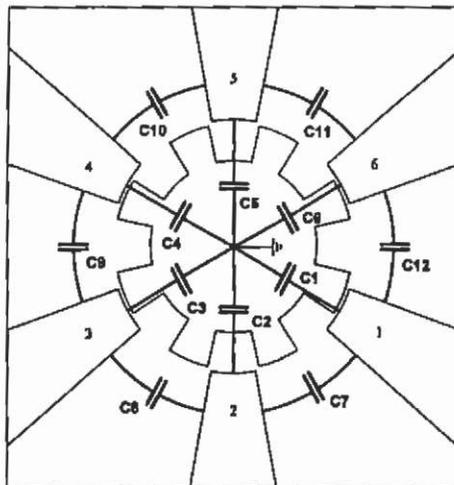
Fig. 10.29. Definition of the elements of the equivalent circuit for a 6/8 pole radial field electrostatic micro motor.

To avoid axial forces on the rotor shaft, the motor must be excited symmetrically. Fig. 10.30 shows the possible symmetric excitations of a motor with 6 stator electrodes. The grey electrodes are excited by 1 V and the rotor electrodes are set to ground potential 0 V.
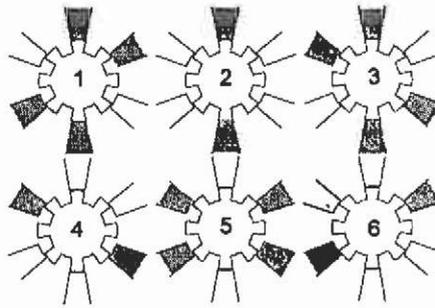
Fig. 10.30. Possible symmetrical excitation sequence to perform one revolution of the rotor.

By applying different excitation cycles to the equivalent circuit, the torque characteristics versus rotor position can be calculated. Using the principles of virtual work, the torque is found by partial differentiating of energy with respect to the angle of rotation:

$$T = \frac{\partial W_{electrostatic}}{\partial \vartheta} \quad . \tag{10.38}$$
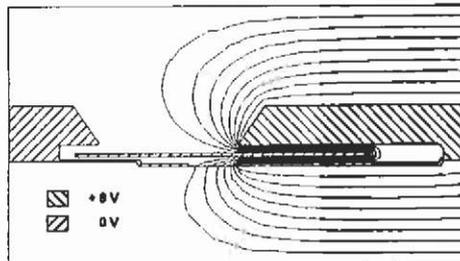
Fig. 10.31. Potential solution of the radial cross-section of the axial field motor.
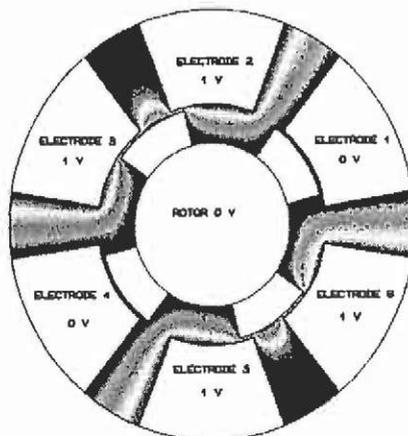
Fig. 10.32. Potential solution of the radial field type micro motor.

10.2   Coupled thermo–electromagnetic problems

Coupled thermo-electromagnetic problems have to be considered in the simulation of realistic electromechanical devices and electroheat installations. To handle this type of problem, an interface program to co-ordinate the finite element method simulations and to perform intermediate calculations, such as heat source evaluation, numerical relaxation and mesh transitions, is recommended.

The material data used to define FEM problems are often strongly dependent on the temperature. Since material data parameters occur in many coefficients of the electromagnetic field equations, the calculation of the electrical and/or the magnetic field coupled to the thermal field is recommended. The heat generation of electromagnetic nature results in a coupling of the source term of the right-hand side of the thermal equation. The combined problem in the electromagnetic and the thermal domain is generally described by Helmholtz-like differential equations. The discretisation yields two or more sets of algebraic equations that have to be coupled numerically: the electric field and/or magnetic field together with the thermal FEM-equations. Each of them can be extended with an algebraic set of circuit equations. These include coupling terms as well, e.g. in resistances.

The discretisations, on which the computations for the single field problem are performed, do not have to be identical. Sometimes, only a sub-mesh has a physical meaning: e.g. air carrying a magnetic leakage flux is replaced by a convection constraint in the thermal model; the solid parts can be identical. Even the mesh in areas with more than one continuous degree of freedom can be discretised with different overlapping geometrical meshes and/or element types. Therefore, mesh transition operations have to be defined. The groups of algebraic equations can be solved with a strong coupled or with a cascade coupled strategy.

The first approach consists of the generation of a large system of non-linear equations with both types of FEM-equations, associated with the coupling terms. The mesh transitions and heat source terms have to be written as linearised algebraic functions. This large linear system may have unfavourable numerical properties due to the different nature of the underlying physical equations, resulting in a difficult to solve problem.

The second method (Fig. 10.33) defines an iterative process in which both sets of equations are solved sequentially. The mesh transitions and heat source calculations are necessary intermediate steps and do not necessarily have to be linearised. This approach can be interpreted as a "decomposition" of the different fields.
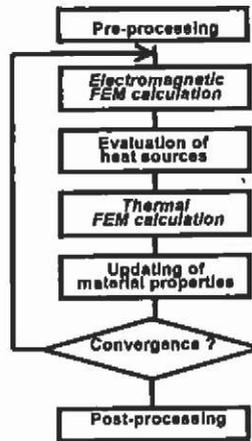
Fig. 10.33. General flow chart of the cascade algorithm.

The following groups distinguish the operations necessary for the iterations:

- Updating of material properties: The various temperature dependent material parameters are updated whenever the algorithm requires it. Therefore, various characteristics must be implemented. These involve electrical conductivities, thermal conductivities, permanent magnet material properties, thermal conductivities, loss coefficients and characteristics.
- External process handling: Calls to execute external FEM-solvers and mesh generators.
- Iteration control: Process commands to control the flow of the iterative loop and evaluation of stopping criteria. These criteria can be based on absolute or relative residuals or solution differences between two consecutive weighted solutions of the total problem or a sub-problem respectively.
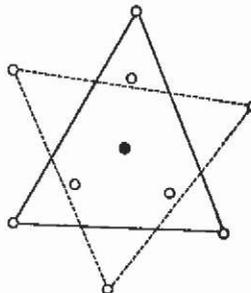
Fig. 10.34. Projection of element related values.

- Data transition commands: When different meshes are used, mechanisms are necessary to project the field variables onto another mesh. Basically, this is a per-node interpolation of the solution.

The projection of the values associated with the element's surface or volume, e.g. a calculated loss density, is not straightforward. If the meshes do not differ very much, the position of the centre of gravity of the element to be filled in, can be located in the other mesh (black dot Fig. 10.34). The corresponding element-related value can then be copied. If the meshes differ much or if a higher order of accuracy is desired, an average can be generated by means of a numerical integration using Gauss points (additional white dots Fig. 10.34).

It can be stated that a large difference between the meshes is not advantageous. It would mean that the related physical domains would not be calculated with a corresponding accuracy. However, mesh differences can arise due to local mesh quality reasons.

- Adaptive relaxation of the convergence process: In order to prevent the non-linear iteration process from divergence and to accelerate the convergence, an appropriate relaxation method must be applied. The damping factor can be predefined according to a certain function of the iteration number or adaptively, based on a minimisation of the total or partial (weighted) residual vector.
- Heat sources calculation: For the area covered by every meaningful element, a heat source density can be calculated based on the electromagnetic solution (Table 10.2).

Table 10.2. Overview of main heat sources in electromagnetic problems.

heat source	occurs in	formula	application
joule losses	problems with electrical current (perpendicular to, or in a plane)	$q_{joule} = \dfrac{J_{rms}^2}{\sigma}$	conductors in machines, eddy currents in induction heating
iron losses	non-static magnetic field problems	$q_{iron} = \left(c_1 f + c_2 f^2\right) B^{(2)}$	magnetic materials with a hysteresis loop
dielectric losses	non-static electric field problems	$q_{diel} = \omega\left(\varepsilon_r \tan\delta\right) E^2$	capacitive heating, losses in isolating materials
external heat sources & sinks	various	look-up table, ...	e.g. ventilation, cooling channels, friction

10.2.1 Three-phase high voltage power cable

In this section, the results of a three-phase power cable simulation with respect to the coupled magnetic/electrostatic/thermal field problem are shown.

Three-phase power cables exist in many variations and types, differing in conductor shape, material choice, conductor arrangement etc. They consist mainly of the following parts, in which several of the previously mentioned loss mechanisms can be found:

- conductor, usually made of copper, suffering from joule losses caused by the high current
- insulation layers and filling materials, loaded with electric fields and therefore subject to dielectric losses
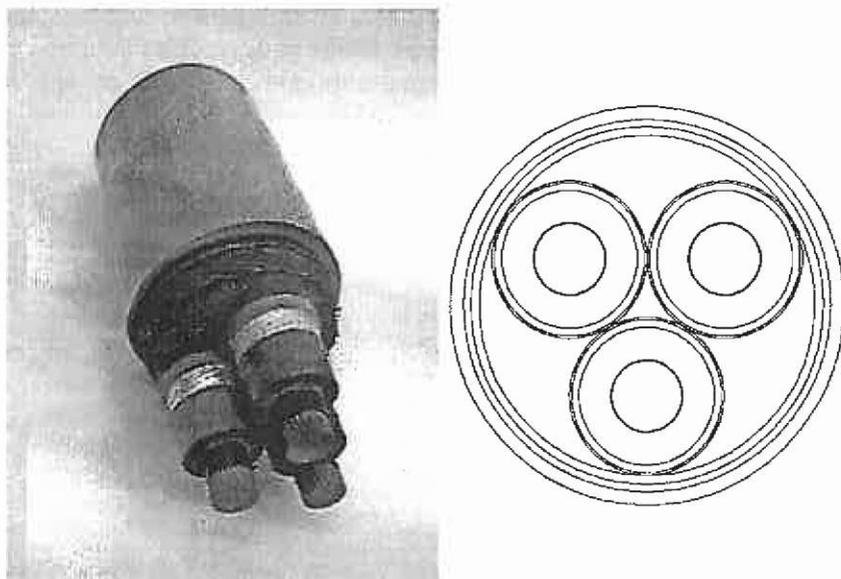
Fig. 10.35. Geometry of the three-phase high voltage power cable.

- grounded lead sheath around the primary isolation, shielding the electric field; due to its relative low conductance, internal eddy currents can develop
- mechanical protections (armour), sometimes made of magnetic steel and therefore subject to hysteresis and eddy current losses

The presence of both electrically related and magnetically related heat sources leads to a combined model consisting of three field types. Electric, magnetic and thermal fields that have to be calculated over a complete or a partial cross-section of the cable and its surrounding.

The electrical field, described by the scalar potential V, is only of interest in the isolation part loaded with an electrical field. Therefore, only a mesh covering this region is required to solve the electrostatic field equations.

The time-harmonic magnetic field is calculated on a larger mesh, since it is only partly shielded by the mechanical protection and thus a leakage field can exist outside the model. This leakage flux is considered by the region surrounding the cable geometry; the far field is modelled by a Kelvin transformed mesh. The losses consist of joule losses in the conducting regions, such as the lead, steel and copper and possible iron losses inside the steel.

The thermal field is represented by the temperature potential distribution T. The static thermal field region consists of the cable with the surrounding soil in which it is buried. From a certain distance, the ground is modelled by a Kelvin transformation and therefore assumed to be infinitely deep. It is assumed that the ground surface is cooled by convection.
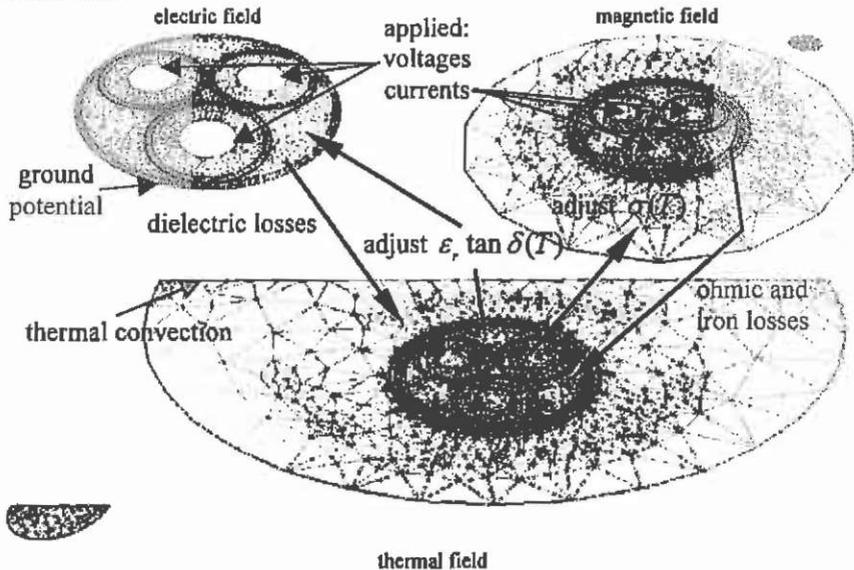
Fig. 10.36. Problem meshes used to compute the different physical fields.

The element-wise over the previous meshes (electro-, magneto-static) calculated losses are projected onto the thermal mesh. The extracted temperatures are used to update the material properties in the other fields. Basically the dielectric loss factor and the thermal conductivity depend on temperature, but this is an effect of minor importance in this example problem. The largest parameter changes are encountered in the conductivity of the copper.

To solve the entire problem, a three-domain mesh must be constructed (Fig. 10.36). The mesh for the electric field contains 9943 first order triangular elements, the magnetic field mesh 15378 and the thermal field mesh 15560 elements.

The results of this threefold-coupled problem are collected in the following figures.

a) Solution of the electric field.   b) Flux lines of the magnetic field.

c) Temperature distribution.
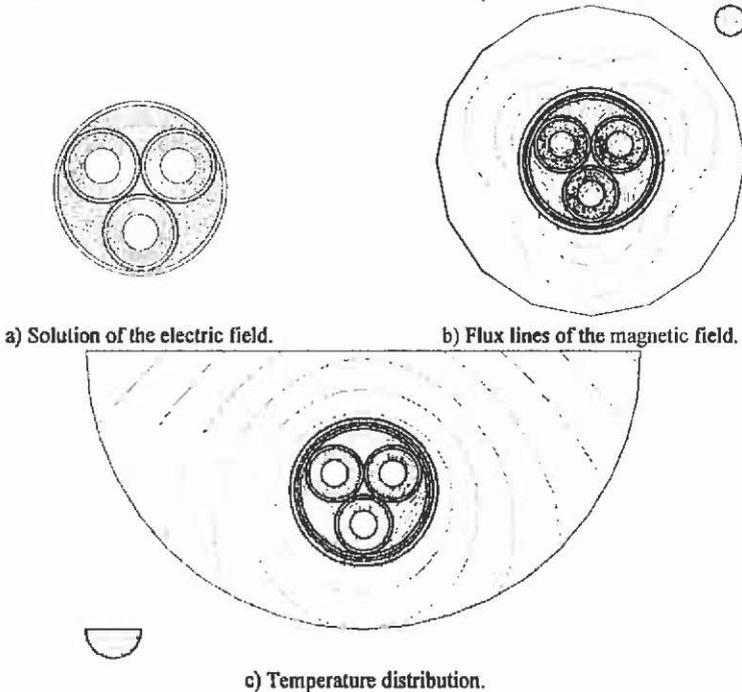
Fig. 10.37. Entire solution of the threefold coupled problem.

The temperature in the centre of a conductor amounts 84°C, which corresponds to reported measurements (Van Dommelen & Germay [116]). Table 10.3 shows the computed values of the heat sources and their location within the models.

Table 10.3. Overview of main heat sources in electromagnetic problems.

location	loss mechanism	value [W/m³]
copper conductors	ohmic	$5,97.10^4$
conductor isolation	dielectric	$1,17.10^2$
inter-conductor filling material	dielectric	$2,55.10^2$
mechanical protection	ohmic + iron	$1,23.10^{-3}$

10.2.2 Coupled simulation for electrical machines

An efficient simulation algorithm for the coupled magnetic-thermal field of electric machines is proposed. This algorithm uses a combined FEM-

circuit approach in both magnetic and thermal field regions. FEM calculations are performed to compute magnetic and thermal phenomena in a xy 2D Cartesian cross-section, whereas a circuit approach is applied to consider thermal phenomena in the z direction (Fig. 10.38).
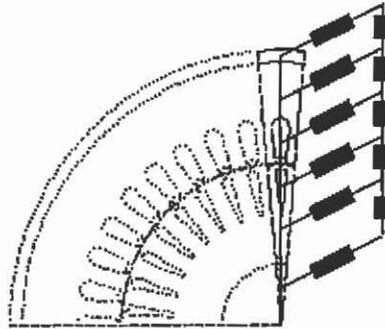
Fig. 10.38. Principle of the combined thermal FEM-circuit-model; the resistances in the model represent a part of the symmetry of the axial thermal phenomena. (The cross-links to the next slot are not shown.)

The computations in the magnetic and the thermal domain result in four sets of equations: the magnetic FEM-equations, electrical circuit equations, the thermal FEM-equations and thermal circuit equations. The entire approach to solve the entire problem is a computationally efficient method. The simulation of a 15 kW TEFC (totally enclosed and fan-cooled) four-pole induction machine demonstrates the coupled approach.

In general, the heat transport perpendicular to the axis of cylindrical electrical machines can be modelled with a high accuracy. Less is known about the effects influencing the heat flow in the axial direction. For example, the flow of the cooling fluids in the end-regions is very complex.   Therefore, 3D-calculations should be coupled to a 'computational fluid dynamics' simulation, asking for huge computational efforts.

The circuit approach offers the advantages that the heat transfer in the axial direction is modelled in a straightforward way, using thermal resistances. However, the determination of their values is troublesome and asks for experience and measurements.

Different methods to solve the overall problem are possible:

- A cascade iteration algorithm in which the systems are solved in successive steps. This results in a poor convergence. On the other hand loss calculation algorithms can be introduced in a simple way.
- The solution of a numerically strong coupled large system of equations in which the four sub-sets together with the coupling

terms are assembled. This may result in a very ill-conditioned system of equations. Therefore, particular equation solvers are recommended to obtain a fast convergence.

- An intermediate approach is possible by placing the FEM-equations together with their corresponding circuit equations in one system of equations describing a single physical domain. These 'physical' sub-systems are then coupled in a cascade-like iteration. This method has a moderate rate of convergence, but requires less memory.

The time-harmonic magnetic model is coupled with circuit equations describing the effect of the end-windings, the bars outside the rotor and the end-rings as already introduced. The thermal model consists of a FEM component, including equivalent convection coefficients to account for ribs and equivalent thermal conductivities for areas such as the air gap and the slots. This thermal FEM model is extended by a circuit approach. The circuit equations represent the thermal paths connecting shaft, yokes, slots and the frame through thermal resistances and represent the internal end-region in air, the end-windings, end-rings, bearings and end-caps of the machine. The resulting temperature distribution of this coupled approach is plotted in Fig. 10.39. The isothermal lines in the shaft are caused by heat flowing through the shaft, a path that is described by a thermal network circuit equation.
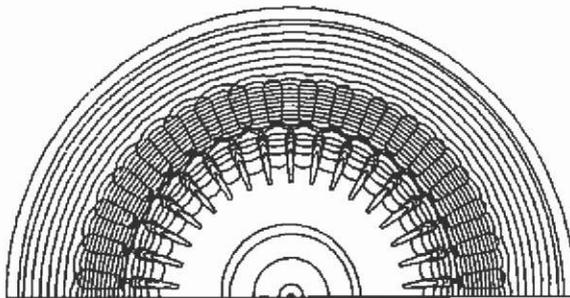
Fig. 10.39. Isothermal lines of the temperature distribution.

10.2.2.1 Modelling of thermal contact resistances In the electromagnetic-thermal coupled modelling of electromagnetic devices, the thermal simulation poses some extra difficulties, not present in the electromagnetic field.

Examples are the mixed boundary conditions as found in convection (linear coefficient) or radiation (non-linear coefficient). These require extra adjustments of the finite-element matrix and right-hand-side.

A second particularity is the presence of thermal contact resistances, although it can be noted that electrical contact resistances exist as well when electrical current in the plane of the simulation is modelled. They are a particular problem encountered in thermal FEM models of the electromagnetic devices. Examples for this thermal contact interface are:

- contact frame-stator yoke
- contact stator winding-slot
- contact rotor winding/bar-slot
- contact rotor yoke-shaft
- glue layer between a permanent magnet and the yoke.

Several approaches are known to model such contact resistances. Corrections to the thermal conductivities of the conductor materials can be applied to consider this effect. This approach causes an error in the internal temperature distribution of the conducting region, but the average temperature is calculated accurately. The temperature distribution is used to update material data of the related sub-problems. Despite the extra computational costs of averaging, this approach has the advantage that there is an obvious geometrical relation between the elements of regions in the different sub-problems. This is advantageous for projection methods.

An extra equivalent contact layer of elements can be inserted, filled with an equivalent contact material. In order to obtain an acceptable aspect ratio of the finite elements, a sufficient number of elements have to be generated inside the contact layer and the adjacent regions. Due to the high number of slots in an electrical machine and the small size of the contact layer, this yields a significant growth of the numerical model and hence computation time. The number of elements can be reduced if the layer is enlarged, but this reduces the size of the conductor and the tooth, and so no clear geometrical relationship exists between the thermal and the magnetic sub-problem. If the same model, with the reduced conductor, were to be used for the magnetic sub-problem, an error would be made in the leakage flux and the joule heat calculation.

Another possibility is to duplicate the nodes lying on the edges marking the border between the conducting region and the iron. Extra terms in the equations modelling the contact are inserted to define the thermal relationship of the nodes. The standard meshing algorithms have to be adapted to generate the extra nodes. Moreover, these extra equations change the numerical properties of the matrix system to be solved.

The third proposed approach will be outlined. Here, the mesh generator must be able to double the nodes lying on the contact resistance

(Fig. 10.40). This must be performed in a consistent way to prevent any 'crossing' of the edges of the contact resistance. Standard mesh generators need to be extended with appropriate search and data management structures.
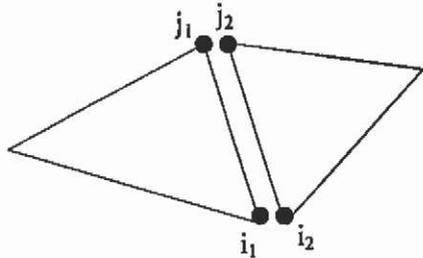
Fig. 10.40. Two elements at the thermal contact resistance interface.

Within the FEM code, the standard element matrices are generated (the right-hand-side remains unaltered). The following extra matrix, suited for first order elements, is added to the system. The exact values of the coefficients are obtained using the Galerkin approach.

$$
\begin{array}{c}
 \\
i_1 \\
j_1 \\
\vdots \\
i_2 \\
j_2
\end{array}
\begin{array}{cccc}
\;\; i_1 \quad\quad j_1 \quad\;\; \cdots \quad\;\; i_2 \quad\quad j_2 \\
\left[
\begin{array}{ccccc}
\dfrac{h_c l_e}{3} & \dfrac{h_c l_e}{6} & \cdots & -\dfrac{h_c l_e}{3} & -\dfrac{h_c l_e}{6} \\[2mm]
\dfrac{h_c l_e}{6} & \dfrac{h_c l_e}{3} & & -\dfrac{h_c l_e}{6} & -\dfrac{h_c l_e}{3} \\[2mm]
\vdots & \vdots & \ddots & & \vdots \\[2mm]
-\dfrac{h_c l_e}{3} & -\dfrac{h_c l_e}{6} & & \dfrac{h_c l_e}{3} & \dfrac{h_c l_e}{6} \\[2mm]
-\dfrac{h_c l_e}{6} & -\dfrac{h_c l_e}{3} & \cdots & \dfrac{h_c l_e}{6} & \dfrac{h_c l_e}{3}
\end{array}
\right]
\end{array}
\qquad (10.39)
$$

with h_c the contact resistance coefficient and l_e the length of the adjacent element edges.

A computed example using this approach is shown in Fig. 10.41. A partial thermal model of an induction machine is shown. The example consists of one stator winding slot and a rotor bar. The above relation models the contacts between the slots and the iron core. This leads to a set of isothermal lines appearing to be discontinuous. This is not true since many isothermal lines lie in the temperature jump inside the thermal contact resistance.
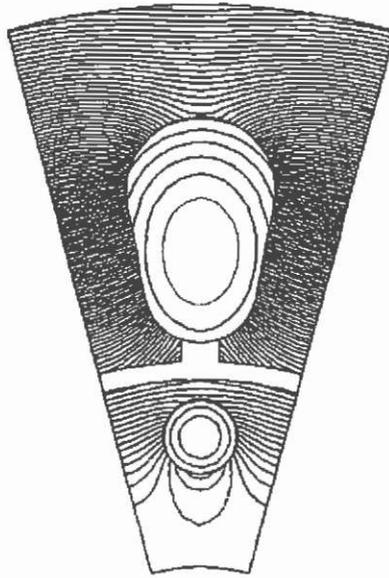
Fig. 10.41. Computed temperature distribution using the thermal contact
resistances in a slot arrangement of an electrical machine example.

10.3   Numerical optimisation

10.3.1   Shape optimisation for small DC motor

The application of the methods used is demonstrated by the optimisation
of a small DC motor. For the optimisation a (4/4, 12) evolution strategy
combined with the simulated annealing algorithm is used.

The objective is to minimise the overall material expenditure,
determined by permanent magnet-, copper- and iron volume subject to a
given torque of the example motor.

$$Z(\mathrm{x}) = 10^{\left(\frac{\mathrm{vol}(x) - \mathrm{vol}_{max}}{\mathrm{vol}_{max}}\right)} + penalty \tag{10.40}$$

The use of penalty term in the form:

$$penalty = \begin{cases} T < T_{min} : 10^{\left(\frac{T_{min} - T(x)}{T(x)}\right)} \\ T \geq T_{min} : \quad 1 \end{cases} \tag{10.41}$$

allows the evaluation of the objective function even if the torque
constraint is violated.

The torque is computed by integrating the Maxwell stress tensor in
the air gap region. Flux density dependent rotor iron losses were taken

into account at a rated speed of 200 rpm and subtracted from the air gap torque to form the resulting output torque.

The overall dimensions and the slot geometry of the DC motor are described by 15 free design parameters. The free parameters are the $n/2$ edges of the polygon describing the rotor slot contour and the outer dimensions of rotor and stator as indicated in Fig. 10.42. The motor consists of a stator back iron with a 2-pole Ferrite permanent magnet system and a rotor with six slots.

1. stator back iron
2. permanent magnet
3. rotor iron
4. winding slot
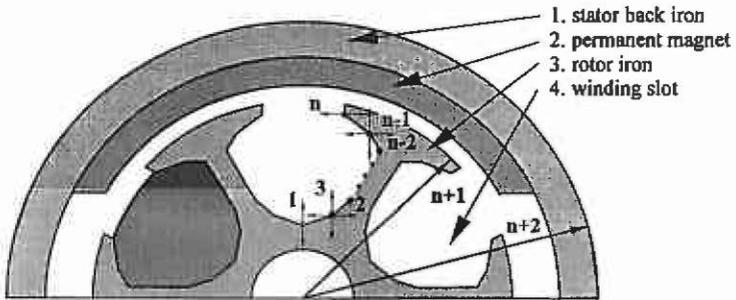
Fig. 10.42. Geometrical definitions and design variables of the DC motor.

The necessary two-dimensional field computation to evaluate the quality function, to compute the torque of the machine, is performed by standard two-dimensional finite element analysis. To ensure controlled accuracy, adaptive mesh generation is applied until a given error bound is fulfilled.

An initial mesh is generated from any geometry represented by non-overlapping polygons. Fig. 10.43 shows an initial and adaptive generated mesh for the example DC motor.

a)   b)
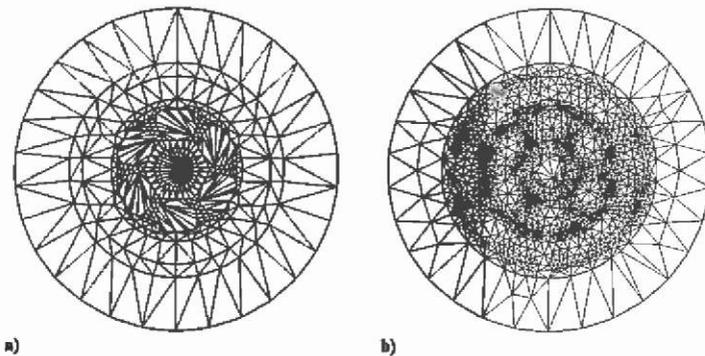
Fig. 10.43. a) initial- and b) adaptively refined mesh.

Constraints result from fabrication conditions. The change of shape from a sub-optimal initial geometry to the final shape of the motor can be

taken from Fig. 10.44. It can be noticed that the iron parts of the initial geometry are over dimensioned. The actual torque of this configuration was approximately 25% lower than the desired value T_{min}. The optimised motor holds the torque recommended, which is achieved mainly by enlarging the winding copper volume by about 20%. The most significant change from start to final geometry can be seen in the halving of the iron volume. Consequently the iron parts are highly saturated, especially the teeth regions. In comparison to this, a test optimisation with neglected rotor iron loss results in a 10% smaller rotor diameter. Unfortunately, the permanent magnet material is brittle, which limits the minimum magnet height. The magnet volume decreases slightly. Along optimisation the overall volume off the motor was reduced by 38%. The rate of convergence is plotted in Fig. 10.45.
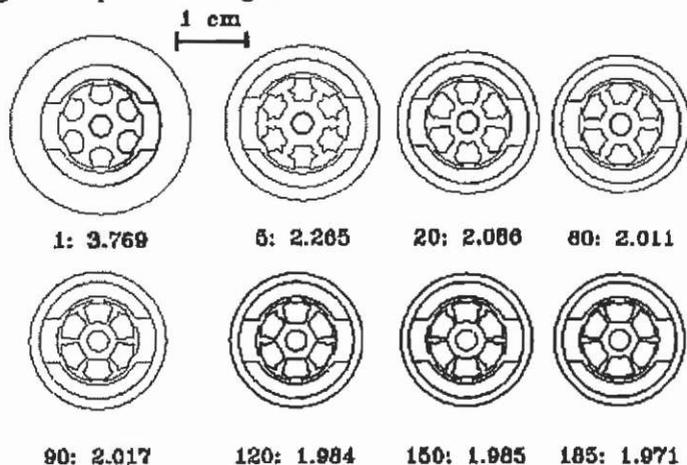
Fig. 10.44. Motor shapes during optimisation (iteration step: quality).
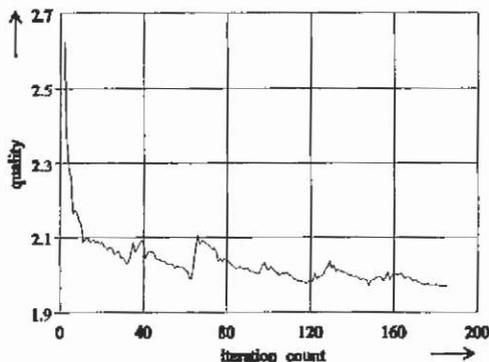
Fig. 10.45. Quality versus iteration counts.

10.3.2 Pole shape optimisation of a synchronous generator

The aim is to optimise the pole shape of a 3-phase synchronous generator. The quality function is evaluated by a simplified analytical approach. At no load of the generator typically the generated voltage is desired to be sinusoidal. The time dependent sinusoidal output voltage requires a position dependent sinusoidal distributed flux distribution in the air gap. With DC field exciting current and a concentrated shaded pole, the sinusoidal field excitation is reached by influencing the air gap reluctance. The air gap length $\delta(\varphi_z)$ is a function of the circumferential angle φ_z.

Due to symmetry only half a pole pitch is used for the evaluation of the air gap flux density distribution. The following assumptions were made:

- the stator of the machine is spotless
- saturation of ferromagnetic parts is neglected
- inside the iron parts it should be $\mu_{Fe} \to \infty$, i.e. flux lines are perpendicular at the iron boundaries
- the flux lines are approximated by circular arcs
- pole flux leakage is neglected
- flux density in the interpolar gap is not present.
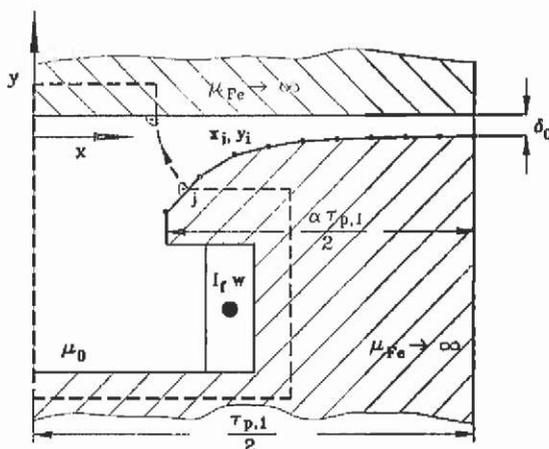
Fig. 10.46. Geometry and co-ordinate system.

Fig. 10.46 shows the geometry and co-ordinate system of the used configuration. The pole shape of half a pole pitch is approximated with n parts of a polygon. Between the sample of the polygon, linear interpolation of the flux density is applied. The objective variables are the

y-co-ordinates $y_1, ..., y_n$ of the samples of the polygon. The air gap length $\delta_x > 0$ restricts the optimisation problem.

$$y_i \leq 0 \quad \text{with} \quad i = 1(1)n \quad .$$

With the objective function:

$$\frac{\sqrt{\sum_{v=2}^{\mu} B_v^2}}{B_p} \rightarrow \min. \tag{10.42}$$

B_p denotes the fundamental of the flux density distribution and the B_v are the harmonics. Determination of the harmonics is performed by a fast Fourier transformation (FFT). With Ampere's law on the path of integration as indicated in Fig. 10.46 the flux density of the position of interest is evaluated with:

$$B_j = \mu_0 \frac{I_f w}{\delta_j} \quad . \tag{10.43}$$

Fig. 10.47 shows the initial and optimised pole shape. The variations of the pole contour for temporary iteration steps can be taken out from Fig. 10.48.
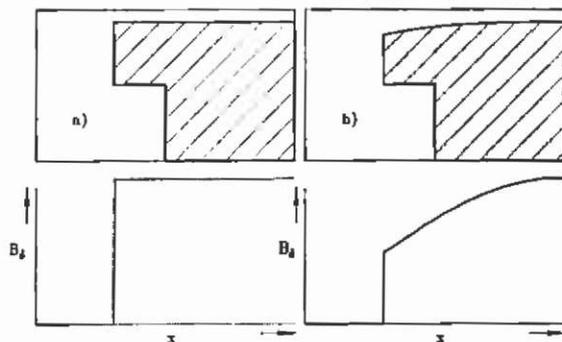
Fig. 10.47. Pole shape and flux density; a) rectangle pole and b) optimised pole shape.
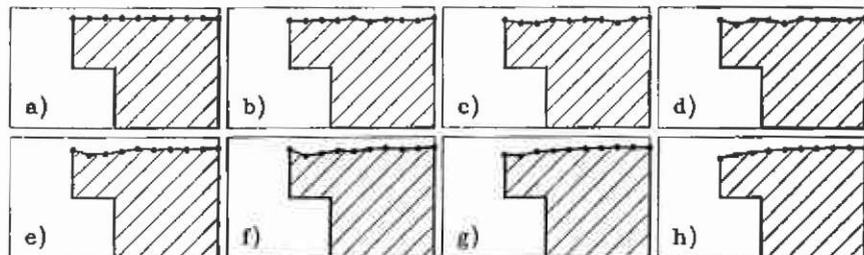
Fig. 10.48. Pole shape and quality $Z^{(k)}$ of iteration step k.

Table 10.4. Quality function during optimisation.

a)	$Z^{(0)}$	=30.39%	e)	$Z^{(1)}$	=29.62%
b)	$Z^{(2)}$	=28.04%	f)	$Z^{(3)}$	=27.26%
c)	$Z^{(4)}$	=25.08%	g)	$Z^{(5)}$	=24.59%
d)	$Z^{(40)}$	=23.53%	h)	$Z^{(244)}$	=23.26%

10.3.3 Optimisation of an actuator using a magnetic equivalent circuit model

The MEC method and the mentioned combined numerical optimisation algorithm will be applied for the example. The task is to optimise the shape of a brushless DC motor. Fig. 10.50 shows the initial shape of the electromagnetic device. The used complete equivalent magnetic circuit can be taken out of the introductory section about numerical techniques.

The armature winding is fixed to closed stator slots and the rotor is axially assembled out of permanent magnet rings. The magnet material used is a plastic bonded NdFeB grade (MQ1). The objective is to minimise the material costs of the construction under the assumption of the same torque production as the initial construction. Objective variables can be taken from Fig. 10.49.

Material costs are estimated and set to 0.7 US$/kg for the lamination, 4.1 US$/kg for the used copper volume of the armature winding and 110 US$M/kg for the used magnet material. The resulting shape of the motor can be seen in Fig. 10.51.

Using a quality function with K as the sum of all material costs, the start value of quality is approximately 42 US$. After the optimisation, the overall material costs decreases to 11 US$. Fig. 10.51 shows the reduction of the cost intensive permanent magnet volume. Due to the expensive armature winding compared to the lamination cost, wide stator teeth can be noticed. Maximum flux density of 1,04 T is enumerated in the teeth.
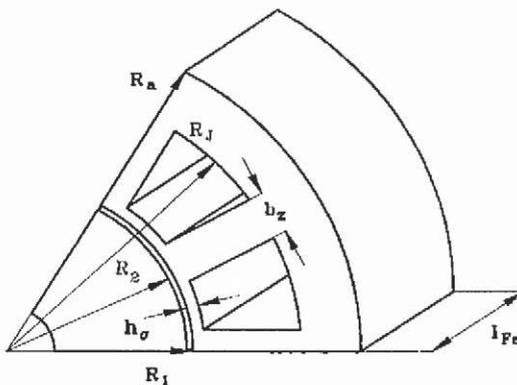
Fig. 10.49. Objective variables.
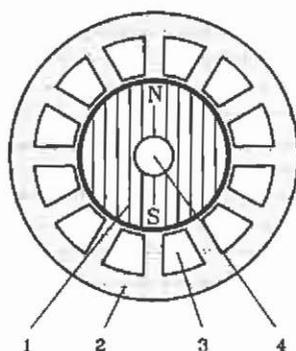
Fig. 10.50. Initial geometry.

1 permanent magnet; 2 stator yoke; 3 area of armature winding; 4 shaft.
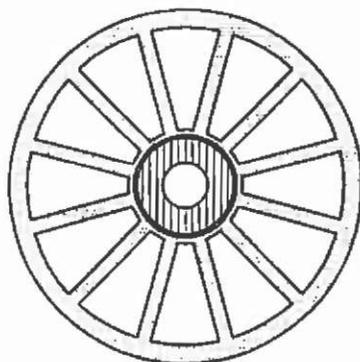
$l_{Fe} = 37,3$ mm

Fig. 10.51. Optimised shape of the DC actuator.

10.3.4 Design of a lifting magnet

FEM method and the numerical optimisation algorithms are applied to the shape optimisation of a lifting magnet. The objective is the reduction of the weight of the device at constant lifting force. There are no additional geometrical constraints. The problem is formulated with 10 free parameters to be optimised as indicated in Fig. 10.52.

Obviously, the problem has to be defined as a non-linear magneto-static field problem. During optimisation an accuracy of at least 1% was required for the FEM field calculation. The size of the mesh was restricted to 4000 elements. The optimisation method used was the evolution strategy (4/3, 12).
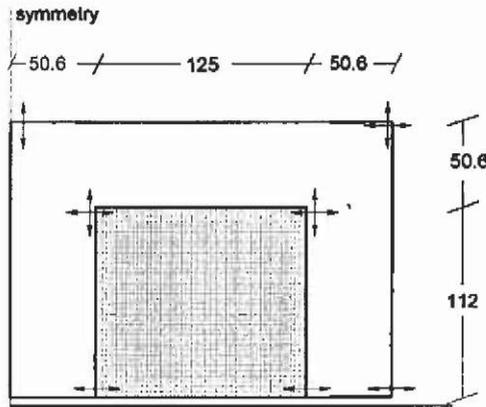
Fig. 10.52. Lifting magnet with initial proportions. Arrows indicate the admissible variation of design parameters.

This method offers a compromise between reliability and performance. A plus-strategy with a smaller number of children and parents would result in faster convergence. The initial and adaptive generated final mesh for the FEM calculation can be seen in Fig. 10.53.
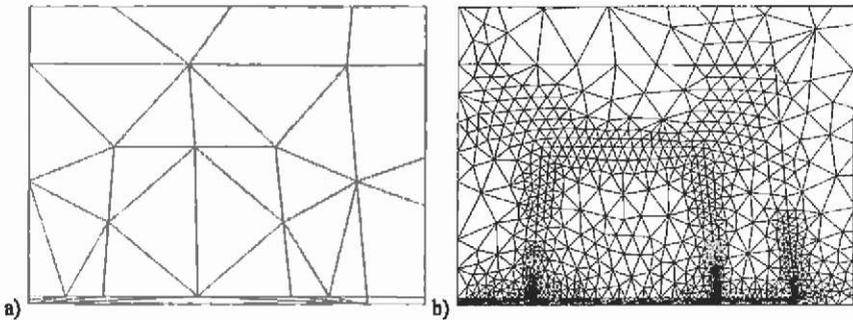
Fig. 10.53. a) Initial and b) adaptive generated final mesh for the optimised geometry.

The resulting field plot of the optimised lifting magnet is shown in Fig. 10.54. The optimisation resulted in a reduction of weight of approximately 6% in comparison to the initial geometry.

The dependence of step length on the iteration steps is illustrated in Fig. 10.55. It serves as convergence and stopping criterion. About 40 iterations, each involving 12 objective function evaluations, seem to be sufficient for an geometrical accuracy of 1 mm. Fig. 10.55 shows that the most significant reduction in weight is achieved during the first 30 iterations. The improvement in the following iterations is less than 1%.
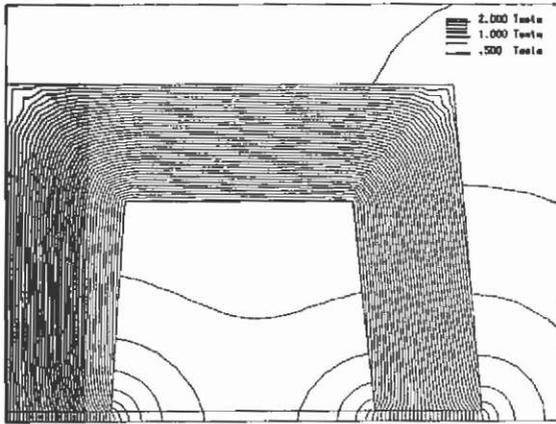
Fig. 10.54. Field distribution of the optimised lifting magnet.
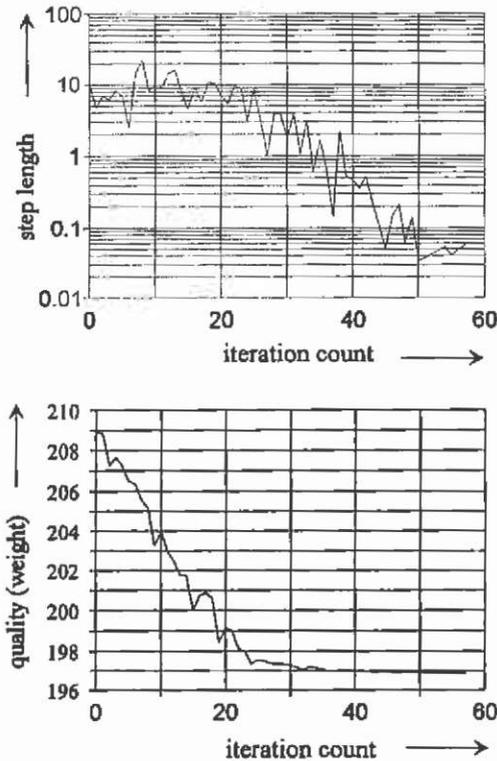
Fig. 10.55. Step length and weight of the lifting magnet versus iteration count
during optimisation.

References

1   Akin, J.E., *Finite elements for analysis and design*, Academic press, London, 1994.

2   Alotto, P., Kuntsevitch, A.V., Mangele, Ch., Molinari, G., Paul, C., Preis, K., Repetto, M., Richter, K.R., Multiobjective Optimization in Magnetostatics: A Proposal for a Benchmark Problem, *IEEE Transactions on Magnetics*, vol. 32, no. 3, pp. 1238-1241, 1996.

3   Alotto, P., Caiti, A., Molinari, G., Repetto, M., A multiquadrics-based algorithm for the acceleration of simulated annealing optimization procedures, *IEEE Transactions on Magnetics*, vol. 32, no. 3, pp. 1198-1201, 1996.

4   Alotto, P., Gaggero, M., Molinari, G., Nervi, M., A Design of Experiment and Statistical Approach to Enhance the Generalised Response Surface Method in the Optimization of Multiminima Problems, *IEEE Transactions on Magnetics*, vol. 33, no. 2, pp. 1896-1899, 1997.

5   Bäck, T., Hammel, U., Schwefel, H.-P., Evolutionary Computation: Comments on the History and Current State, *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, pp. 3-17, April 1997.

6   Babuska, I., Aziz, A.K., On the angle condition in the finite element method, *SIAM Journal on numerical analysis*, pp. 214-226, 1976.

7   Balachandran, M., *Knowledge-based optimum design*, Computational Mechanics Publications, Southampton, 1996.

8   Barrett, R., Berry, M., Chan, T., Demmel, J., Donato, J., Dongarra, J., Eijkhout, V., Pozo, R., Romine, Ch., van der Vorst, H.,

Templates for the solution of linear systems: Building blocks for iterative methods, SIAM, Philadelphia, templates@cs.utk.edu, 1994.

9   Bedrosian, G., A New Method for Coupling Finite Element Field Solutions with External Circuits and Kinematics, *IEEE Trans. on Magnetics*, vol. 29, pp. 1664-1668.

10   Bertsekas, D.P., *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, 1982.

11   Beyer, H.G., Toward a Theory of Evolution Strategies: On the Benefits of Sex - the $(\mu/\mu, \lambda)$ Theory, *Evolutionary Computation*, vol. 3, no. 1, pp. 81-111, 1995.

12   Beyer, H.G., Toward a Theory of Evolution Strategies: Self Adaption, *Evolutionary Computation*, vol. 3, no. 3, pp. 311-347, 1995.

13   Binns, K.J., Lawrenson, P.J., Trowbridge, C.W., *The analytical and numerical solution of electric and magnetic fields*, John Wiley & Sons, 1995.

14   Bischof, C., Roh, L., Mauer-Oats, A., ADIC: An Extensible Automatic Differentiation Tool for ANSI-C, Argonne Preprint ANL/MCS-P626-1196, revised May 1997.

15   Blakely, R.J., *Potential theory in gravity and magnetic applications*, Cambridge University Press, Cambridge, 1996.

16   Bossavit, A., *Electromagnétisme, en vue de la modélisation*, Springer-Verlag, Berlin, 1991.

17   Box, G.E.P., Draper, N.R., *Empirical Model-Building and Response Surfaces*, John Wiley & Sons, New York, 1987.

18   Brebbia, C.A., Dominguez, J., *Boundary elements, An Introductory Course*, Computational Mechanics Publications, Southampton, 1992.

19   Brandiski, K., Pahner, U., Belmans, R., Comparative analysis of three classes of experiment design applied to optimization of PM DC machines, *International Workshop on Electric and Magnetic Fields*, Leuven, Belgium, pp. 79-82, 1994.

20   Brandiski, K., Pahner, U., Belmans, R., Optimal design of a segmental PM DC motor using statistical experiment design

method in combination with numerical field analysis, *ICEM 1994*, Paris, France, vol.3, pp. 210-215, September 5-8, 1994.

21 Brebbia, C.A., *Boundary element techniques: theory and applications in engineering*, Springer Verlag, Berlin, 1984.

22 Breitschuh, U., Jurisch, R., *Die Finite-Element-Methode*, Akademie Verlag, 1993.

23 Briggs, W., *A multigrid tutorial*, S.I.A.M., Philadelphia, PA, 1987.

24 Chua, L.O., Lin, P.M., *Computer Aided Analysis of Electronic Circuits - Algorithms and Computational Techniques*, New Jersey: Prentice-Hall, 1975.

25 Clarke, G.M., Kempson, R.E., *Introduction to the design and analysis of experiments*, Arnold, London, U.K., 1997.

26 Coleman, T.F., Jonsson, G.F., The Efficient Computation of Structured Gradients using Automatic Differentiation, *Cornell Theory Center, Technical Report*, CTC97TR272, 1997.

27 Comini, G., Del Giudice, St., Nonino, C., *Finite element analysis in heat transfer, Basic formulation and linear problems*, Taylor & Francis, 1994.

28 Cook, R.D., *Finite element modelling for stress analysis*, John Wiley & Sons, Inc., New York, 1994.

29 Dappen, S., *Numerische Verfahren zur Berechnung und Optimierung von Induktionstiegelöfen*, PhD-thesis, RWTH Aachen, IEM, Verlag Shaker, Germany, 1996.

30 De Weerdt, R., Hameyer, K., Belmans, R., Dynamic analysis of non-sinusoidally supplied induction machine based on the finite element method, *Proc. conf. 6th European Conference on Power electronics and applications*, Sevilla, Spain, Sept. 19-21, pp. 3.574-3.579, 1995.

31 De Weerdt, R., Hameyer, K., Belmans, R., End-winding leakage calculation of a squirrel-cage induction motor for different load conditions, *Int. Journal for Computation and Mathematics in Electrical & Electronic Engineering*, vol. 14, no. 4, pp. 85-88, 1995.

32 Director, S.W., Rohrer, B.A., The generalized adjoint network and network sensitivities, *IEEE Transactions on Circuit Theory*, pp. 318-322, 1969.

33   Dobmann, M., Liepelt, M., Schittkowski, K., Algorithm 746: PCOMP: A Fortran Code for Automatic Differentiation, *ACM Transactions on Mathematical Software*, vol. 21, no. 3, pp. 233-266, September 1995.

34   Draper, N.R., Small composite design, *Technometrics*, vol. 27, no.2, pp. 173-180, 1985.

35   Dreher, T., Meunier, G., 3D Modeling of Electromagnets Fed by Alternating Voltage Sources, *IEEE Trans. on Magnetics*, vol. 29, pp. 1341-1344, 1993.

36   Dular, P., Meys, B., Remacle, J.F., Henrotte, F., Hameyer, K., W.Legros, Error estimation and adaptive meshing using a dual approach for the study of a microwave cavity, *3rd International Workshop on Electric & Magnetic Fields*, Liège, Belgium, May 6-9, pp. 495-500, 1996.

37   Edminister, J.A., *Theory and Problems of Electromagnetics*, 2nd edition, Schaum's Outline Series, McGraw-Hill Inc., 1993.

38   Eriksson, K., Estep, D., Hansbo, P., Johnson, C., *Computational differential equations*, Cambridge University Press, Cambridge, 1996.

39   Fischer, J., *Elektrodynamik*, Springer Verlag, 1976.

40   Fletcher, R., *Practical Methods of Optimization, Volume 2 - Constrained Optimization*, John Wiley & Sons, Chichester, 1981.

41   Geman, S., Geman, D., Stochastic relaxation, Gibbs distribution and the Bayesian restoration in images, *IEEE Transactions on Patt. Anal. Mach. Int.*, vol. 6, no. 6, pp. 721-741, 1984.

42   George, J.P., Borouchaki, H., *Delaunay triangulation and meshing, Application to finite elements*, Hermes, 1998.

43   George, A., Liu, J., *Computer Solution of Large Sparse Positive Definite Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1981.

44   Goering, H., Roos, H.-G., Tobiska, L., *Finite-Element Methode*, Akademie Verlag 1993.

45   Gyselinck, J. and Melkebeek, J., Numerical methods for time stepping coupled field-circuit systems, *Proc. conf. ELECTRIMACS' 96*, pp. 227-232, 1996.

46 Hackbusch, W., *Theorie und Numerik elliptischer Differentialgleichungen*, Stuttgart 1986.

47 Hafner, Ch., *Numerische Berechnung elektromagnetischer Felder*, Springer Verlag, 1987.

48 Hajek, B., Cooling schedules for optimal annealing, *Mathematics of Operational Research*, vol. 13, no. 4, pp. 563-571, 1988.

49 Hameyer, K., *Beitrag zum automatischen optimalen Entwurf von elektromechanischen Wandlern am Beispiel eines Stellantriebs*, Verlag Shaker, Reihe Elektrotechnik, 1992.

50 Hameyer, K., Kasper, M., Shape optimization of a fractional horsepower dc-motor by stochastic methods, ed. in: *Computer Aided Optimum Design of Structures III, Optimization of Structural Systems and Applications*, CMP, Elsevier Applied Science, London, New York, pp. 15-30, 1993.

51 Hameyer, K., Hanitsch, R., Numerical optimization of the electromagnetic field by stochastic search and MEC-model, *IEEE Trans. on Magnetics*, vol. 30, No. 5, pp. 3431-3434, September 1994.

52 Hameyer, K., Belmans, R., Computation of the electric and magnetic field below high-voltage lines, *8th International Symposium on theoretical electrical engineering (ISTET)*, Thessaloniki, Greece, Sept. 22-23, pp. 278-281, 1995.

53 Hameyer, K. , Mertens, R., Pahner, U., Belmans, R, New technique to enhance the accuracy of 2-D/3-D field quantities and forces obtained by standard finite-element solutions, *IEE Proc.- Sci. Meas. Technol.*, vol. 145, no. 2, pp. 67-75, March 1998.

54 Hanitsch, R., Park, C.-S., Hameyer, K., Novel disc type motors, *Journal of Magnetism and Magnetic Materials*, no. 80, pp. 131-134, 1989.

55 Heath, M., *Scientific Computing, an introductory survey*, McGraw-Hill, New York, NY, 1997.

56 Huebner, K.H., Thornton, E.A., Byrom, T.G., *The finite element method for engineers*, John Wiley & Sons, 1995.

57 Ingber, L., Simulated annealing: Practice versus theory, *Mathematical and Computer Modelling*, vol. 18, no. 11, pp. 29-57, 1993.

58   Iványi, A., *Hysteresis models in electromagnetic computations*, Akadémiai Kiadó, Budapest, 1997.

59   Johansson, T.B., Hameyer, K., Van Dessel, M., Belmans, R., 3D Optimization of average torque of electrostatic micromotors, *Microsystem Technologies*, vol. 1, no. 2, March 1995.

60   Johnson, C., *Numerical solution of partial differential equations by the finite element method*, Cambridge University Press, Cambridge, 1987.

61   Kasper, M., Die Optimierung elektromagnetischer Felder mit Hilfe der Finiten Element Methode und deren Anwendung auf ein Wirbelstromproblem, *VDI Verlag, Reihe Elektrotechnik*, No. 70, 1990.

62   Kasper, M., Hameyer, K., Kost, A., Automated optimal design of a permanent magnet DC motor using global evolution strategy and FEM, *Int. Journal of Applied Electromagnetics & Mechanics*, no. 6, pp. 367-376, 1995.

63   Kellog, O.D., *Foundations of potential theory*, Dover Publications, 1954.

64   Kosmol, P., *Methoden zur numerischen Behandlung von nichtlinearen Gleichungen und Optimierungsaufgaben*, Teubner Studienskripten, Mathematik, Stuttgart 1998.

65   Kost, A., *Numerische Methoden in der Berechnung elektromagnetischer Felder*, Springer Verlag, 1994.

66   Krawczyk, A, Tegopoulos, J.A., *Numerical modelling of eddy currents*, Oxford Science Publisher, 1993.

67   Lin, P.M., *Symbolic Network Analysis*, Elsevier, Amsterdam, 1991.

68   Liu, Z.J., Howe, D., Mellor, P.H., Jenkins, M.K., Coupled thermal and electromagnetic analysis of a permanent magnet brushless DC servo motor, *Proc. conf. 6th IEE International Conference on Electrical Machines and Drives*, Oxford, U.K., pp. 631-635, 1993.

69   Liwschitz-Garik, *Alternating Current Machines*, D. Van Nostrand Company, Ltd., 1946.

70   Lombard, P. and Meunier, G., A General Purpose Method for Electric and Magnetic Combined Problems for 2D, Axisymmetric

and Transient Systems, *IEEE Trans. on Magnetics*, vol. 29, pp. 1737-1740, 1993.

71   Lowther, D.A., Silvester, P.P., *Computer-aided design in magnetics*, Springer Verlag, 1986.

72   Lowther, D.A., Rong, R., Forghani, B., Field Smoothing and Adaptive Mesh Generation, *IEEE Transactions on Magnetics*, vol. 29, no. 2, pp 1890-1893, March 1993.

73   Marsal, D., *Finite Differenzen und Elemente*, Numerische Lösungen von Variationsproblemen und partiellen Differentialgleichungen, Springer Verlag, Berlin, 1989.

74   Maxwell, J.C., *A treatise on electricity and magnetism*, vol. I&II, Clarendon Press, 1891.

75   Mayergoyz, I.D., *Hysteresis - Mathematical models*, Springer New York, New York, 1991.

76   Mecrow, B.C., Jack, A.G., Masterman, J.M., Determination of rotor eddy current losses in permanent magnet machines, *Proc. conf. of the 6th IEE International Conference on Electrical Machines and Drives*, Oxford, UK, pp. 299-304, 1993.

77   Mertens, R., De Gersem, H., Belmans, R., Vandewalle, S., Lahaye, D., Roose, D., Hameyer K., An algebraic multigrid method for solving very large electromagnetic systems, *IEEE Transactions on magnetics*, vol. 34, no. 5, pp. 3327-3330, September 1998.

78   Mertens, R., Pahner, U., Hameyer, K., Belmans, R., De Weerdt, R., Force calculation based on a local solution of Laplace's equation, *IEEE-Transactions on Magnetics*, part II, vol. 33, no. 2, pp. 1216-1218, 1997.

79   Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., Teller, E., Equation of state calculations by fast computing machines, *Journal of Chem. Physics*, 21, pp. 1087-1092, 1953.

80   Mohammed, O.A., Optimal Design of Magnetic Circuits by Finite Elements and Dynamic Search Procedure, *COMPEL -The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, vol. 9, suppl. A, pp. 106-110.

81   Montgomery, D.C., *Design and analysis of experiments*, 4th edition, John Wiley & Sons, New York, 1996.

82   Muramatsu, K., Nakata, T., Takahashi, N., Fujiwara, K., Ukita, H., Investigation of demagnetizing factor for permanent magnet, *Proc. 3rd International Workshop on Electric and Magnetic Fields*, Liege, Belgium, pp. 429-434, May 6-9, 1996.

83   Øksendal, B., *Stochastic differential equations*, Springer Verlag, 1995.

84   Pahner, U., *A general design tool for the numerical optimisation of electromagnetic energy transducers*, Ph.D. thesis, K.U. Leuven, 1998.

85   Palko, S., *Structural Optimisation of an Induction Motor using a Genetic Algorithm and a Finite Element Method.* Acta Polytechnica Scandinavica No. 84, PhD-thesis, Helsinki University of Technology, 1996.

86   Park, I.H., Lee, H.B., Hahn, S.Y., Pole Shape Optimization for Reduction of Cogging Torque by Sensitivity Analysis, *COMPEL-The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, vol. 9, suppl. A, pp. 111-114.

87   Park, I.H., Lee, H.B., Kwak, I.G., Hahn, S.Y., Design Sensitivity Analysis for Steady State Eddy Current Problems by Continuum Approach, *IEEE Transactions on Magnetics*, vol. 30, no. 5, pp. 3411-3414. September 1994.

88   Plejic, M., Hribernik, B., Parameter identification for FEM thermal models of electrical machines, *Proc. conf. ICEM'96*, vol. 1, pp. 282-285, 1996.

89   Preis, K., Ziegler, A., Optimal Design of Electromagnetic Devices with Evolution Strategies, *COMPEL - The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, vol. 9, suppl. A, pp. 119-122, 1990.

90   Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T., *Numerical recipes in C, The art of scientific computing*, Cambridge University Press, New York, 2nd edition, 1996.

91   Quarteroni, A., Valli, A., *Numerical approximation of partial differential equations*, Springer Verlag, 1994.

92   Rahman, M.A., Zhou, P., Determination of saturated parameters of PM motors using loading magnetic fields, *IEEE Transactions on Magnetics*, vol. 27, no. 5, pp. 3947-3950, September 1991.

93   Ramirez, J.A. ,Freeman, E.M., Sensitivity Analysis for the Automatic Shape Design of Electromagnetic Devices in 3D Using FEM, *IEEE Transactions on Magnetics*, vol. 33, no. 2, pp. 1856-1859, 1997.

94   Rao, S.S., *Engineering optimisation*, John Wiley & Sons, 1996.

95   Rechenberg, I., *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution.* Verlag Frommann-Holzboog, Stuttgart, Germany, 1973.

96   Rudnev, V.I., Cook, R.L., Loveless, D.L., Black, M.R., *Induction heat treatment, Basic principles, computation, coil construction and design considerations, Steel heat treatment handbook*, Marcel Dekker Inc., 1997.

97   Saad, Y., *Iterative Methods for Sparse Linear Systems*, PWS Publ., Boston, MA, 1996.

98   Sabonnadière, J.-C., Coulomb, J.-L., *Finite Element Methods in CAD: Electrical and Magnetic Fields*, North Oxford Academic Publishers, 1987.

99   Salon, S.J., *Finite element analysis of electrical machines*, Kluwer Academic Publisher, Boston, London, 1995.

100   Schwefel, H.-P., *Numerical Optimization of Computer Models*, John Wiley & Sons, Chichester, New York, 1981.

101   Schwab, A.J., *Begriffswelt der Feldtheorie*, dritte Auflage, Springer Verlag, 1990.

102   Schwarz, H.R., *Methode der finiten Elemente*, Teubner Studienbücher, Mathematik, 1984.

103   Silvester, P.P., Ferrari, R.L., *Finite elements for electrical engineers*, Cambridge, 1996.

104   Sloan, S.W., A Fast Algorithm for constructing Delaunay triangulations in the plane, *Software for Engineering Problems III, Progress in Engineering*, vol. 11, Computational Mechanics Publications, 1993.

105   Sommerfeld, A., *Elektrodynamik*, Verlag Harri Deutsch, 1988.

106   Stumberger, B., Kreca, B., Hribernik, B., Coupled two-axis model of synchronous motor with permanent magnets, *Proc. of the 3rd International Workshop on Electric and Magnetic Fields*, Liege, Belgium, pp. 399-403, May 6-9, 1996.

107   Storn, R., Price, K., Differential Evolution - a simple and efficient adaptive scheme for global optimization over continuous spaces, *Technical Report TR-95-012*, ICSI, http://http.icsi.berkeley.edu ⌐storn/litera.html, 1992

108   Storn, R., Price, K., Minimizing the real functions of the ICEC 96 contest by Differential Evolution, *Proc. of the International Conference on Evolutionary Computation*, Nagoya, Japan, 1996.

109   Suykens, J., Vandewalle, J., Nonconvex Optimization using a Fokker-Planck Learning Machine, *Proc. of the European Conference on Circuit Theory & Design (ECCTD'95)*, pp. 983-986, 1995.

110   Sykulski, J. K.: *Computational Magnetics*, Chapman & Hall, London, 1st edition, 1995.

111   Tsukerman, I.A., Konrad, A., Meunier, G., Sabonnadière, J.C., Coupled Field-Circuit Problems: Trends and Accomplishments, *IEEE Trans. on Magnetics*, vol. 29, pp. 1701-1704, 1993.

112   Tsukerman, I.A., Konrad, A., Bedrosian, G., Chari, M.V.K., A Survey of Numerical Methods for Transient Eddy Current Problems, *IEEE Trans. on Magnetics*, vol. 29, pp. 1711-1716, 1993.

113   Tsukerman, I.A., A Stability Paradox for Time-Stepping Schemes in Coupled Field-Circuit Problems, *IEEE Trans. on Magnetics*, vol. 31, pp. 1857-1860, 1993.

114   Tsukerman, I.A., Plaks, A., Comparison of accuracy criteria for approximation of conservative fields on tetrahedra, *IEEE Trans. on Magnetics*, vol. 34, no. 5, pp. 3252-3255, September 1998.

115   Turner, P.R., *Numerical Analysis*, MacMillan Press Ltd, 1994.

116   Van Dommelen, D., Germay, N., Complementary results of a new method for the calculation of buried cable heating, *Revue E*, vol. VI, no. 12, pp. 347-354.

117 Vandewalle, J., De Man, H. Rabaey, J., The adjoint switched capacitor network and its applications, *Proc. IEEE ISCAS Conference*, pp. 1031-1034, 1980.

118 Van Dessel, M., Belmans, R., Hanitsch, R., Hemead, E.A., Design of a small, axial-flux permanet-magnet motor, *ETEP European Transactions on Electrical Power*, vol. 6, no. 2, pp. 81-90, March/April 1996.

119 Walkhoff, L.-R., *Theoretische und Experimentelle Untersuchungen an einem permanentmagneterregten bürstenlosen Scheibenläufermotor mit Luftspaltwicklungen*, PhD thesis, TU-Berlin, 1998.

120 Wait, R. and Mitchell, A.R., *Finite Element Analysis and Applications*, John Wiley & Sons, Chichester, 1985.

121 Wunsch, G., *Elektromagnetische Felder*, VEB Verlag Technik, 1989.

122 Zienkiewicz, O.C., Taylor, R.L., *The finite element method*, 4th edition, Volume I, Basic formulations and linear problems, McGraw-Hill, London, 1994.

123 Zienkiewicz, O.C., Taylor, R.L., *The finite element method*, 4th edition, Volume II, Solid and fluid mechanics dynamics and non-linearity, McGraw-Hill, London, 1991.

A postscript on software development at Leuven

Without the constant lively enthusiasm for the subjects of field computations that exists within the research group at Katholieke Universiteit Leuven it would have been impossible to collect this amount of knowledge and expertise in such a short time. Here we would like to present the work that is done within our research group in the frame of field computations and simulations. This book can serve as some thanks to such enthusiastic co-workers and partly as a document of their work during their Ph.D. studies in the electrical engineering department at the Katholieke Universiteit in Leuven.

A university is a place of education. Besides some particular people, students are constantly coming to join the group to work scientifically and are leaving the group to use their acquired knowledge in industry or elsewhere.

At the beginning of 1996 the enthusiasts, Herbert, Ronny, Uwe, Johan D., Johan F., Peter, Koen, Geoffrey, Hans and us, started the development of some special FEM solver and software modules to be coupled to commercial packages. This development has now grown into an independent software package for two-dimensional field problems, including coupled field effects, enhanced force computation, and automated numerical optimisation, with its own name OLYMPOS-2D.

The most important solvers for two-dimensional field problems are present in this package, such as static, time-harmonic, in-plane and transient (magnetic, thermal, electric, mechanical strain/stress,...). Today the group manages more than half a million lines of software code. Various developed programs, post-processor tools and graphical interfaces support the single FEM solver. Parameterisation tools, scripting, graphical coil definition modules back up users such as scientific visitors and students. Further software development and scientific studies will focus on the coupling of the effects of the different field types, the automated optimisation of entire electromagnetic systems, investigations in the representation of non-linear materials will of course work towards a three-dimensional package. Therefore, this book can serve as an intermediate report of the scientific field computation activities inside the Electrical Energy group of the Katholieke Universiteit Leuven.

Index